DESIGN AND DEVELOPMENT OF IOT-ENABLED SYSTEM WITH NEAR-INFRARED(NIR) SPECTROSCOPY FOR DETERMINING THE QUALITY OF SUGARCANE JUICE (*Saccharum Officinarum)*

BALINDAN, GERALD CHRISTIAN REY R.
ELALTO, HANNS JASPHER A.
MASANGCAY, PAUL FRANCIS B.
MOJICA, JEROH LEE M.

An undergraduate thesis outline submitted to the faculty of the Department of Computer and Electronics Engineering, College of Engineering and Information Technology, Cavite State University, Indang, Cavite in partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Engineering with Contribution No. _______. Prepared under the supervision of Prof. Bienvenido C. Sarmiento Jr.

INTRODUCTION

As global agriculture continues to shift toward data-driven practices, the ability to accurately monitor crop quality has become important. Reliable quality assessment not only helps improve harvest decisions but also strengthens the economic stability of farming communities. This is especially true for sugarcane, one of the world's most widely cultivated crops, responsible for about 80% of global sugar production and supporting millions of livelihoods worldwide (FAO, 2020). In the Philippines, sugarcane is a key agricultural commodity, contributing significantly to rural economies and serving as a major source of income in regions like Negros, Batangas and Bukidnon (Briones, 2020).

Modern agricultural practices increasingly depend on fast, accurate and user-friendly tools for quality assessment. However, current technologies for analyzing sugarcane juice are often not accessible, especially for rural farmers who lack the resources or training to operate lab-grade equipment. Although near-infrared

(NIR) spectroscopy has shown strong potential in agriculture, for example, in grain and fruit quality analysis, it remains underused in sugarcane farming, particularly in portable or field-deployable systems (Ciaccheri et al., 2021). This gap in technology highlights the need for a solution that is both reliable and easy to use on-site.

A promising approach is to integrate NIR spectroscopy with Internet of Things (IoT) technology, allowing real-time monitoring and digital access to quality data. Such systems can provide rapid feedback on juice composition, including key metrics like Brix, moisture content and purity, without the need for chemical reagents or laboratory procedures. In the Philippines, national agencies like the Department of Science and Technology (DOST) have called for the development of smart farming tools that can help modernize traditional practices and support small-scale producers (DOST, 2020). These innovations are crucial for addressing long-standing challenges in rural farming communities and improving competitiveness in the sugar industry.

To meet this need, the study aims to design and develop a cost-effective, IoT-enabled NIR spectroscopy device that can accurately evaluate sugarcane juice quality in real-time. By using spectral data and machine learning models, the system will predict essential quality indicators and display the results on a local screen and connected mobile application. This solution will help farmers make informed decisions, reduce reliance on manual testing and promote more efficient sugarcane production. Ultimately, the goal is to create a practical and accessible tool that supports data-driven farming and helps local producers achieve better outcomes.

Statement of the Problem

Traditional methods for sugarcane quality assessment are labor-intensive, time-consuming, and often subjective, leading to inconsistencies in evaluation (Aparatana et al., 2022). This remains a critical factor in the sugar production industry for small-scale farmers, influencing both yield and processing efficiency. Conventional approaches rely on manual sampling and laboratory analysis. Furthermore, the reliance on refractometers creating delays in optimizing harvesting and processing operations (Rudolph Research, n.d.).

Another challenge in sugarcane production is the variability in sucrose content due to environmental factors and crop management practices. Without precise and timely quality assessment inefficiencies in sugar milling and refining affects overall sugar rating. In addition, inaccurate or delayed detection of sugarcane quality variations can result in suboptimal processing conditions which leads to increased production costs and wastage. While refractometers play a crucial role in sugar industries their effectiveness can be influenced by sample preparation and environmental conditions which may introduce inconsistencies in sugar readings (Technology Networks, 2024).

This research aims to address these challenges by developing an IoT-enabled system incorporating NIR spectroscopy for real-time sugarcane quality evaluation. By integrating advanced predictive models such as Partial Least Squares Regression (PLSR) and Support Vector Regression (SVR), this system seeks to enhance accuracy in quality assessment, efficiency, and improving overall sugar production sustainability.

Objectives of the Study

General Objective: To design and develop an IoT-enabled system with Near-Infrared (NIR) spectroscopy for determining the quality of sugarcane juice (Saccharum officinarum).

Specifically, the study seeks to:

1. Implement the system through:
    a. Designing and developing a system for obtaining parameters such as Brix, Pol, and moisture content using NIR spectroscopy
    b. Conversion of the parameters for the development of a classification system for determining the LKgTC, sucrose content, maturity and purity
    c. Incorporating machine learning to improve the efficiency of sugarcane quality assessment.
    d. Integrating hardware and IoT software to enable real-time data acquisition and transmission for remote monitoring
2. Test the system through:
    a. Conducting functional testing using unit testing, integration testing, system testing and acceptance testing.
    b. Conducting a non-functional testing using performance testing.
3. Evaluate the system through:
    a. Functional suitability, performance efficiency, interaction capability, reliability, and flexibility.
    b. Comparing the traditional methods through refractometer for Brix and polarimeter for Pol in terms of execution time, accuracy/precision, throughput, and repeatability.

c. Comparing the laboratory testing for LKgTC, sucrose content, maturity and purity through execution time, accuracy/precision, throughput, and repeatability.

4. Determine the cost of the system.

Significance of the Study

By creating an IOT system with incorporation of NIR spectroscopy and Machine Learning algorithms for determination of sugarcane quality can have immense potential to provide immense benefits for all the different sectors and stakeholders involving sugarcane and related industries such as:

Sugarcane Farm Owners and Sugarcane Wine (Basi) Producers - This study enables sugarcane farm owners and sugarcane wine brewers to have an affordable method to record LKg/TC, maturity, and purity in real-time to reduce errors alongside profit estimation for improved data gathering and decision-making. By having accurate sucrose level monitoring, producers ensure the consistency of their products and improve their production efficiency. (SDG 1: No Poverty, SDG 2: Zero Hunger, SDG 8: Decent Work and Economic Growth)

Sugarcane Industry and Researchers - This study encourages standardization, and acceptance of new technology which amplifies the productivity and competitiveness of different sectors related to sugarcane farming and production. By giving real-time data on various main sugarcane quality parameters such as LKg/TC, pol, and brix, the study provides essential function to researchers for analysis of new sugarcane farming technologies and sugarcane varieties. (SDG 9: Industry, Innovation, and Infrastructure, SDG 12: Responsible Consumption and Production)

Technology Developers and Future Researchers - This study elevates advancement in various fields of technology such as Internet of Things, NIR Spectroscopy, and Automation for manufacturing and agricultural industries. These technological advancements provide a step-forward information to future researchers regarding IOT-based agricultural studies. (SDG 9: Industry, Innovation, and Infrastructure, SDG 17: Partnerships for the Goals)

Time and Place of the Study

The development of the system will begin in June 2025, with fabrication starting in June 2025 and completing by July 2025. The data will be collected within two weeks after fabrication. The system testing and evaluation will take place at Magallanes, Cavite from July 2025 to August 2025. The time frame of the study is from June 2025 until August 2025.

Scope and Limitation of the Study

This study aims to design and develop an IoT-enabled system with Near-Infrared (NIR) spectroscopy for determining the quality of sugarcane for local sugarcane farmers. The system will incorporate machine learning techniques to enhance the efficiency and accuracy of sugarcane quality assessment. The system will also predict the weekly sugarcane bidding price based on the data from the Sugar Regulatory Administration (SRA) to support informed pricing decisions.The system will integrate NIR spectroscopy that measures critical sugarcane quality parameters such as Brix and Pol. These parameters are then used to compute metrics like LKg/TC, sucrose content, maturity, and purity. The system is equipped with IoT features to allow real-time monitoring of sugarcane data such as the LKgTC, sucrose content, maturity and purity through a mobile application.

The system will not include a juicer for the sugarcane, as only a small amount of sugarcane juice is required for the scanner to work. The system will utilize PLSR as its machine learning algorithm to assess the qualities of the sugarcane, and IoT which will integrate a mobile application for monitoring and recording of the tested parameters of the quality assessment. The system will also not be designed to test the quality parameters of the other parts of the sugarcane such as the skin or the cross-sectional scanning of billets and defibrillated samples. The system is limited to testing the specified parameters and may not provide the same level of precision as industrial-grade analyzers. The sensors and microcontroller are designed for small-scale use and will be mostly for small-scale or local sugarcane farmers to utilize for the monitoring of the sugarcane.

Definition of Terms

Brix (%Brix) – A measurement of the total soluble solids in juice while apparent purity quantifies the percentage of sucrose in the juice's total solids (Verlag, 2019).

IoT (Internet of Things) - Refers to a network of interconnected devices that communicate and exchange data, enabling smart automation and efficient monitoring across various domains (Wang et al., 2019).

Machine Learning (ML) – An AI application that enables systems to automatically learn and improve from experience without explicit programming, encompassing supervised, unsupervised, and reinforcement learning techniques (Alam, 2023).

Moisture Content – The amount of water present in a substance, often expressed as a percentage of the total weight.

Near-Infrared (NIR) Spectroscopy – Analytical technique that utilizes near-infrared radiation to evaluate the composition and characteristics of materials (Pike n.d.).

Partial Least Square Regression (PLSR) - A machine learning regression technique that integrates principal component analysis and multiple regression, ideal for predicting multiple dependent variables from a large set of independent variables (Abdi, 2007).

Sugarcane (Saccharum officinarum) – A perennial grass belonging to the Poaceae family that is primarily cultivated for the extraction of juice used in sugar production (Yamane 2025).

Sugarcane Juice – The extracted liquid from crushed sugarcane stalks, rich in sugars like sucrose, and used as the primary raw material in sugar production (Chen & Chou, 1993).

Pol (Polarity) – A measurement of the mass percentage of sucrose in a solution, determined by measuring the optical rotation of polarized light passing through the sugar solution (Alluri, 2019).

Random Forest Regression (RFR) - A machine learning regression technique that utilizes bagging and random subspace methods to create multiple decision trees, which are then combined to generate the final prediction (N., G., Jain, P., Choudhury, A., Dutta, P., Kalita, K., & Barsocchi, P., 2021).

LKgTC (50-kilogram per ton of cane) – A standard unit of measurement used in the sugar industry to estimate the amount of sugar recoverable from one ton of sugarcane.

Conceptual Framework

This section follows the Input-Process-Output (IPO) model to guide the development of the IoT-based system for analyzing sugarcane juice quality. The model outlines the key elements of the system: the input stage includes sugarcane juice samples, software tools and hardware components; the process stage covers designing the system's circuitry, development and training of the machine learning algorithm, mobile app development, system integration, testing, evaluation and cost computation; and the output is a functional device that uses NIR spectroscopy and machine learning to provide real-time data analysis. This framework ensures a structured approach to building a reliable and efficient solution for farmers and other stakeholders.

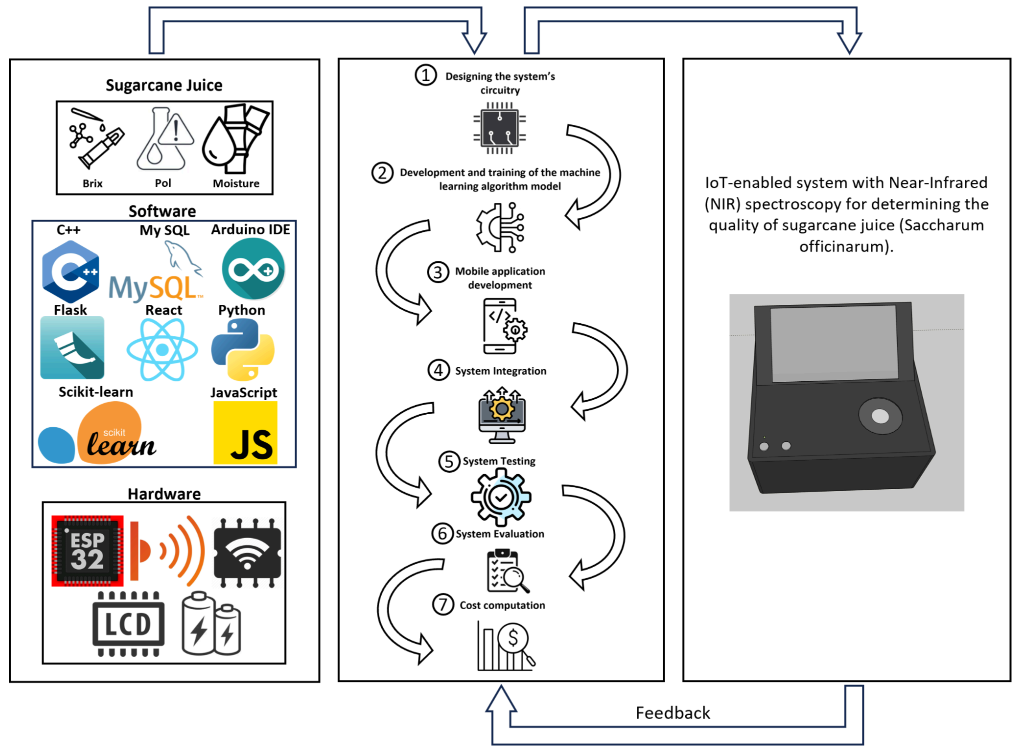


Figure 1. Conceptual framework of the proposed system

The input stage of the system will include the sugarcane juice, software and hardware components for its development. The sugarcane juice samples will provide the data needed to predict the qualities. The programming languages will include C++ for firmware programming, Python for handling the machine learning algorithms, JavaScript and React for the front-end development, and MySQL for database management. The system will also utilize machine learning libraries such as scikit-learn to build the predictive models. On the hardware side, the system will use ESP32 microcontroller, which will act as the main control unit, interfacing with various modules including the NIR spectroscopy sensor for detecting juice composition, and LCD display for visual output, a rechargeable lithium-ion battery for portable power supply and a main switch for user control.

The process phase outlines the major development steps involved in building the IoT-based NIR spectroscopy system for sugarcane juice quality analysis. It begins with designing the system's circuitry, where the electronic components are planned and connected, including the ESP32 microcontroller, NIR sensor, display, power source, and other modules. This stage ensures that all hardware parts are compatible and can function reliably together. Next is the development and training of the machine learning algorithm model. Spectral data collected from sugarcane juice samples is used to train a model using techniques such as Partial Least Squares Regression (PLSR). This model will be responsible for predicting key juice quality parameters like Brix and Pol from the raw NIR data. Following the model development, a mobile application will be created to provide a user-friendly interface for farmers or stakeholders. The app will display real-time results and allow users to access historical data, helping them make informed decisions about harvest timing. After these individual components are completed, system integration takes place. This step combines the hardware, firmware, machine learning model and mobile application into one fully functional unit. Once integrated, system testing is performed to ensure that all parts work together correctly and consistently. This includes checking the accuracy of predictions, communication between devices and overall system stability. Finally, the system will undergo evaluation to assess its performance in real-world conditions and identify areas for improvement. A cost computation will also be conducted to determine the financial feasibility and affordability of deploying the device at scale for local farmers.

The primary output is a fully functioning IoT-enabled device that uses NIR spectroscopy to determine the quality of sugarcane juice in real-time. The system is capable of measuring important quality indicators such as Brix, Pol, moisture content and purity without the need for chemical reagents or time-consuming laboratory procedures. Through the integration of the PLSR model, the system will generate reliable and accurate predictions of juice quality from raw spectral data. The results will be presented to the user through a user-friendly display interface, and may also be accessed via mobile application. This will allow the farmers and stakeholders to make informed decisions about the optimal timing for harvest, based on objective quality data rather than traditional guesswork.

The development of this IoT-enabled NIR spectroscopy system will demonstrate the potential of integrating modern sensors, machine learning and embedded systems in agriculture. By providing an accurate method to assess sugarcane juice quality parameters, the system will empower sugarcane farm owners to make data-informed decisions that can lead to better yield quality, higher profit margins and reduced post-harvest losses. The inclusion of cost analysis ensures that the system remains accessible and practical. This technology supports the broader goals of smart agriculture by enhancing productivity, sustainability and precision in crop management.

REVIEW OF RELATED LITERATURE

This chapter will provide a comprehensive summary of the related studies and literature, gathered through extensive research conducted by the researchers. The information presented in this chapter is essential to understand the context and significance of the proposed study.

Importance of brix, pol, moisture content for sugarcane farming

Brix is a widely used parameter in the sugarcane industry that indicates the percentage of soluble solids in a liquid solution. According to Grimaldo (2023), Brix is essential for estimating sugar content in sugarcane juice, which directly impacts extraction efficiency and overall product quality. Monitoring Brix levels ensures that sugarcane is harvested at optimal maturity, maximizing yield and enhancing profitability. This makes Brix a crucial parameter in evaluating and maintaining the quality of sugarcane juice in both industrial and agricultural settings.

Polarization (Pol) is a critical parameter in assessing the quality of sugarcane juice, as it directly correlates with sucrose content, which is the primary sugar present in the juice. The study of Xiao, et al.(2017) has shown a strong linear relationship between Pol and sucrose levels, with high coefficients of determination indicating a robust predictive capacity. Moreover, Pol measurements are essential for monitoring and controlling sugarcane juice quality during processing. By analyzing Pol alongside other quality indices like brix (Bx) and apparent purity (Ap), it is possible to gain comprehensive insights into juice quality, facilitating optimization of processing conditions and ensuring consistent product quality.

Another important parameter for the assessment of sugarcane is the moisture content. A study by Mequanent and Ayele (2014) proved the importance of moisture content as an indicator of sugarcane maturity and sugar quality. As sugarcane ripens, the moisture content in the leaves decreases, making it a useful parameter for determining the optimal time to harvest. The researchers found that lower moisture levels were linked to higher sugar concentration and better juice purity. Additionally, they observed that sugarcane harvested later, when moisture levels had decreased, produced sweeter juice. These findings highlight the role of moisture content in influencing sugar recovery and overall cane quality. Monitoring this parameter can help farmers make better harvesting decisions and improve the economic return from their crops.

Role of LKGTC, sucrose content, maturity, purity in sugarcane farming

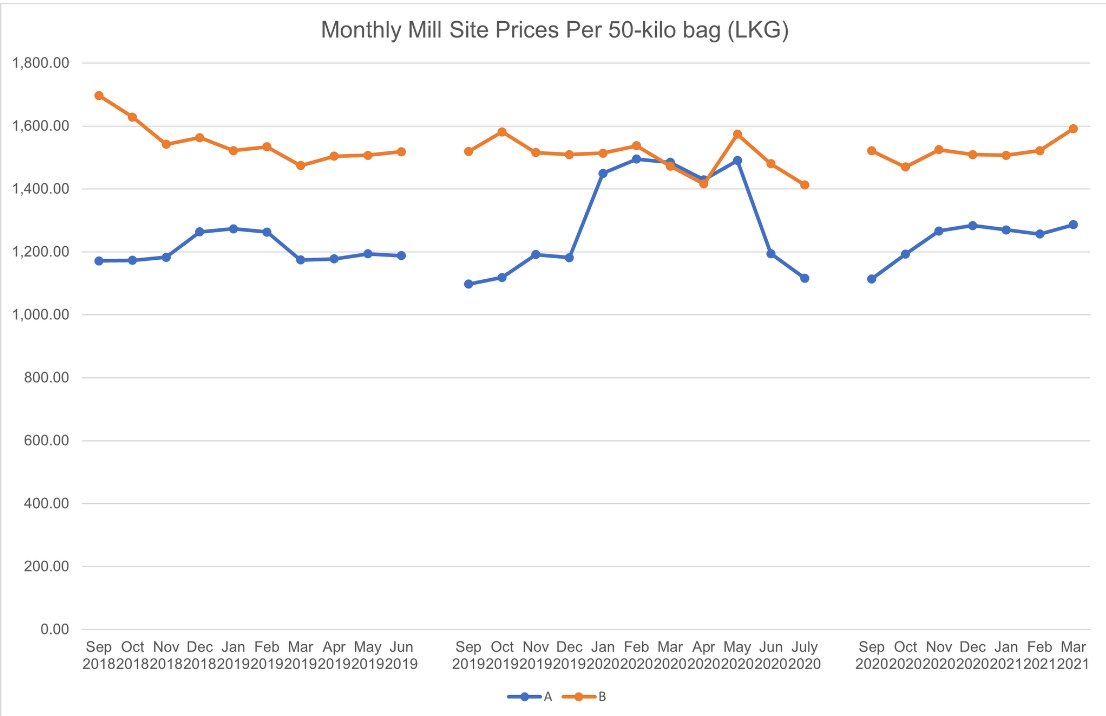
LKgTC is an important factor in sugarcane farming as it stands for having 50-kilogram bags of sugar per ton of cane wherein it measures the efficiency of sugarcane recovery or how much sugar can be extracted for a ton of sugarcane and United Sugar Producers Federation of the Philippines (UNIFED) recorded for the year of 2024 a sugarcane extraction of 1.44 LKGTC during the start of the current milling season, against an average of 1.7 LKGTC (Business World, 2024).

According to the Sugar Regulatory Administration (SRA), the monthly average sugar prices in Metro Manila for the crop year 2022–2023 experienced a significant occurrence of unexpected changes.Table 1. Monthly Average Sugar Prices in Metro Manila Crop Year 2023-2024

Price is an important consideration that can drive a farmer's decision to plant sugarcane. When the price of sugar drops relative to other options, farmers can shift to more profitable crops. High sugar prices are a good signal for farmers to plant sugarcane, with the projected income computed in terms of mill site prices using sugar yield (LKG/TC) and the prevailing sharing scheme implemented in the mill district. According to a report of the United States Department of Agriculture (2021), the sugar industry proved resilient during the COVID-19 pandemic, with prices within the P1,500 ($30.9) per LKG bag level, which industry views as profitable to producers and fair to consumers. The mill site composite price of raw sugar in CY 2020/2021 through March ranged from P1,492.70 ($30.77) to P1,570.16 ($32.37) per LKG bag and had a national average of P1,501.04 ($30.94) per LKG bag.

The average mill site prices of U.S. quota sugar and domestic prices midway through CY 2020/2021 were PhP1,238.7 ($25.54) and P1,520.79 ($31.36) per LKG bag, respectively. For the past three years, the prices of U.S. quota sugar have been significantly lower than domestic prices. SRA justifies the “A” quedan allocation due to the U.S. quota’s utility in the event of a surplus and considering U.S. prices being higher than the prevailing world prices (see table 1). Lower sugar demand during the COVID-19 pandemic resulted in lower prices compared to the pre-pandemic CY 2018/2019.

Table 1. Monthly mill site prices based on LKgTC (Sevilla, 2021)



The sucrose content or termed as brix percentage or sugar content where one degree Brix is equal to one gram of sucrose in 100 gram of solution which does simply mean that for example, 20% Brix is equivalent to 20% sucrose and a higher sucrose content in the base sugarcane juice of the harvested sugarcane leads to a higher yield of sugar and sugarcane with brix % closer to 23% is considered to produce the highest cane sugar quality (Hanna Instruments, 2016)

Understanding the maturity of sugarcane is critical in sugarcane farming as it directly influences yield, juice quality, and sugar recovery. Harvesting at the optimal stage of maturity ensures the maximization of millable cane weight and sugar content, while minimizing field losses. Conversely, premature or overmature harvesting can result in reduced cane yield, lower sugar recovery, compromised juice quality, and increased milling challenges due to the presence of extraneous matter. Efficient sugarcane harvesting practices, including accurate determination of maturity and the adoption of proper techniques, are thus essential for achieving sustainable and economically viable production (Agritech, 2020).

Apparent purity is a vital parameter for sugarcane farmers, as it directly influences sugar extraction efficiency and the quality of the final product. This is calculated as the polarization percentage divided by corrected Brix which provides insights into the proportion of sucrose in the juice relative to impurities. Higher apparent purity signifies improved sugar recovery during processing, which enhances productivity and reduces waste. Additionally, assessing apparent purity helps identify mineral and organic impurities introduced during harvesting, enabling farmers to adopt practices that improve crop health and juice quality. By understanding this metric, farmers can evaluate and refine their cultivation techniques, negotiate better prices for high-purity produce, and contribute to sustainable sugar production by reducing environmental impact. Apparent purity, therefore, serves as both an economic and ecological tool for optimizing sugarcane farming and aligning production with industry standards (Abdusalam & Afolayan, 2022).

NIR application for obtaining brix, pol, and moisture content on Sugarcane Juice

According to Corred, et. al. (2024), Near-Infrared (NIR) spectroscopy is a well-established technique to monitor the quality of raw sugarcane received by sugar mills, and consequently, for pricing and trading with producers and growers. NIR spectroscopy is being utilized in detecting the Brix level of the sugarcane by sugar mills and manufacturers to detect sugarcane quality. NIR is favored because of its speed of operation, ranging from 0.2 - 1 minute only, rather than other methods of determining Brix such as refractometry, chromatography or polarimeter which ranges from 20 - 30 minutes before it can determine the Brix values of the sugarcane.

For the application of NIR to sugarcane, a study of Corredo, et. al.(2024) shows the vis-NIR (visible - near infrared) raw spectral data obtained for 302 samples of each sugarcane sample type are shown in Figure. The study observed a noisy aspect in the region corresponding to the visible spectrum (400 to 698 nm), Thus, only bands in the spectral range between 699 and 1010 nm and between 1070 and 2153 nm (303 spectral bands) were retained.

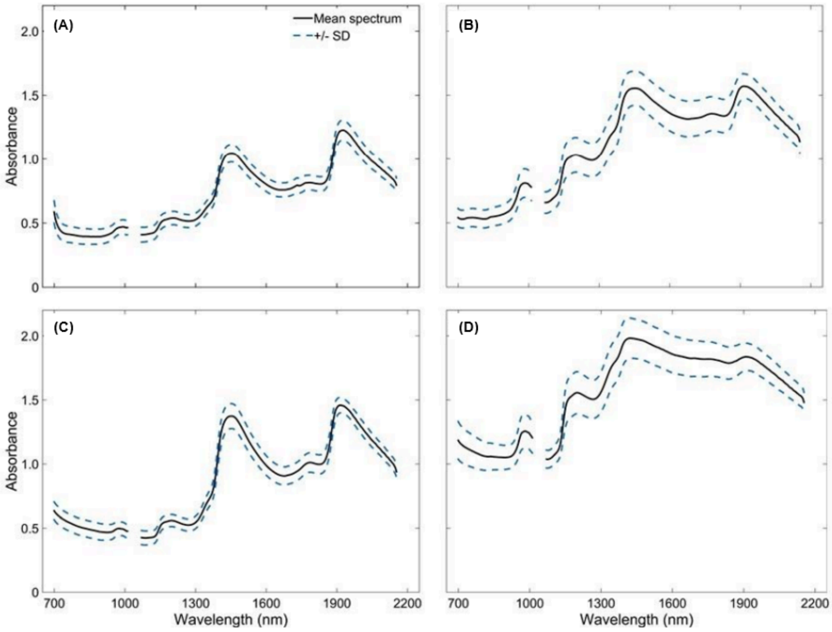


Figure 2. vis-NIR mean spectra and standard deviation (SD) of all 302 sugarcane samples for (A) skins and (B) cross-sectional scanning of billets, (C) defibrated, and (D), raw juice samples (Corredo, et. al., 2024).

The figure above determines the wavelength for vis-NIR of sugarcanes samples; the (d) raw juice samples proves the utilization of NIR in sugarcane juices. Vis-NIR spectroscopy may be used in a number of applications, including the classification of sugarcane varieties, with promising results. The same authors showed that the spectral regions between 650 and 750 nm, corresponding to the visible spectrum, was the most suitable for sugarcane discrimination. The vis-NIR technique principle is based on the detection of compounds and molecules through their molecular vibration states. Different varieties naturally have different concentrations of parameters such as sucrose and fibre according to genetics. Furthermore, for all of them, the plant matrix is essentially composed of water (75–82%), insoluble solids content (Fibre, 10–18%), and soluble solids (Brix, 18–25%), which are composed of nonsugars (1–2%), sucrose (14–24%), and reducing sugars (0–1.5%). However, the prediction of quality parameters related to chemical compounds of interest should be independent of sugarcane varieties.

According to a study of Kuswurjanto and Triantarti (2019), The result obtained shows the potential use of NIRS for determining %brix and %pol of sugarcane juice quality. NIRS have a promising result to replace conventional analysis methods. The performance of the prediction set is presented by the scatter plots in figure 3 and 4, respectively.

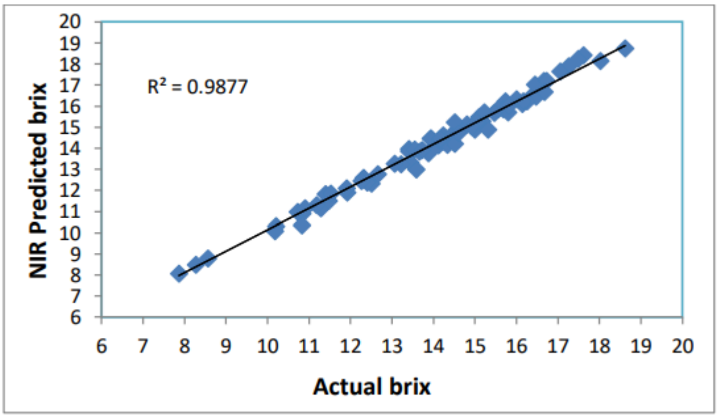


Figure 3. Actual vs NIR predicted values for brix of sugarcane juice (R Kuswurjanto and Triantarti, 2019)

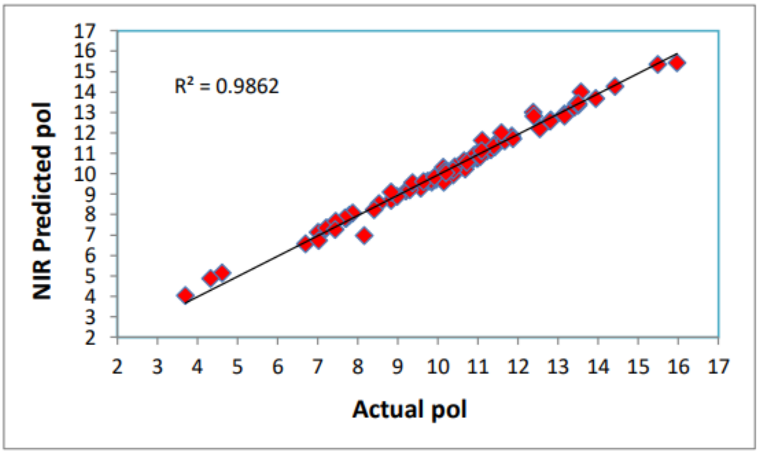


Figure 4. Actual vs NIR predicted values for pol of sugarcane juice (R Kuswurjanto and Triantarti, 2019)

The performances of the model were evaluated using 102 independent samples in the prediction set. It shows that the r2 were good for both brix and pol parameters. The result of NIR prediction is consistent with the study

Near-infrared (NIR) spectroscopy has emerged as a promising alternative to traditional analysis methods in the sugarcane industry due to its speed, efficiency, and ability to handle large sample volumes. It demonstrated its effectiveness in various applications, including plant breeding trials for fiber quality assessment, online analysis in sugar factories, and cane payment systems in countries like Australia, Japan, Kenya, and South Africa. NIR technology has shown accurate results in estimating fiber, ash, % brix, and % pol levels, offering a reliable and rapid solution for juice quality analysis (Kuswurjanto and Triantarti, 2019).

Required amount of sugarcane sample

A study by Nawi et.al (2014) predicted the sugarcane juice quality brix and pol for both raw and clarified sugarcane juice by using a 10mm square cuvette with 10mm path length as a container for the sugarcane juice samples. A square type with an internal size of 10mm x 10 mm and a 43.75 mm internal height and a 10mm square cuvette volume can be calculated by using the formula inner length x inner width x inner height x 80% and for every 10mm can accommodate 1mL of sample and by substitution, 1mL x 1mL x 4.375mL x 80% = 3.5 mL (Spectrocology, 2022), therefore 3.5 mL can be used as the standard sample needed for the sugarcane quality assessment wherein spillage of sample is considered.

Machine Learning Algorithms

The application of machine learning (ML) to near-infrared (NIR) is used to assess sugarcane quality and brings about increased efficiency and accuracy to farmers. Measurement of essential quality indicators like Brix, Pol, moisture, and sucrose are generally done manually which is both labor-intensive and time-consuming. However, by using ML techniques on NIR spectroscopy means they will be automated and provide real-time and highly accurate results. ML algorithms can extract the features from the NIR spectral data that will help farmers make better decisions about pricing and harvesting based on the estimated sugarcane quality.

According to a study of Aparatana, et al.(2023), In the first experiment, 213 sugarcane juice spectra measured by the benchtop NIR spectrometer and 213 sugar quality measurements obtained using the conventional method were used to develop six partial least squares (PLS) regression models. PLS and MLR were used to develop multiple regression models. They are powerful statistical methods that can generate a linear regression model by projecting the predicted and observable variables into a new space. Another study from Corredo, et al. (2021), states the examination of the potential of NIRS for predicting and mapping Brix, Pol and Fibre content in a commercial sugarcane field. The quality attributes models were adjusted considering the spectral reflectance from the 1100–1800 nm wavelengths by using partial least squares regressions (PLSR). A total of 350 samples were collected in a sugar mill laboratory for calibration and cross-validation models development.

According to the study of Raafi'udin et al. (2024). The utilization of Random Forest Regressor (RFR) with default parameters was used to develop a feature selection model for Near-Infrared Spectroscopy (NIRS) data. Their study focused on predicting six quality parameters of beef, including color, drip loss, pH, storage time, Total Plate Colony (TPC), and water moisture. This demonstrates the effectiveness of RFR in handling complex spectral data and improving prediction accuracy in NIRS-based quality assessment.

Testing Metrics

Evaluating the performance of predictive models for sugarcane juice quality involves various testing metrics. Commonly used metrics include the coefficient of determination (R^2), root mean square error (RMSE), and standard error of prediction (SEP) (Bao, et al., 2009). These metrics assess the accuracy and reliability of the models in predicting quality parameters. For instance, a study utilizing a portable visible and near-infrared spectrometer reported R^2 values of 0.85 for Pol and 0.84 for Brix, indicating good prediction performance (Cheng, et al., 2014).

Evaluation Metrics

In addition to testing metrics, evaluation metrics such as the ratio of performance to deviation (RPD) and standard error of calibration (SEC) are used to assess the robustness and generalizability of predictive models. High RPD values and low SEC indicate strong model performance. For example, Kuswurjanto and Triantarti (2019) reported RPD values of 5.441 for Brix and 6.125 for Pol, suggesting that their NIR-based models were highly effective in predicting sugarcane juice quality parameters (Kuswurjanto and Triantarti, 2019).

Over-Mature Harvesting of Sugarcane

Over-mature harvesting of sugarcane negatively impacts both the quality of the cane and the economic returns for farmers. When cane is left unharvested from a previous season and becomes over-mature at the start of a new crop season, its sucrose content and overall quality decline. This leads to reduced prices per ton of cane delivered to the mill. Maintaining a proper harvest schedule and monitoring cane maturity is therefore essential to prevent losses in both quality and income (Sugar Industry of Belize, 2017).

According to Briceño, et al. (2018), sucrose losses between harvesting and milling begin soon after cutting, increasing with the time the cane remains in the field or in the mill yards. The deterioration rate depends upon environmental conditions, the cane variety, and the management of the harvesting system.

Therefore, implementing a well-structured harvesting plan that considers the last date of harvest and the maturity of the cane is crucial to ensure optimal sugar recovery and economic returns.

Maturity of cane is a major factor in the inversion and subsequent reduction of stored sucrose. As physiological maturity increases the extent of sucrose loss is enhanced. Harvesting of immature (or) over mature cane should be avoided to cut down sugar losses. In India, the crop is cultivated in almost all the states barring the cold ones, under diverse conditions. In India sugarcane is cultivated in 5.02 M ha,with 342.1 Mt production and productivity of 68.1 t/ha. In Andhra Pradesh area under cultivation is 1.80 M ha and the production is 140.4 Mt with productivity of 78.0 t/ ha (Reddy, 2018).

According to an article from University of Florida IFAS Extension (2023), maturity curves for different sugarcane varieties and their impact on harvest scheduling. It highlights how harvesting at the optimal maturity stage ensures maximum sucrose accumulation, preventing losses from harvesting too early (immature cane) or too late (over-mature cane). The study emphasizes the importance of timely harvesting to maximize yield and profit.

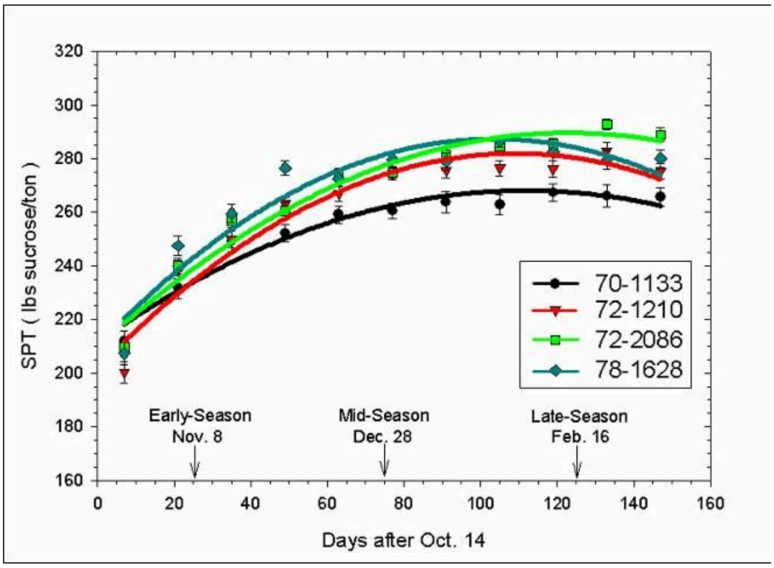


Figure 5. Maturity curve of different sugarcane varieties (Gilbert, et. al., 2023)

Sustainable Development Goals (SDG) covered by the system

The proposed system aims to contribute to the attainment of a number of the Sustainable Development Goals (SDGs). It helps achieve SDG 2: Zero Hunger by offering information on juice quality which lets farmers know when to pick and enables them to make higher harvests and earn more money (World Food Programme, 2025). Supporting sustainable food production and hardy agricultural procedures is also a result.

It implements SDG 8: Decent Work and Economic Growth by improving the agriculture industry's technology which increases farmers' productivity. It ensures that products and processes are of good quality which helps in saving resources and avoiding harm to the environment while the economy develops (United Nations, 2025).

For SDG 9: Industry, Innovation and Infrastructure, this system exemplifies technological innovation applied to agriculture, enhancing resource efficiency and sustainability in sugarcane processing. Its development, involving machine learning and embedded systems, directly boosts scientific research and technological advancement in the sector (GOVPH, 2020).

Regarding SDG 12: Responsible Consumption and Production, the system enables the sustainable management and efficient use of natural resources. By providing data for optimized harvest and reduced post-harvest losses, it ensures more efficient sugarcane utilization and significantly contributes to waste reduction in the overall production cycle (United Nations, 2025).

The project also contributes to SDG 17: Partnerships for the Goals. The development and sharing of this system foster knowledge sharing and technological cooperation for sustainable development. It also encourages effective public, public-private, and civil society partnerships essential for bringing such technology to fruition and widespread deployment (United Nations, 2025).

Comparison Table for Research Gap and Contribution

Table 2. Summary of related studies

Study (Author, Year)	Approach	Evaluation Method	Limitations	Differences with Proposed Study
Hiriotappa et al. (2022)	Applied NIR spectroscopy for rapid analysis of TSS and TS in fresh sugarcane juice.	NIR spectroscopy with PLS regression.	Controlled sample temperature; laboratory-based.	Expands scope by including more quality parameters using ML
Kuswurjanto et al. (2019)	Evaluated sugarcane quality using NIR in core sampler systems for payment purposes.	NIR reflectance spectroscopy; compared with Wet Disintegrator method.	Used stand-alone spectral anaylsis	Integrates NIR IoT sensors for continuous juice quality monitoring
Taira et al.	Non-destructiv	Portable NIR	Limited to	Integrates full

(2013)	e measurement of sugar content in whole stalk sugarcane using portable NIR.	spectrophotometer in transmittance mode.	sugar content quality parameters; no integration with data systems.	quality suite and IoT connectivity
Sanseechan et al. (2018)	Utilized Vis-NIR spectroscopy for sugarcane quality prediction.	Vis-NIR spectroscopy; statistical analysis.	Focused on specific parameters; no real-time monitoring.	Integrates IoT sensors and predictive modeling features.
Rani et al. (2020)	Proposed IoT-based spectroscopic systems for agriculture.	Conceptual framework; NIR spectroscopy.	Lacked practical implementation; theoretical approach.	Proposed system is validated, integrated with ML.
Villanueva et al. (2019)	Studied traditional Brix testing methods in Philippine sugarcane farms.	Manual sampling; refractometry.	Labor-intensive; prone to human error.	Highlighted need for automation and IoT solutions.
Escobar et al. (2021)	Assessed modernization barriers in agriculture.	Surveys; case studies.	Generalized findings; not crop-specific.	Emphasized the importance of IoT in farming modernization.
DOST-PCAARRD (2019)	Reported on technological adoption in Philippine agriculture.	Surveys; technology assessments.	Broad scope; not specific to sugarcane.	Underlined the necessity for accessible tech solutions like IoT devices.
SRA (2023)	Provided data on sugarcane production regions in the Philippines.	Statistical analysis; regional reports.	Data-centric; lacked technological focus.	Offering local context for implementing IoT-based solutions.
Singh et al. (2021)	Explored pricing models based on sugarcane quality.	Economic modeling; quality metrics.	Dependent on accurate quality assessment.	Supported the importance of precise, quality data.

Shao et al. (2021)	Developed IoT systems for precision agriculture.	IoT device integration; data analytics.	General application; not sugarcane-focused	Demonstrates IoT feasibility; proposed study applies similar tech to sugarcane.
Balasubramanian et al. (2022)	Implemented IoT solutions in small-scale farming.	IoT devices; mobile applications.	Limited to specific regions; scalability concerns.	Showed successful IoT adoption that applies in sugarcane farming.
Wang et al. (2020)	Applied NIR spectroscopy in fruit quality analysis.	NIR spectroscopy; machine learning models.	Not applicable to sugarcane.	Validated NIR's effectiveness, supporting its use in sugarcane quality assessment.

Synthesis

The review has focused on the challenges and data parameters involved in sugarcane farming particularly on the importance of real-time quality monitoring and prediction of quality of the sugarcane. Over-mature sugarcane significantly reduces sucrose content and affects both juice quality and economic returns which makes timely harvesting crucial. Raw data parameters like Brix, Pol, and moisture content are essential indicators of sugarcane maturity and sugar recovery potential. Near-Infrared (NIR) spectroscopy has emerged as a rapid and efficient alternative to traditional quality assessment methods that offers accurate real-time data analysis. Additionally, the integration of machine learning algorithms such as Partial Least Squares Regression (PLSR) and Random Forest enhances the prediction accuracy of obtained data from the (NIR) spectroscopy that allows for optimized harvesting and pricing strategies. These technologies address the limitations of traditional analysis methods and contribute to maximizing yield, profitability, and sustainability in sugarcane farming.

METHODOLOGY

This chapter presents the detailed overview of the materials, methods, and data collection procedures that will be utilized for this study. By outlining these components and processes, it will ensure a clear framework of how the study will be conducted.

1. Requirements and Specification

Functional Requirements

NIR-Based Juice Quality Assessment

- The system shall activate the NIR sensor upon pressing Button 2.
- The system shall collect and preprocess spectral data from the sugarcane juice sample.
- The system shall compute quality parameters including Brix, Pol, Moisture Content, Maturity, and Purity using embedded machine learning.
- The system shall display the processed results on a 7-inch TFT LCD screen via I2C.

Scanning and Feedback

- The system shall be powered on using Button 1 and initiate scanning using Button 2.
- The system shall provide status feedback through a green LED with the following behavior:
  - Solid green: system is turned on
  - Fast blinking green: scanning in progress
  - Slow blinking green: charging
  - No light: system is off

Data Logging and IoT Connectivity

- The system shall transmit processed scan results to a cloud-based IoT server via Wi-Fi.
- The system shall associate each scan result with a timestamp for traceability.
- The system shall allow remote access to historical data through a connected mobile or web application.

Mobile App Interface

- The mobile app shall display real-time results sent from the system via Wi-Fi.
- The app shall provide graphical visualization of stored scan data.
- The app shall allow users to review previous scan logs and access system history.

Power Management

- The system shall wake from idle mode through user input or mobile interaction.
- The system shall indicate low battery levels through the mobile app.

Non-functional Require

Functional Suitability

- The system shall accurately measure sugarcane juice parameters such as Brix, Pol, and Purity.
- It shall consistently deliver correct and usable results aligned with standard quality assessments.

Performance Efficiency

- The system shall process scans and display results within 3 seconds of initiation.
- It must maintain fast and reliable performance even during batch or repeated testing sessions.

Interaction Capability

- The system shall feature an interface that is simple and intuitive for farmers and technicians.
- Users shall be able to operate the device and interpret results with minimal training.

Reliability

- The system shall operate continuously without crashes, system freezes, or data loss.
- It shall maintain stable performance during prolonged use in real-world farming conditions.

Flexibility

- The system shall adapt to different sugarcane farm setups without requiring reconfiguration.
- Both hardware and software shall support varying field conditions and user preferences.

2. System Architecture

The system architecture of the IoT-enabled NIR spectroscopy device is illustrated in figure 6. It integrates hardware and software components for analyzing the quality of sugarcane juice using spectral data. The primary components include the ESP32 microcontroller, NIR spectroscopy sensor, 7-inch TFT LCD, mobile application, cloud database, and supporting peripherals like battery, buttons and led indicators. These modules are connected through various communication protocols and operate under a regulated power supply.

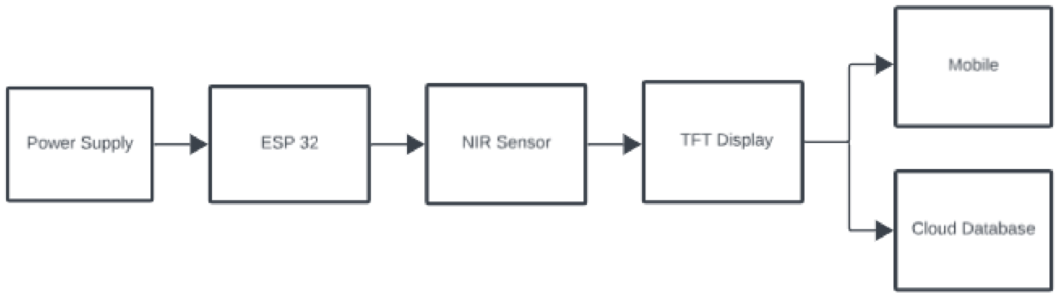


Figure 6. Block diagram of major components

Data Flow Between Modules

The input stage of the system begins when the user presses Button 1, which powers on the device, activating the ESP32 microcontroller, NIR sensor, LCD display, and Wi-Fi module. Once the system is turned on, the green LED lights up solid, indicating that the device is active and ready. The user then presses Button 2 to initiate the scanning process. During this time, the green LED starts blinking rapidly, signaling that the NIR sensor is actively scanning the sugarcane juice sample. The NIR sensor captures the sample's spectral response and sends the raw data via SPI communication to the ESP32 for processing.

In the processing stage, the ESP32 microcontroller receives the spectral data and begins preprocessing it to remove noise and normalize the signal. The cleaned data is then passed through an embedded machine learning model trained to analyze and classify the juice sample. The model evaluates key quality parameters such as Brix, Pol, Moisture Content, Maturity, and Purity based on the scanned data. The ESP32 performs all computation locally, enabling fast and offline analysis. If the device is charging during this phase, the green LED blinks slowly, indicating that the system is receiving power while operating.

Finally, in the output stage, the processed results are displayed on a 7-inch TFT LCD screen using I2C communication, allowing the user to immediately view the quality parameters of the scanned juice sample. At the same time, the ESP32 transmits the same data wirelessly via Wi-Fi to a connected IoT server, where it is stored for future reference and can be accessed remotely through a mobile or web application. When the device is powered off or disconnected, the green LED turns off, indicating that the system is inactive. This structured data flow ensures that users receive timely, accurate results while benefiting from intuitive visual feedback and cloud integration.

Major Functional Block

Functional Block	**Description**
Input	Push Buttons (Power On/Scan), NIR Spectroscopy Sensor
Processing	ESP32 Microcontroller with Machine learning algorithm
Output	7” TFT Display, LED Indicator, Mobile App
Power Supply	Li-ion Battery + TP4056 + 5V Boost Converter

Hardware-Software Interaction and Design Justification

The ESP32-S3-WROOM-1 microcontroller serves as the central control unit, managing input from the NIR sensor and push buttons, executing machine learning models, and handling display and wireless communication. The hardware-software interaction is optimized through the use of:

- SPI protocol for fast, reliable data transfer from the sensor,
- I2C for low-pin-count display interfacing,
- and Wi-Fi for remote data transmission to cloud infrastructure.

The embedded software is responsible for sensor initialization, data acquisition, preprocessing, model inference, and communication. By deploying the

PLSR model directly on the ESP32, the system avoids dependency on external computation resources, making it portable and cost-effective.

Design decisions were based on the following criteria:

- Low power consumption for field deployment,
- Remote monitoring via cloud integration,
- User-friendly interaction using visual and mobile outputs.

3. Hardware Design

Materials

Hardware

ESP32 S3 WROOM-32. This will serve as the central processing unit of the system which manages communication between devices and controlling network connectivity for IoT features.

AS7263 Spectral Sensor Breakout. This will be responsible for capturing the NIR spectral data of the sugarcane juice. It will analyze the absorbance and transmittance of specific wavelengths to extract key parameters such as Brix, Pol, and Moisture Content.

18650 3.7V Li-Ion Rechargeable Battery. This will provide portable power to the entire system and ensures operational flexibility where continuous power supply may not be available.

TP4056 Lithium Battery Charger Module. This will handle safe charging of the 18650 battery using USB input.

5V Boost Converter Module. This will convert the 3.7V battery output to a stable 5V supply, which is required by most modules including the ESP32 and display.

Quartz glass. This allows NIR light to pass through accurately without interference, enabling precise scanning by the NIR sensor.

TFT Display. This will display real-time sugarcane quality parameters for user convenience during testing.

5mm Green LED. This will indicate when the device is powered on and operational.

Casing. The system will be enclosed in a durable casing designed to protect the components from environmental exposure and physical damage during field operations.

Software

C++. C++ will be used as the programming language for ESP32 to collect data from sensors and transmit it to the backend for data processing.

Python. Python will be the programming language used for backend development, handling database interactions and machine learning tools.

Javascript. Javascript will be used as the programming language to build a mobile app that will display the gathered data and the predicted data.

Arduino IDE. This software will be used to write and upload the C++ code to the ESP32 to allow communication between the sensors and backend.

MySQL. MySQL is the database that will be used in storing sensor reading and predicted values from machine learning.

Scikit-learn. Is a python library that will provide machine learning tools that are needed for processing sensor readings.

React Native. Is a mobile app development tool that will help in organizing a user interface.

Flask. Flask is a backend API that will process sensor reading, make predictions using the machine learning model then store the results in the database.

Circuit Diagram

In Figure 9, shows the proposed circuit diagram, the ESP32 is at the center of the system, serving as the central processing unit, controlling the NIR spectroscopy sensor and display.

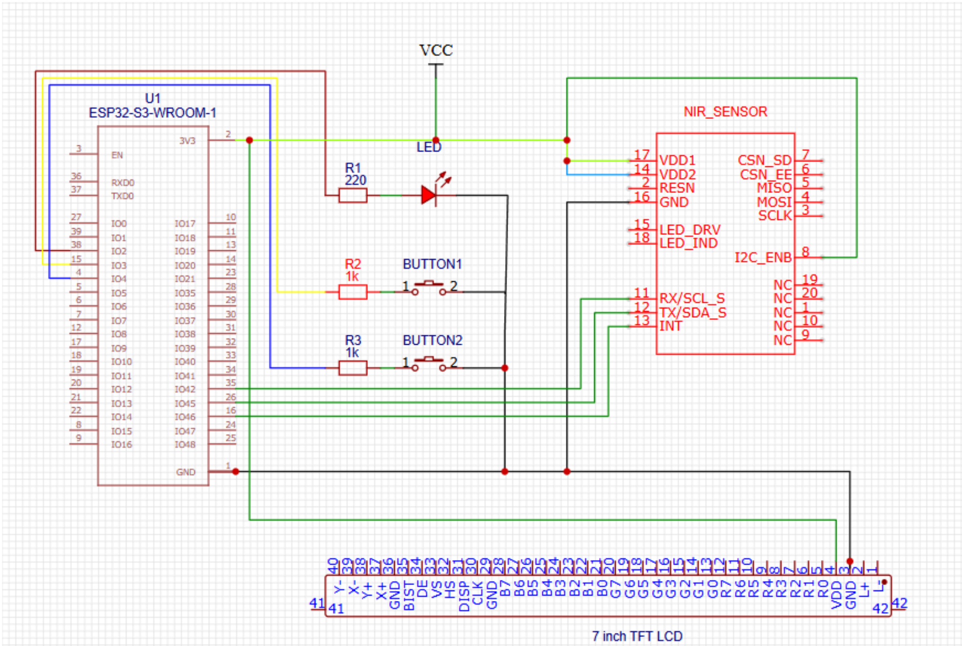


Figure 7. Schematic diagram of the proposed system

Figure 7 shows the Near Infrared (NIR) spectroscopy system using the ESP32-S3-WROOM-1 as the central processing unit. The NIR sensor is connected to the ESP32 via SPI interface lines: MISO (GPIO 12), MOSI (GPIO 11), SCLK (GPIO 13), and CSN_EE (GPIO 10). The power for the sensor is supplied through VDD1 and VDD2 pins connected to VCC, while GND is grounded. Control and indicator lines such as RESN, LED_DRV, LED_IND, and I2C_ENB are also appropriately wired to GPIOs on the ESP32. The 7-inch TFT LCD display is interfaced using I2C communication via SCL (GPIO 39) and SDA (GPIO 38) lines, allowing the ESP32 to render processed data visually. Power supply lines (VCC and GND) are shared across all modules to ensure synchronized operation.

Two momentary push buttons (BUTTON1 and BUTTON2) are wired to the ESP32 through GPIO 21 and GPIO 33, respectively, with 1kΩ pull-down resistors (R2 and R3) to keep the pins low when idle. These buttons control system power and initiate the NIR scanning process. A 220Ω resistor (R1) is used in series with a 5mm green LED, which acts as a power indicator. The LED is connected to GPIO 2 and illuminates when the system is active. The entire setup is powered by a regulated 5V DC power source, possibly backed by a 3.7V 18650 Li-Ion battery with a TP4056 charging module and 5V boost converter for stable output. This configuration provides both autonomous data analysis and wireless communication to an IoT server for real-time monitoring, supporting agricultural quality control of sugarcane juice.

Power Management

The system is designed with efficient and reliable power management to support autonomous field deployment. The 18650 3.7V Li-Ion rechargeable battery acts as the primary power source, providing energy for all system components. A TP4056 charger module safely charges the battery via USB, while the 5V boost converter ensures a consistent voltage output needed by critical modules such as the ESP32, AS7263 sensor, and TFT display.

The power path is carefully designed to support both charging and operational modes simultaneously. The green LED connected via a 220Ω resistor to GPIO 2 on the ESP32 provides visual confirmation of system status. Additional components such as the momentary buttons are powered through the ESP32's regulated GPIO outputs, maintaining low power consumption when idle.

3D Model

Figure 8 illustrates the physical design of the developed device through multiple views. Subfigure (A) shows an isometric view, highlighting the general shape, dimensions and component placement. The device has a compact dimensions of 20 cm by 20 cm by 10 cm. Subfigure (B) presents the front view, where the lid covering the scanning area is clearly visible. Subfigure (C) displays the top view, revealing the layout of the 3.5-inch TFT LCD screen, the two buttons (power and scan), and the protective lid. Subfigure (D) provides the left side view, showing the height and slope of the display section and the main body.

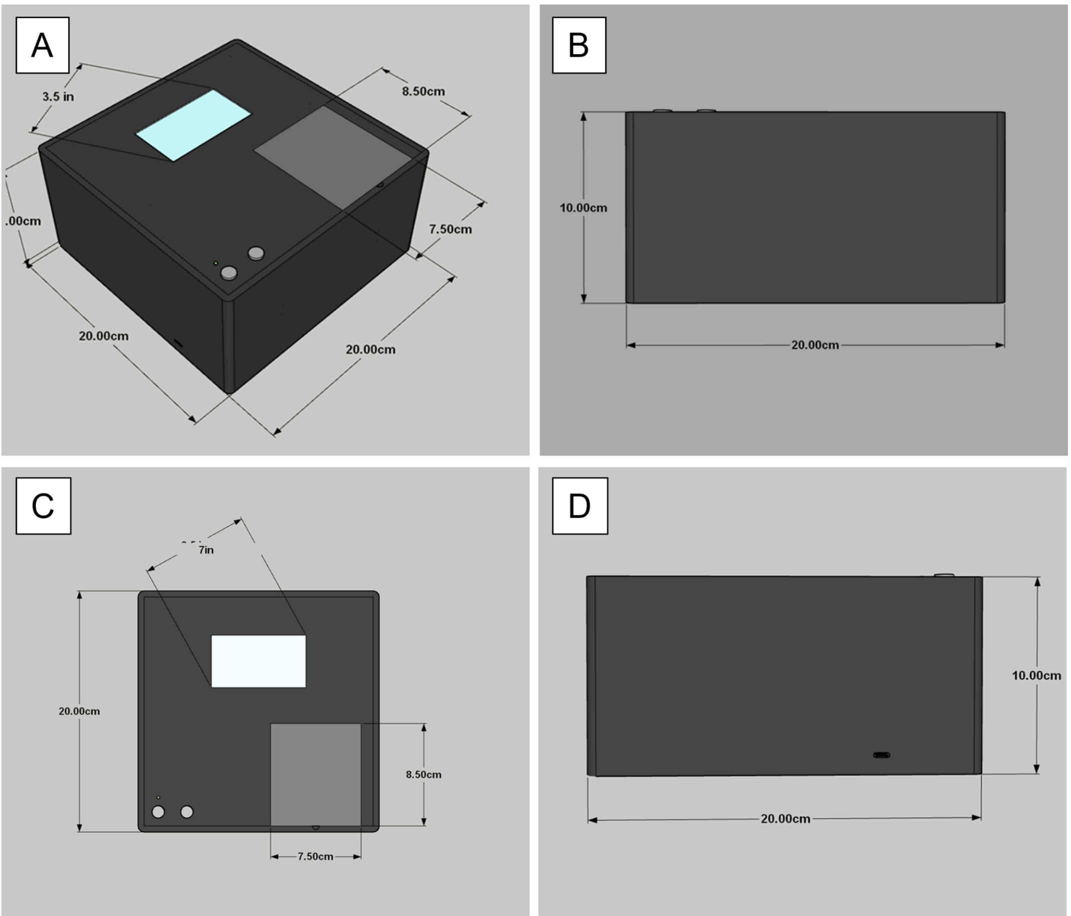


Figure 8. The 3D model of (A) isometric view, (B) front view, (C) top view and (D) left side view of the proposed system

Figure 9 shows the key external features of the developed device that support user interaction and system operation. The LCD screen displays real-time information such as Brix, Pol, moisture content, LKgTC, maturity and purity levels, making the system easy to monitor during use. The scan button allows the user to start the scanning process once a sample is in place. The power button turns the device on or off. The scanning area is protected by a lid , which helps block ambient light and ensures more accurate readings. The charging port will be used to supply power to the device.

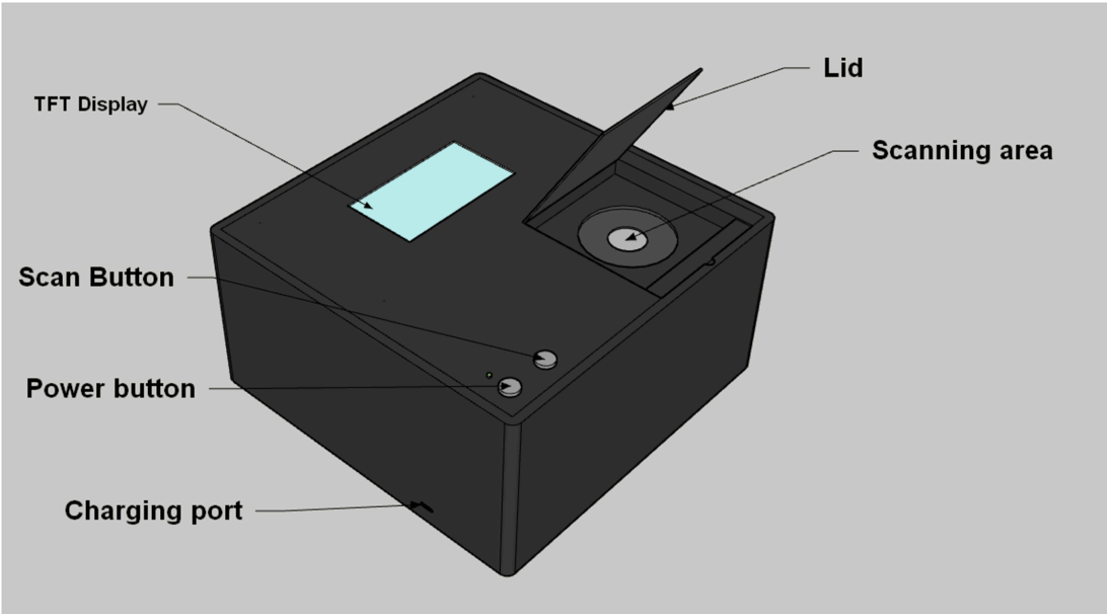


Figure 9. Labeled 3D model of the device

4. Software Design

Development and Training of the Machine Learning Algorithm Model

Data Collection

Fresh sugarcane juice will be extracted from different sugarcane varieties at varying maturity levels. The sugarcanes will be sliced and pressed, and the extracted juice will be stored in airtight containers to avoid fermentation. To ensure representativeness and capture spectral variability, each sample will be collected from multiple regions and maturity stages, with a target of at least 150–200 unique samples.

Each sample will undergo reference analysis for three quality indicators:

- Brix% (measured via digital refractometer)
- Pol% (measured via polarimeter)

- Moisture content (measured through a moisture analyzer)

Table 3. Ranges and classification criteria for sugarcane juice quality parameters (Briones, 2020)

Parameter	Description	Typical Range	Classification System
Brix (%)	Total soluble solids in juice, primarily sugars; indicates overall sugar content	12–22%	**Low**: <14% - **Medium**: 14–18% **High**: >18%
Pol (%)	Apparent sucrose content determined via polarimetry	10–20%	**Low**: <12% - **Acceptable**: 12–16% **High**: >16%
Moisture (%)	Amount of water content in the juice or cane sample	60–75% (in whole cane)	**Low**: <65% - **Optimal**: 65–70% **High**: >70%

This reference data serves as ground truth for supervised machine learning.

Data Cleaning and Normalization

After obtaining raw NIR spectral data (400–980 nm range) for each sample, the data will undergo preprocessing to reduce noise and standardize input. The preprocessing steps include:

- Savitzky-Golay Smoothing for noise reduction
- Multiplicative Scatter Correction (MSC) and Standard Normal Variate (SNV) for light scatter correction
- First and second derivative transformations to enhance peak separation and remove baseline drift

All preprocessing will be implemented using SciPy and NumPy libraries in Python. After preprocessing, the data will be normalized using Min-Max scaling or

Z-score standardization to bring all features to a uniform scale.

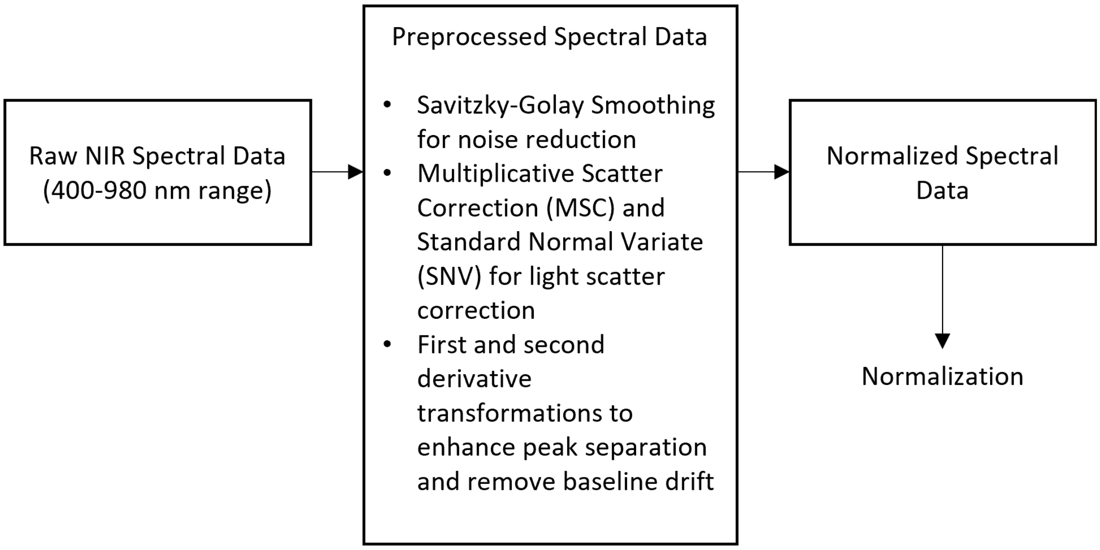


Figure 10. NIR Spectral Data Preprocessing to Normalization Flowchart

Model Development and Training

The dataset for training the PLSR model will be divided into two primary subsets: the training set (90%) and the validation set (10%). This approach allows the model to effectively learn from the NIR spectral inputs and associated quality parameters, such as Brix, Pol, and purity, while using the validation set to fine-tune model parameters and evaluate performance on unseen data. This strategy supports optimal model development by maximizing data usage and minimizing overfitting, which is especially important given the limited size of the dataset.

Table 4. Dataset subset

Dataset Subset	Percentage of Total Data	Purpose
Training Set	90%	Training
Validation Set	10%	Hyperparameter Tuning

Due to the limited dataset size (n=25), K-Fold Cross-Validation (K=10) will be used instead of a static train-test split to ensure robust evaluation. In each fold, the dataset is divided into a training set (90%) and a validation set (10%). This process is repeated 10 times, each time with a different validation subset, and the results are averaged to evaluate model performance. Hyperparameter tuning (e.g., number of components for PLSR, C and gamma for SVR, number of estimators for RF) will be conducted using GridSearchCV.

The three regression algorithms, Partial Least Squares Regression (PLSR), Support Vector Regression (SVR), and Random Forest Regression (RFR), were selected based on their established performance in spectroscopic and sugarcane juice quality prediction tasks based in another studies:

Table 5. Spectroscopic performance of PLSR, SVR and RFR (Taira et al., 2013-2022)

Model	**Brix RMSE**	**Pol RMSE**	**Moisture RMSE**	**R^2 Score**	**Reference**
PLSR	0.65	0.72	0.89	0.91	Taira et al., 2013; Kuswurjanto et al., 2019
SVR	0.58	0.64	0.81	0.94	Shao et al., 2021; Wang et al., 2020
RFR	0.55	0.59	0.76	0.96	Balasubramanian et al., 2022; Breiman, 2001

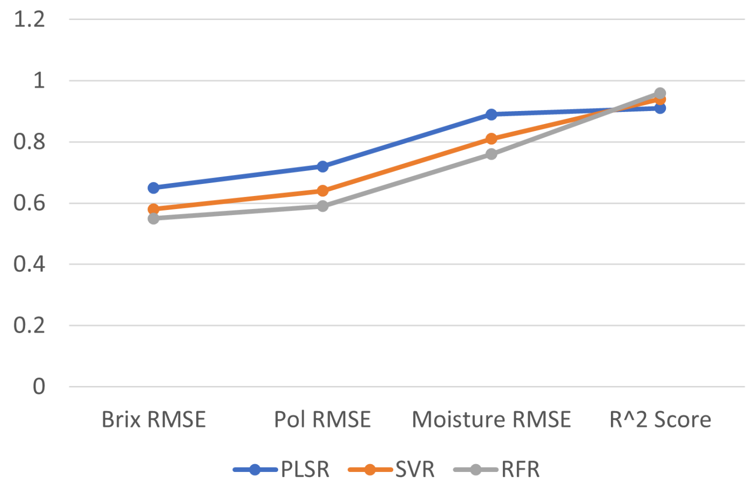


Figure 11. Visualized spectroscopic performance of PLSR, SVR and RFR (Taira et al., 2013-2022)

The PLSR model will be developed using the PLSRegression module from scikit-learn. The SVR model will use the SVR module with radial basis function (RBF) kernel. The Random Forest will be implemented using RandomForestRegressor. The independent variables (features) are the preprocessed spectral intensities, while the dependent variables (targets) are the Brix, Pol, and Moisture contents.

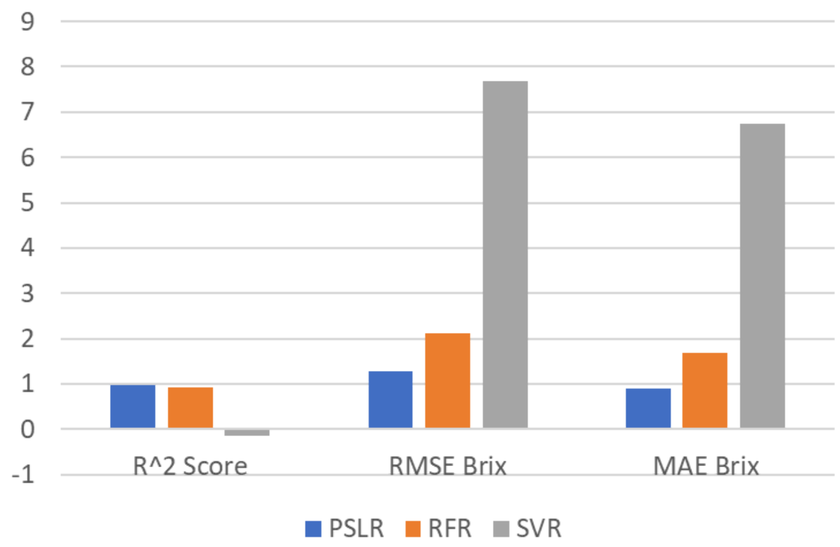


Figure 12. Pre-simulated Model Results (Brix Prediction on 25 Smoothed Sugarcane Samples) (Renganathan et al., 2021)

The dataset is sourced from a study by Renganathan et al., involving 25 sugarcane juice samples analyzed using Near-Infrared (NIR) spectroscopy across the 700–1100 nm range. Each sample's spectrum will be paired with corresponding Brix values obtained through laboratory refractometry. To improve spectral quality, Multiplicative Scatter Correction (MSC) will be applied to reduce scattering artifacts, followed by Savitzky-Golay filtering to smooth the data while preserving key spectral features.

The preprocessed data is then used to train and evaluate three machine learning models: Partial Least Squares Regression (PLSR), Support Vector Regression (SVR), and Random Forest Regression (RFR). Model performance will be assessed using 10-fold cross-validation to ensure generalizability and avoid overfitting. Among the models, PLSR achieved the highest predictive accuracy for Brix content ($R^2 = 0.968$), outperforming both RFR ($R^2 = 0.913$) and SVR ($R^2 = -0.133$).

These results indicate that PLSR is highly effective in modeling the complex relationship between near-infrared spectra and sugar content in sugarcane juice. Its ability to handle multicollinearity and process high-dimensional spectral inputs makes it particularly suitable for applications involving NIR spectroscopy. Given these strengths, PLSR is the most appropriate choice for developing a robust Brix prediction model in an IoT-enabled sugarcane juice quality assessment system.

Model Evaluation

The trained models will be evaluated using the following regression metrics:

- **R² Score**

  Since the traditional method will be replaced by spectroscopic + machine learning, the evaluation of the PLSR model, the accuracy and the reliability of the model in predicting key quality parameters of the sugarcane juice (Brix, Pol, and moisture content) will be calculated by using the equation []. This will quantify the model's performance using statistical metrics that would tell how close the predictions are to the lab-measured values.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(yi - \hat{y}i)^2}{\sum_{i=1}^{n}(yi - \bar{y})^2}$$

yi = Actual (lab-measured) value

$\hat{y}i$ = Predicted value from the model

$\bar{y}$ = Mean of actual values

R^2 is the coefficient of determination, yi is the actual lab-measured value, $\hat{y}i$ is the predicted value from the model and the $\bar{y}$ is the mean of the actual values. The formula would tell us how well the PLSR model explains the variance in the lab values (Brix, Pol, Moisture) using only the NIR spectra.

The closer the coefficient of determination is close to 1 is the indicator of excellent model performance.

- **Root Mean Square Error (RMSE) – to penalize large errors**

Since the spectroscopy method would be utilized, the evaluation of how large the average prediction of the machines needed. It will be done using the formula [], which will be the same units as the target (e.g. %Brix).

$$RMSE = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}(yi - \hat{y}i)^{2}}$$

The lower the computed RMSE, the more accurate the system is in predicting the values. A target is needed to be met for it to be acceptable on field use.

- **Mean Absolute Error (MAE) – to assess average model deviation**

Mean Absolute Error is also used to compute the clear average error magnitude. It is the conservative version of the RMSE, which is not sensitive to outliers.

$$MAE = \frac{1}{n}\Sigma_{i=1}^{n}|yi - \hat{y}i|$$

The lower the MAE, the more consistent the model is to the actual values.

Calculation for LKgTC (50-kg Bags of Sugar per Ton of Cane)

To estimate the recoverable sugar yield per metric ton of harvested cane:

$$LKgTC = Brix\ x\ Recovery\ Coefficient$$

Calculation for Sucrose Content

To determine the amount of sucrose present in the juice:

$$Sucrose\ Content\ (\%) = Brix\ x\ (\frac{Purity}{100})$$

Calculation for Maturity

Maturity is assessed to identify the optimum harvest time of sugarcane:

$$Maturity\ Ratio = (\frac{Brix_{Top}}{Brix_{Bottom}})$$

Calculation for Purity

Purity indicates the proportion of sucrose in the total soluble solids:

$$Purity\ (\%) = \left(\frac{Pol}{Brix}\right) x\ 100$$

Model Deployment

After selecting the best-performing model (based on RMSE and R^2 scores), it will be deployed in the system's backend, integrated with the IoT-NIR device and a web-based interface. The chosen model will take live input from the NIR sensor, preprocess the spectra on-device or cloud, and return predictions for Brix, Pol, and Moisture in real time.

Deployment will include:

- Conversion of the model using joblib or ONNX for portability
- Integration into a Python Flask/Node.js API
- Frontend display of real-time predictions via a web dashboard

Mobile Application Development

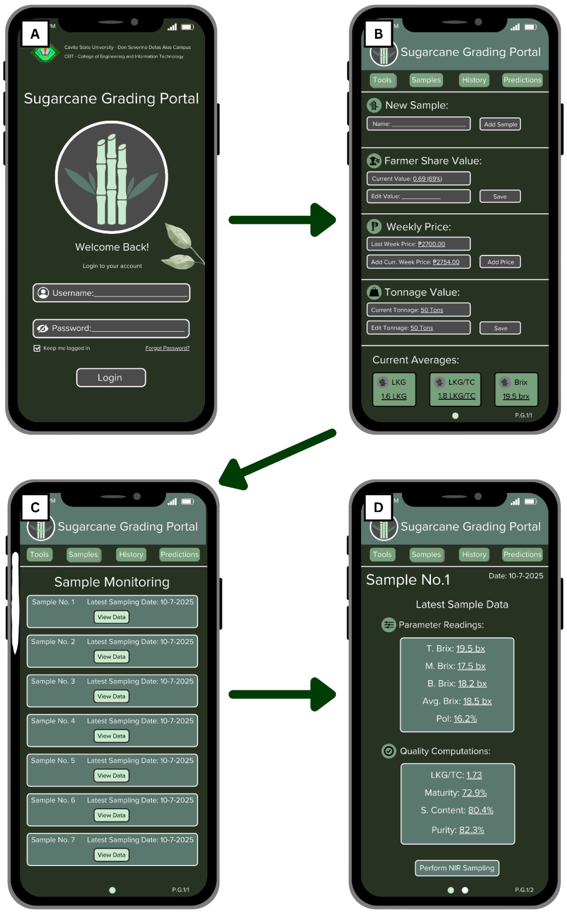


Figure 13. (A) Login page of the mobile application , (B) User tool tab, (C) Sample monitoring tab (D) Sample latest data readings

Figure 13. Section A shows the login page of the mobile application where the user can login using their username and password and it also includes the keep me logged in checkbox and forgot password. Section B is the tool tab where the user can add new samples and weekly price, edit share value and tonnage. Section C is the sample monitoring tab where users can access the latest readings of the samples. Lastly, Section D shows the latest values read by the NIR sensor for the specific samples to track the sugarcanes qualities and also conduct a new NIR scanning.

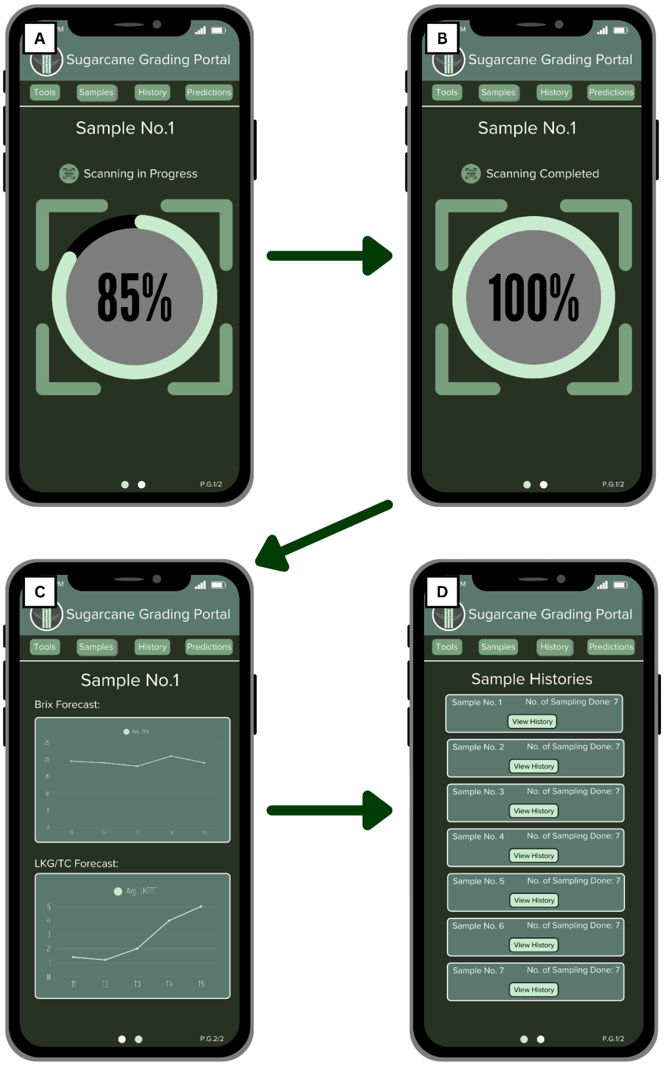


Figure 14. (A) Scanning in progress, (B) Scanning Completed, (C) Sample LKGTC and Brix graph forecast (D) Sampling history tab

Figure 14. Section A is the loading progress when NIR scanning is ongoing and Section B is the scan completion message to inform the user about the progression of their sampling. Section C shows the graph of the Brix and LKGTC forecasting of the specific samples. Section D is the sampling history tab where the user can access the history of their sugarcane samples.

Figure 15. (A) Specific sample history data tab, (B) Sampling history value, (C) Specific sample forecast graph and (D) Weekly price history

Figure 15. Section A is the tab to access the specific sample history data wherein Section B shows the sampling history NIR readings for the parameter reading and quality reading. Section C shows the historical graph forecast of the specific sampling and Section D shows the tab for the user to access the history of the weekly price to inform them about the status of the sugarcane market.

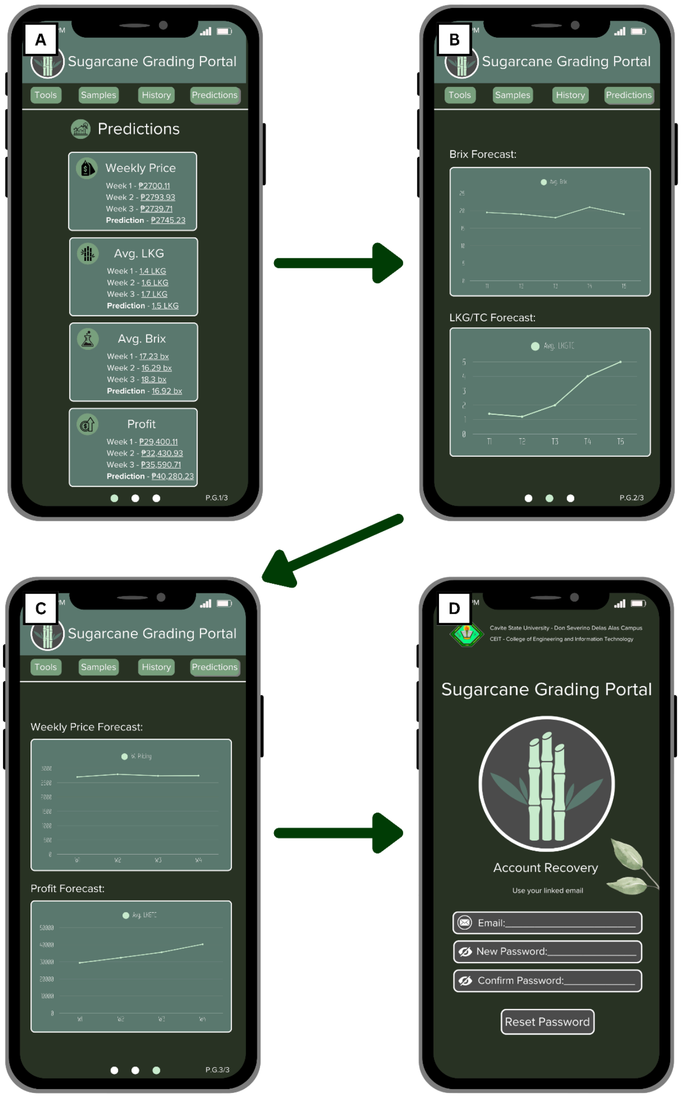


Figure 16. (A) Predictions tab, (B) Predictions graph for LKGTC and Brix, (C) Predictions graph for Profit and Weekly Price, (D) Forgot Password Tool

Figure 16. Section A shows the predictions tab where the user can access information regarding the predictions of weekly price, LKG, Brix and Profit and Section B shows the graphical representation of the Brix and LKG forecasting continued by the Section C where the graphical representation of the Weekly Sugarcane Price and Profit calculation and forecast. Lastly, Section D is the forgot password tool to help the user recover their account.

1. Integration

System Integration

The diagram in Figure 17 represents the layers and the required components and software for the mobile application development. The sensor utilized is an NIR Spectrophotometer Sensor that reads the required sugarcane qualities such as brix, pol and moisture content and the recorded qualities are then sent to an MySQL database and the Flask API that is then displayed by the mobile application where the UI/UX are integrated for the users ease of use and functionality. The hardware, which includes the NIR spectroscopy sensor and ESP32 microcontroller, sends real-time data about the sugarcane juice to the cloud computing software. This software, powered by Flask for processing and MySQL for database management, provides processed data and predictions (like Brix and Pol levels) to the mobile app. The mobile app, in turn, allows users to interact with the system by sending user requests and credentials back to the cloud software. This integrated flow ensures that juice quality insights are efficiently captured, analyzed, and delivered to users for informed decision-making.

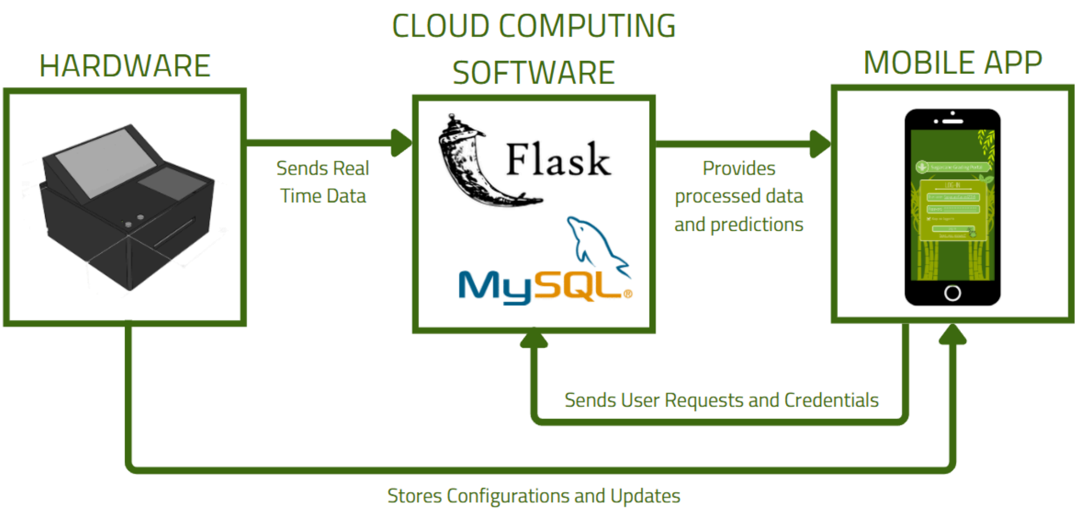


Figure 17. IoT Architecture Layers and Component

5. System Testing and Evaluation

System Testing

Testing and performance of the model

Before deployment, the integrated IoT-enabled NIR spectroscopy system for sugarcane juice quality prediction is subjected to rigorous testing. This phase ensures that all components—from sensors to prediction models—function accurately, reliably, and consistently under defined thresholds. For each test case, 20 trials were conducted to ensure statistical validity, following the rule of thumb used in engineering systems testing (Mason et al., 2003; Ahmar et al., 2024), which states that a minimum of 20 samples provides sufficient statistical power for preliminary system validation.

Functional Tests

Functional testing evaluates each component of the system performing its intended function correctly such as unit testing, integration testing, system testing and acceptance testing.

a. Unit Testing

Unit testing checks each component of the system individually to ensure it functions accurately before integration. To evaluate these units, statistical measures such as Percentage Accuracy, Standard Deviation (σ), and Root Mean Squared Error (RMSE) were used. The passing criteria required each unit to meet defined thresholds, such as variance ≤ ±2% or prediction within ±1%. The overall pass rate will be computed using:

$$Passing\ Rate\ (\%) = \left(\frac{Number\ of\ Successful\ Trials}{Total\ Trials}\right) \times 100$$

$$Variance = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

The NIR sensor's accuracy will be verified by scanning juice samples and ensuring that at least 80% of the outputs exhibited a variance of less than ±2%. The preprocessing function will be evaluated using synthetic spectral data, with a passing criterion defined as 80% or more of the denoised outputs matching the expected results. Lastly, the PLSR model is tested using known inputs, and it is considered accurate if 90% or more of its predictions fell within ±1% of the corresponding reference values. Each of these test cases will be conducted over 20 trials and runs to ensure consistency and reliability of the results.

Table 6. Unit testing (Ahmar, et al., 2024)

Test Case ID	Test Name	Description	Steps to Execute	Expected Result
UT-001	NIR Sensor Output Accuracy	Near Infrared Sensor	Illuminate juice sample, capture NIR signal	≥ 80% of the scans show signal variance < ±2%
UT-002	Preprocessing Functionality	Preprocessing	Input synthetic spectral data	≥ 80% of outputs match expected denoised data
UT-003	PLSR Prediction Consistency	PLSR Model	Feed known inputs into PLSR mode	≥ 90% predictions within ±1% of reference

b. Integration Testing

Integration testing ensures smooth data flow and communication between units. Statistical success is based on packet loss, data integrity, and output accuracy. A system passed if 90–100% of 20 trials were correct, using the same passing rate formula.

The data flow from sensor to processor will be confirmed by verifying the correct transmission of structured data packets from the NIR sensor to the processor without any corruption. The connection between the processor and the model will be validated by ensuring that the preprocessed spectral data is accurately passed to the PLSR model for analysis. Finally, the flow from the model to the user interface is tested by confirming that the prediction results displayed on the screen matched the expected values in at least 90% of the cases. Each of these test cases will be conducted over 20 trials and runs to ensure consistency and reliability of the results.

Table 7. Integration testing (Ahmar, et al., 2024)

Test Case ID	Test Name	Description	Steps to Execute	Expected Result
IT-001	Sensor-to-Processor Flow	NIR Sensor, Microcontroller	Capture real sample and send data to processor	100% correct data structure & packet transmission
IT-002	Processor to Model	Processor, Preprocessing, PLSR Model	Run full pipeline from sensor to prediction	Preprocessed data is correctly fed into model
IT-003	Model-to-UI Display	PLSR Model, Output Interface	Run prediction, verify displayed output	≥ 90% accurate display of values

c. System Testing

System testing evaluates the entire hardware and software system running together in real-world conditions. This ensures the system works under normal and continuous operation.

Real-time responsiveness will be evaluated by measuring the time taken from scanning a sample to displaying the results, with the expected response time falling within 5 to 10 seconds. Additionally, batch sample evaluation will be conducted to test the system's stability by processing multiple samples consecutively, ensuring that the system operates continuously without crashing or performance degradation. Each of these test cases will be conducted over 20 trials and runs to ensure consistency and reliability of the results.

Table 8. System testing (Ahmar, et al., 2024)

Test Case ID	Test Name	Description	Steps to Execute	Expected Result
ST-001	Real-time Responsiveness	Evaluate processing speed	Time the interval from scanning to result display	Response time 5 to 10 seconds
ST-002	Batch Sample Evaluation	Test system on multiple sequential samples	Insert 10 samples in sequence	No system crash in ≥ 90% of trials

d. Acceptance Testing

Acceptance testing checks whether the system met the design project objectives and is ready for deployment based on reliability.

Continuous use testing involved running the system for a duration of one hour, during which it analyzed samples at three-minute intervals. The objective will ensure that at least 90% of the results remained consistent throughout the test period, and that the system did not experience any freezing, crashes, or operational failures during continuous operation. Each of these test cases will be conducted over 20 trials and runs to ensure consistency and reliability of the results.

Table 9. Acceptance testing (Ahmar, et al., 2024)

Test Case ID	Test Name	Description	Steps to Execute	Expected Result
AT-001	Reliability Over Time	Continuous operation test	Operate for 1 hour with samples every 180 seconds	≥ 90% consistent results and uptime

Non-functional Tests

Non-functional testing evaluates the system's overall quality characteristics such as performance, usability, and scalability.

a. Performance Testing

Performance testing evaluates how fast and accurately the system performs under standard conditions, focusing on processing time and prediction quality.

The latency test will be conducted to ensure that the system delivers results within 10 seconds of scanning a sample, maintaining real-time performance. Model accuracy is assessed using real validation data, with the prediction model required to achieve a coefficient of determination (R^2) of at least 0.85 and a root mean squared error (RMSE) of 1.0 or lower to be considered reliable. Each of these test cases will be conducted over 20 trials and runs to ensure consistency and reliability of the results.

Table 10. Performance Testing (Ahmar, et al., 2024)

Test Case ID	Test Name	Description	Steps to Execute	Expected Result
PT-001	System Latency Test	Response Time (sec)	Scan sample and time from input to UI display	Result displayed within ≤ 10 seconds
PT-002	Model Accuracy Evaluation	Accuracy, RMSE, R^2	Use validation dataset and compare predictions to lab reference values	$R^2 \geq 0.85$; RMSE ≤ 1.0

System Evaluation

The researchers will create a survey questionnaire following ISO (International Organization for Standardization) standards during usability testing operations. According to the ISO25010 standards which outlines the quality model for system and software product evaluation. This standard defines eight key quality characteristics such as functional suitability, reliability, usability, and performance efficiency each of which can be evaluated through user perception and experience. The survey will also gather satisfaction ratings from the evaluators. It will be given to sugarcane farmers. A Likert scale will be used to collect numerical responses, providing a detailed evaluation of the evaluators' experiences with the machine's performance and their satisfaction with it.

Table 13. Response categories of the survey questionnaire used and their numerical values (ISO, 2011)

EQUIVALENT RATING	RATING SCALE
Strongly Agree	5
Agree	4
Neutral	3
Disagree	2
Strongly Disagree	1

a. Accuracy Evaluation

To evaluate the accuracy of the developed NIR Spectrometry System, the researchers will conduct a comparative analysis using data from traditional and laboratory methods. This evaluation will focus on the system’s ability to accurately measure Brix and Pol values which are critical indicators of sugarcane juice quality. The researchers will perform a series of 20 tests and compare the NIR system’s results against readings obtained using a refractometer, polarimeter and standard laboratory procedures. For each test percent error will be calculated to determine how much the results deviate from the reference values. This data driven approach will provide a clear and objective assessment of the system’s accuracy and reliability in real world conditions.

Comparison Table of Traditional - Laboratory - NIR Spectrometry System

Table 11. Comparison of Traditional Method and NIR Spectrometry System for getting Brix and Pol

Test No.	Brix(Refractometer)	Pol(Polarimeter)	NIR		Percent Error
			Brix	Pol	
1 2 ~ 20					
				Average	

Table 12. Comparison of Laboratory Method and NIR Spectrometry System forgetting Brix and Pol

Test No.	Laboratory				NIR				Percent Error
	LKgTC	Sucrose Content	Maturity	Purity	LKgTC	Sucrose Content	Maturity	Purity	
1 2 ~ 20									
							Average		

b. Software Evaluation

The researchers will also design a survey questionnaire based on the ISO/IEC 25010:2011 standard, which defines the quality model for system and software product evaluation. This evaluation will focus on critical software attributes such as functional suitability, performance efficiency, reliability, usability, and maintainability. The survey aims to capture both objective and subjective feedback, particularly the satisfaction levels of users and system administrators who interact with the software. Using a Likert scale, the questionnaire will collect numerical data that reflects the evaluators' experiences with the software's functionality, ease of use, and overall performance. This approach ensures a standardized and comprehensive assessment of the software's quality from the perspective of its actual users.

SOFTWARE EVALUATION FORM

DATA PRIVACY NOTICE

In accordance with the Data Privacy Act of 2012 (RA 10173), all personal information collected in this form shall be kept strictly confidential and used only for research evaluation purposes. Your responses will be stored securely and will not be shared with third parties without your consent. By completing this form, you consent to the collection and use of your data as stated above.

Part 1: PERSONAL INFORMATION

Name: ____________________________ Sex: ____ Age: ____

Respondent: Admin ( ) User ( ) Faculty ( )

Address: ____

Part 2: ISO 25010 EVALUATION FORM

Please evaluate the design project entitled "**Design and develop an IoT-enabled system with Near-Infrared (NIR) spectroscopy for determining the quality of sugarcane juice (Saccharum officinarum).**." by rating each item in the questionnaire using the rating scale shown below. Put a check mark (✔) on the column that corresponds to your assessment.

Numeric Value	Equivalent Rating
5	Strongly Agree
4	Agree
3	Neutral
2	Disagree
1	Strongly Disagree

A. Functional Suitability	5	4	3	2	1
1. **Functional completeness** - The web application does all specified tasks and intended users' objectives.					

2. **Functional correctness** - The web application provides accurate results when used by intended users.					
3. **Functional appropriateness** - The web application meets all the functional requirements of the users.					
B. Performance Efficiency					
1. **Time behavior** - The web application responds quickly to user interactions and commands.					
2. **Resource Utilization**- The web application operates smoothly without excessive demand on hardware or network bandwidth.					
C. Interaction Capability					
1. **Appropriateness Recognizability** - The web application's interface makes its purpose and functions immediately clear to users.					
2. **Learnability** - The web application is easy to learn and use for both admin and normal users.					
3. **Operability** - The web application is straightforward to operate, with user-friendly controls.					
4. **User Error Protection** - The web application provides clear feedback and correction options for any errors made.					
5. **User Engagement** - The web application presents functions and information in an engaging and motivating way.					
6. **Self-Descriptiveness** - The web application provides appropriate information to make its features and capabilities immediately understandable.					
D. Reliability					
1. **Faultlessness** - The web application consistently delivers accurate and reliable results.					
2. **Availability** - The web application is always operational and accessible when needed by users.					
E. Flexibility					
1. **Scability** - The web application can handle increased workloads without performance degradation.					

Feedback/Suggestions:

____________________________ __________________

Evaluator Signature Date

Thank you for taking the time to complete this evaluation form.

c. User Evaluation

To ensure the system's overall effectiveness and usability in agricultural settings, a comprehensive evaluation framework is designed that emphasized critical criteria such as functional suitability, performance efficiency, Interaction Capability, Reliability, and Flexibility. This evaluation not only measured how accurately and reliable the system predicts sugarcane juice quality parameters (Brix, Pol, and Purity), but also how sugarcane farmers could operate it without prior technical training. Feedback will be collected through user feedback and system performance assessments to identify areas for improvement. The insights gathered from this process will ensure that the proposed system is both practically useful and user-friendly, contributing to smart farming practices and data-driven agricultural decision-making.

USER ACCEPTABILITY EVALUATION FORM

Part 1: PERSONAL INFORMATION

Name / Pangalan: ______________________Sex (Kasarian): ____ Age (Edad): ____

Respondent: Sugarcane Farmers (Magsasaka ng Tubo) ( )

Address: ____

Part 2: ISO 25010 EVALUATION FORM

Please evaluate the design project entitled "**Design and develop an IoT-enabled system with Near-Infrared (NIR) spectroscopy for determining the**

quality of sugarcane juice (Saccharum officinarum)." by rating each item in the questionnaire using the rating scale shown below.

Put a check mark (✔) on the column that corresponds to your assessment.

Mangyaring sagutan ang pagsusuri ng aming proyekto na nagngangalang **"Design and develop an IoT-enabled system with Near-Infrared (NIR) spectroscopy for determining the quality of sugarcane juice (Saccharum officinarum)."** sa pamamagitan ng pag-rate sa bawat item sa talatanungan gamit ang antas ng pagsusuri na makikita sa ibaba.

Lagyan ng check (✔) ang kahon na tumutugma sa iyong opinyon.

Numeric Value	Equivalent Rating
5	Strongly Agree (Lubos na Sumasang-ayon)
4	Agree (Sumasang-ayon)
3	Neutral
2	Disagree (Hindi Sumasang-ayon)
1	Strongly Disagree (Lubos na Hindi Sumasang-ayon)

A. Functional Suitability (Kakayahan sa pang-aangkop)	**5**	**4**	**3**	**2**	**1**
1. Functional completeness (Kumpleto ang Gamit ng Sistema) The system does all specified tasks and intended users' objectives. (Nagagawa ng sistema ang lahat ng dapat nitong gawin ayon sa layunin ng mga gagamit.)					
2. Functional correctness (Tama ang Kinalabasan) - The system provides accurate results when used by intended users. (Nagbibigay ang sistema ng tama at maaasahang resulta kapag ginagamit ng tamang tao.)					
3. Functional appropriateness (Akma sa Pangangailangan) - The system meets all the functional requirements of the users. (Natutugunan ng sistema ang mga kailangang gamit ng mga gumagamit nito.)					
B. Performance Efficiency (Pagkabisa ng pagsasagawa)					
1. Time behavior (Bilis ng Tugon)- The system responds quickly to commands and delivers results within the required time. (Mabilis mag-react ang sistema sa mga utos at nagbibigay agad ng resulta.)					
2. Capacity (Kapasidad) - The system reliably handles the expected workload without performance degradation.					

(Kayang-kaya ng sistema ang dami ng trabaho nang hindi bumabagal o nagkakaproblema.)					
C. Interaction Capability (Dali ng Paggamit)					
1. Learnability (Madaling Matutunan) - The system is easy to learn and use within a short period. (Hindi mahirap pag-aralan at gamitin ang sistema kahit sa unang subok.)					
2. Operability (Madaling Gamitin) - The system is designed to ensure smooth and trouble-free usage. (Maayos at walang sabit sa paggamit ang sistema.)					
3. User Error Protection (May Gabay sa Mali) - The system provides clear feedback or notifications to correct user mistakes. (Nagbibigay ang sistema ng malinaw na babala o tulong kapag may maling nagawa.)					
4. Inclusivity (Angkop sa Lahat) - The system accommodates users of various educational backgrounds. (Naaangkop gamitin ng mga tao anuman ang kanilang antas ng pinag-aralan.)					
D. Reliability (Kapanatagan)					
1. Faultlessness (Walang Abala) - The system performs its intended functions without errors under normal conditions. (Maayos gumagana ang sistema at bihira itong magka-error.)					
2. Availability (Laging Magagamit) - The system minimizes downtime and ensures continuous usability. (Hindi madaling masira at laging puwedeng gamitin ang sistema.)					
E. Maintainability (Pagpapanatili)					
1. Modularity (May Bahagi-bahagi) - The system's modular design makes it easy to isolate and modify individual components. (Maayos ang pagkakadisenyo ng sistema kaya madaling palitan o ayusin ang mga bahagi nito.)					
2. Modifiability (Madaling Baguhin) - The system can be effectively and efficiently modified without introducing new defects. (Kung kailangan ng pagbabago, puwedeng gawin ito nang hindi nagkakaroon ng bagong problema.)					
F. Flexibility (Pag Aangkop)					
1. Scalability (kakayahan sa pagtaas ng kapasidad) - The system can handle increased workloads effectively without compromising performance.					

(Kahit dumami ang trabaho ng sistema, nananatili pa rin ang ganda ng takbo nito.)					
2. **Portability (Madaling I-set up)** - The system can be installed successfully and efficiently in its intended environment. (Walang kahirap-hirap i-install ang sistema sa lugar kung saan ito gagamitin.)					

Feedback/Suggestions (Mga Komento/Suhestiyon):

____

____

____

___________________________________ ___________________

Evaluator Signature (Lagda ng Tagasuri) Date (Petsa)

Thank you for taking the time to complete this evaluation form.

(Maraming salamat sa inyong oras sa pagsagot ng formulariong ito.)

Functional Suitability

Functional Suitability evaluates the extent to which the system delivers the required functionalities to enable sugarcane farmers and agricultural technicians to accurately determine the quality of sugarcane juice. It focuses on whether the system can measure essential indicators such as Brix, Pol, and Purity accurately and consistently, in alignment with laboratory reference results. The goal is to ensure that the system effectively supports the completion of user tasks, without having functional errors that could impact the decision making.

Performance Efficiency

Performance Efficiency refers to how quickly and effectively the system responds to user commands and processes real-time data. This includes the duration taken to scan a sugarcane sample and display the predicted values through the LCD screen and mobile devices. The system must sustain fast and reliable performance, even during high-volume operations like batch testing, to ensure timely data delivery for field-based users.

Interaction Capability

Interaction Capability encompasses the usability of the system, including ease of navigation, interface clarity, and intuitiveness for non-technical users. It considers whether sugarcane farmers can easily learn to use the device, conduct sample scans, and interpret results with minimal training. A user-friendly and informative interface is vital to ensure that the system is both practical and accessible to its intended audience.

Reliability

Reliability assesses the consistency and stability of the system during prolonged operation. It evaluates whether the system can function without crashes or significant prediction errors, especially under real-world agricultural conditions. Maintaining reliable performance is essential to earning user trust, as system failures and inconsistencies may lead to incorrect harvest timing or reduced productivity.

Flexibility

Flexibility examines the system's ability to adapt to varying user requirements including different farming practices. Both hardware and software components should be designed to function effectively without reconfiguration. This ensures that the system remains relevant and usable across a wide range of agricultural scenarios.

Cost Computation

Table 18 provides an estimated breakdown of the materials required for this study, including their respective unit costs, quantities, and total cost for production.

The total cost for each material is calculated by multiplying the unit cost by the quantity needed for the system.

Table 18. Budget estimate for the IoT-enabled system with Near-Infrared (NIR) spectroscopy for determining the quality of sugarcane juice (Saccharum officinarum).

Materials	Unit Price (₱)	Quantity	Total Price (₱)
ESP32 S3 WROOM-32	350.00	1	350.00
AS7263 NIR Spectral Sensor	2,500.00	1	2,500.00
18650 Rechargeable Li-ion Battery	90.00	3	270.00
TP4056 Battery Charger Module	30.00	1	30.00
3.5" TFT LCD Screen	1,500.00	1	1,500.00
Acrylic Plastic Casing	400.00	1	400.00
Push Button	10.00	1	10.00
Quartz Glass Sample Holder	150.00	1	150.00
5V Boost Converter Module	50.00	1	50.00
TOTAL			**5,260.00**

Cost Benefits Analysis

Development Costs

The cost of the automated system involves hardware components, software integration, equipment, and development. Based on the updated cost computation, the total development cost of the IoT-enabled system with Near-Infrared (NIR) spectroscopy for determining the quality of sugarcane juice (Saccharum officinarum) is ₱5,260.00.

Operational Costs

Operational costs include electricity, software hosting, and system maintenance. Table 19 shows the annual operational cost of the IoT-enabled system:

Table 19. Annual operational cost of the IoT-enabled system with Near-Infrared (NIR) spectroscopy for determining the quality of sugarcane juice (Saccharum officinarum).

Cost	Monthly (₱)	Annually (₱)
Electricity	150.00	1,800.00
Software Hosting	-	1,500.00
Maintenance	-	1,350.00
TOTAL		**4,650.00**

Using Table 18 and Table 19, the calculated annual estimated cost for developing and operating the IoT-enabled system with Near-Infrared (NIR) spectroscopy for determining the quality of sugarcane juice (Saccharum officinarum) will be Php 9,910.00.

Benefits

Time Efficiency

The IoT-enabled system with Near-Infrared (NIR) spectroscopy for determining the quality of sugarcane juice (Saccharum officinarum) uses NIR sensor which has an average sampling time of 0.2-1 minutes unlike the traditional method of using a handheld refractometer that has an average sampling time of 5-20 minutes (Jensen et.al., 2014). The averages for both brix % and LKg/TC of the tested sugarcane sample for reliable and accurate sampling are manually calculated by the sugarcane farmers which takes significant time compared to the usage of the NIR spectroscopy device which automatically record and calculate the values that were recorded by the NIR sensor.

Time Efficiency Ratio (TER):

$$Time\ Efficiency\ Ratio\ (TER)\ =\ TRef/TNir$$

Where:

TER = Time Efficiency Ratio

TRef = Time in Refractometer

TNir = Time in NIR Spectroscopy

Interpretation:

TER > 1: NIR spectroscopy is faster and more efficient.

TER < 1: Refractometer is faster.

TER = 1: Both methods are equally efficient.

Minimum Time:

TER = 5/0.2 = 25 (NIR Method is 25 times faster than Refractometer Method).

Maximum Time:

TER = 20/1 = 20 (NIR Method is 20 times faster than Refractometer Method).

Predictive Analysis for Sugarcane Harvest

The use of Near-Infrared (NIR) Spectroscopy in sugarcane farming has become an essential tool for evaluating key quality parameters such as Brix, Pol, and moisture content. These measurements help determine critical indicators like the Estimated LKG per Ton Cane (LKgTC), sucrose content, maturity, and overall purity of the sampled sugarcane.

By incorporating predictive analysis for values such as Brix and LKG, farmers can better anticipate the ideal harvest window which cannot be conducted by relying only on using handheld refractometers. This ensures that the sugarcane is harvested at peak quality, resulting in improved sugar yield and higher profits.

Profit Estimation Formula:

Profit = (Tons of harvested sugarcane × LKG × 0.69(farmers share) × Sugarcane bidding price)

Example Scenario:

Let's say a farmer harvests 20 tons of sugarcane. Based on NIR analysis, the estimated LKG is 1.8, and the current bidding price is ₱1,800 per LKG.

Applying the formula:

Profit = 20 × 1.8 × 0.69 × 1,800

Profit = 36 x .69 × 1,242

Profit = ₱69,552

In this scenario, the farmer can expect a projected profit of ₱69,552 from the harvest. By using predictive tools and real-time quality data, farmers can time their harvest for optimal sugar content and market value, leading to better planning and increased income.

Return of Investment (ROI)

By analyzing the initial investment for the development and deployment of the IoT-enabled NIR spectroscopy system for sugarcane juice quality assessment. The ROI calculation will highlight the significant cost savings and efficiency improvements offered by the proposed system. Equation 10 shows the general formula used to calculate the potential ROI of the system:

$$ROI = \frac{Net\ Profit}{Investment\ cost}\ x\ 100 \quad [10]$$

The Net Profit is the difference between the savings generated by the system and the total monthly cost of operating the system. The Investment Cost refers to the initial cost of the system. Equation 11 will illustrate how ROI will be calculated for this study.

$$ROI\ = \frac{(Total\ Estimated\ Benefits - Total\ Estimated\ Cost\ )}{(Investment\ cost\ )}\ x\ 100\ [11]$$

Using these estimates, the calculated potential ROI after a month will be 756.08%. The computation will be presented in Equation 12.

$$ROI\ = \frac{(\ 49{,}680 - 9{,}910)}{(5{,}260)}\ x\ 100\ [12]$$

REFERENCES

Ayele, N., & Mequanent, Y. (2014, December 16). The effect of harvest age on maturity indices of quality parameters of sugar cane varieties at Metahara sugar estate in cool season. Agriculture and Natural

Resources. https://www.academia.edu/83505168/The_Effect_of_Harvest_Age_on_Maturity_Indices_of_Quality_Parameters_of_Sugar_Cane_Varieties_at_Metahara_Sugar_Estate_in_Cool_Season

Bhatia, A., & Kumar, S. (2024). Efficacy and impact of some insecticides used to control aphids. Agricultural Sciences, 8(12), 129. http://dx.doi.org/10.3390/agronomy8080129

Bsi, A. (n.d.). Newsletter vol. 3. Northern Sugar Industry of Belize. http://www.sugarindustryofbelize.com/new-blog-1

CEDTyClea. (2024, November 13). Sugar growers call for probe of Millers claiming low cane yields. BusinessWorld Online. https://www.bworldonline.com/economy/2024/11/13/634804/sugar-growers-call-for-probe-of-millers-claiming-low-cane-yields/

Expert system for sugarcane. (n.d.). Expert System for Sugarcane. http://www.agritech.tnau.ac.in/expert_system/sugar/contactus.html

Foodsciencejournal. (n.d.-c). https://www.foodsciencejournal.com/assets/archives/2022/vol7issue3/7-4-26-992.pdf

Google. (n.d.). Datasets: Dividing the original dataset | machine learning | google for developers.Google.https://developers.google.com/machine-learning/crash-course/overfitting/dividing-datasets?fbclid=IwY2xjawKzvcpleHRuA2FlbQIxMQABHhr_vPfLQgwWV9RnoyLuBt2ROtYJQb4ne_fntcbN8VWtioLIGqDxRWxESUID_aem_DDGok6nzT1x8GctnJWc3IA

Grimaldo, S. (2023, October 9). Understanding brix: The measure of sweetness in the sugar industry – Schaffer Sugar. https://schaffersugar.com/understanding-brix-the-measure-of-sweetness-in-the-sugar-industry/

Hanna Instruments. (2014, November 24). Measuring sucrose content in the Sugar Processing Industry. Hanna Instruments Blog. https://blog.hannainst.com/measuring-sucrose-content-in-the-sugar-processing-industry/#:~:text=Different%20types%20of%20sugarcane%20can,highest%20quality%20of%20cane%20sugar.

ISO 25000 Portal. iso25000.com. (n.d.). https://www.iso25000.com/index.php/en/iso-25000-standards/iso-25010

Khumaidi, A., Raafi'udin, R., & Triastuti, N. S. (2024, December 24). Enhancing Ship Coating Quality Detection via Machine Learning-Optimized Visible Near Infrared Spectroscopy. https://www.iieta.org/download/file/fid/150425

Understanding brix: The measure of sweetness in the sugar industry. Schaffer Sugar. (n.d.). https://schaffersugar.com/understanding-brix-the-measure-of-sweetness-in-the-sugar-industry/#:~:text=Brix%20is%20used%20to%20estimate,rates%20and%20ensuring%20product%20quality.

Wetzel, W. (2025, March 31). Near-infrared spectroscopy enhances sugar cane quality analysis, study finds. Spectroscopy Online. https://www.spectroscopyonline.com/view/near-infrared-spectroscopy-enhances-sugar-cane-quality-analysis-study-finds

Xiao, Z., Liao, X., & Guo, S. (2017, February 21). Wiley Online Library | Scientific Research Articles, journals, books, and reference works. https://onlinelibrary.wiley.com/doi/epdf/10.1155/2017/1746982

Leandro, E. R., Heenkenda, M. K., & Romero, K. F. (2024). Estimating sugarcane maturity using high spatial resolution remote sensing images. Crops, 4(3), 333–347. https://doi.org/10.3390/crops4030024

Nassar, M., El Magharby, S., EL-Sorady, G., & Abdalla, A. (2022). Productivity and quality of sugar beet in relation to sowing methods and plant density under center pivot irrigation system. Egyptian Academic Journal of Biological Sciences, H. Botany, 13(2), 31–42. https://doi.org/10.21608/eajbsh.2022.252621

Tripathi, S., Singh, N., Mali, S., Naik, J. R., & Pritesh, S. M. (2017). Sugarcane/sugarcane juice quality evaluation by FT-nir spectrometer. *International Journal of Current Microbiology and Applied Sciences*, *6*(8), 3025–3032. https://doi.org/10.20546/ijcmas.2017.609.371

Chen, J.C.P. and Chou, C.C. (1993) Cane Sugar Handbook a manual for cane sugar manufacturer and chemists. John Willey and Sons, Inc., New York, 401-403. - references - scientific research publishing. (n.d.). https://www.scirp.org/reference/referencespapers?referenceid=1377415

Appendices

	A	B
1	Index	Brix
2	1	9
3	2	10
4	3	10.6
5	4	10.6
6	5	11.6
7	6	12
8	7	12.1
9	8	12.5
10	9	13.3
11	10	13.6
12	11	13.7
13	12	13.8
14	13	15.2
15	14	15.2
16	15	15.3
17	16	15.9
18	17	16.4
19	18	16.7
20	19	17
21	20	17.6
22	21	17.6
23	22	19.2
24	23	21.6
25	24	23.5
26	25	23.9

Appendix 1. Reference brix value of presimulation dataset for model selection and evaluation

	A	B	C	D	E	F	G	H	I	J	K	L
1	Index	704.4	704.9	705.4	705.8	706.3	706.8	707.3	707.8	708.3	708.8	709.3
2	1	18.4	18.7	19.7	19.9	20.3	20.8	20.8	21.4	21.3	21.6	22.5
3	2	37.2	38.1	39.2	40.6	42.7	43.2	44.2	44.9	45	45.5	45.4
4	3	34.7	35.5	38.3	40.3	41.6	43.4	43.8	44.1	45.7	45.8	46
5	4	39.2	40.6	43.1	45.6	47.9	50.2	50.4	50.6	52.2	53.1	52.7
6	5	25.22	26.33	27	28	29.67	29.89	30.78	31	31.89	32.22	31.89
7	6	16.67	18.11	19	19.78	20.11	20.89	20.78	21.11	22.22	22.22	22.22
8	7	43.78	45	47.22	49.33	50.78	52.44	53	53.22	53.78	53.89	54.22
9	8	8	8	8	8	8	8	8	8	8	8	8
10	9	34	36.3	37.7	39.8	41.7	43.1	43.6	44.3	45	45.3	45.5
11	10	17.89	18.33	18.89	19.22	20.67	20.11	21.33	21.44	21.78	22.22	21.56
12	11	8	8	8	8	8	8.11	8	8.11	8.11	8.11	8
13	12	22.78	23.22	24.33	24.67	26.11	26.78	27.67	28.22	27.78	28	28.67
14	13	25	25.67	27	28.67	30	30.44	31.44	32.22	32.67	32.56	33
15	14	7.78	8.11	8.11	7.44	8	8.11	8.11	8.56	8.11	8	8
16	15	9.78	10.33	10.44	10.33	10.56	10.89	10.78	11.33	11.44	11.67	11.44
17	16	48.1	49.7	52.2	53.9	55.8	58.2	58	58.7	58.9	58.9	58.9
18	17	11.89	12.11	12	12.67	12.89	13.56	13.22	13.78	13.11	13.78	13.11
19	18	15.44	16	16.11	16.44	17.33	17.78	17.67	18.33	18.22	18.56	18.22
20	19	24	25.22	25.22	27.44	28.78	29.56	30	30.89	31.56	31.33	31.33
21	20	20.44	21.89	22.56	23.67	24.44	26.11	26.44	26.67	27.44	27.11	28.11
22	21	21.11	22	22.33	23.11	24.56	25	25.67	25.78	26.44	26.78	26.33
23	22	8	8	8	8	8	8	8	8.22	8	8	8
24	23	29.44	30.44	33.67	34.33	36.11	37.44	37.89	39	39.44	40.33	39.78
25	24	28.78	29.78	31.44	33.22	34.67	36.11	35.67	36.78	37.67	37.78	37
26	25	20.56	21.22	22.33	23.56	24.56	25.44	25.78	26.22	27	27	26.89

	M	N	O	P	Q	R	S	T	U	V	W	X
1	709.8	710.2	710.7	711.2	711.7	712.2	712.7	713.2	713.7	714.2	714.8	715.3
2	22.9	22.6	22.8	22.6	22.9	24.2	23.9	24.3	24.5	24	24.4	24.3
3	46.7	46.5	47.6	47.4	47.2	49.6	50.2	50.3	50.5	50.2	51.4	51.2
4	48.2	47.8	48.9	49.3	49.5	52.5	52.5	54.3	54.3	53.4	54.5	55.1
5	54.7	54.4	56	56.6	57.3	58.2	60	60.8	60.8	59.9	61.9	62.4
6	33.11	33.56	33.89	33.22	34.22	35.56	35	36.44	36.67	35.56	35.89	36.56
7	22.56	22.44	23	22.67	23.33	23.67	24.56	24.56	24.33	24.67	25.44	25.33
8	55.44	55.78	56.22	55.56	56.11	58.22	59.22	59.33	59.11	58.56	59.78	59.78
9	8	8	8	8	8	8	8	8	8	8	8	8
10	46.9	46.7	47.7	47.8	48.4	51	51.4	51.4	51.5	50.8	52.5	53.2
11	22.33	22.56	22.56	22.22	23.11	24.56	24.44	24.22	24.44	24.22	24.67	24.78
12	8	8.11	8.22	8.11	8.33	8.89	8.67	8.89	8.89	8.89	9	9.11
13	29.56	29.56	29.56	29.33	29.89	30.89	31	31.22	31.78	31.67	31.56	32.44
14	33.89	35	36	35.22	36.56	38	38.56	38.78	38.67	39.11	39.78	39.44
15	8.22	7.78	7.89	7.89	7.67	8.33	7.89	7.78	8.11	8.44	7.78	8
16	11.78	12	11.89	11.44	11.67	12.33	11.78	12.33	12.11	12.22	12	12.22
17	61.2	60.6	61.7	60.7	61.1	63.5	63.7	65	64.1	64.4	64.5	64.8
18	14	13.67	13.56	13.67	13.44	14.33	14.33	14	14	14.44	14.44	14.67
19	18.67	18.67	19.22	18.78	19.11	20.22	19.89	20.11	20.44	20.56	20.56	20.56
20	32.44	33.11	32.89	33.22	34.56	35.44	35.44	35.89	36	36.11	36.33	36.89
21	28.56	29.44	29.78	28.89	29.78	31.78	31.78	31.78	32.11	32	32.78	33
22	28.22	28	27.78	28.22	28.67	29.89	29.67	30.67	30.56	30.56	31.11	31
23	8	8	8	8	8	8.33	8.22	8.22	8.44	8.22	8.44	8.33
24	40.78	41.78	41.67	41.89	42.11	44.33	44.78	45.56	45.56	46.22	46.78	47
25	39.22	39.44	39.67	40	39.78	41.78	42.33	42.22	43.44	42	44.11	43.44
26	28.11	28.56	28.22	28.44	28.89	29.56	30.33	30.44	30.44	30.11	31.22	30.78

	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ
1	715.8	716.3	716.8	717.3	717.8	718.3	718.8	719.4	719.9	720.4	720.9	715.8
2	25.7	24.6	25.5	25.3	27.6	25.9	25.4	26.6	26.5	26.8	27.1	25.7
3	52.6	52.6	53.1	52.7	55.8	54	53.6	54.5	53.8	54.2	56	52.6
4	56.9	57.5	57.4	57.9	61.3	59.1	58.2	60.1	59.2	61.2	62	56.9
5	63.6	64	63.9	63.6	68	65.3	64.8	66.4	65.8	67.3	67.9	63.6
6	37.11	38	37.67	37.78	39.89	38.22	38.67	39.22	38.22	40.33	39.11	37.11
7	26	25.44	25.89	26	27.56	26.56	26.89	27.11	26.67	26.89	27.67	26
8	60.67	60.89	60.56	61	64.11	62.22	61.56	62.78	61.22	63.11	62.22	60.67
9	8	8	8	8	8	8	8	8	8	8	8	8
10	53.6	53.5	54.6	54.2	57.9	55.9	55.1	57.1	56.4	58.3	56.7	53.6
11	25.22	25.44	25.11	26	27.56	26	25.89	25.78	26.44	27	27.11	25.22
12	9.44	9.56	9.56	9.78	11.44	10.11	9.67	10.22	9.89	9.89	10.33	9.44
13	33	33	32.89	33.11	35.33	34.11	33.56	34.89	34.11	34.89	34.89	33
14	41.22	40.67	41.56	42.11	44.11	42.89	43.56	44.11	43.56	44.22	44.56	41.22
15	8.11	8.44	8.67	8.56	9.78	8.22	8.44	8.11	8.22	8.67	8.33	8.11
16	12.56	12.33	12.67	12.56	13.33	12.56	12.67	13.11	13	12.56	13	12.56
17	66	65.8	66.6	66.4	69.5	67.2	66.8	68.3	66.6	68.3	68.8	66
18	14.89	14.67	14.67	14.89	16.44	15.11	15	15.56	15.33	15.56	15	14.89
19	20.67	21	21.44	21	23.22	21.56	21.11	22.22	21.56	22.44	22.67	20.67
20	37.44	38.33	38.56	38.22	41.89	39.56	39.33	39.44	40	40.11	40.22	37.44
21	34	33.67	34.33	34.33	37.44	35.44	34.89	35.89	35.22	36	36	34
22	32.33	32	32.22	32.11	35.56	32.78	32.67	33.44	33.56	34.22	34	32.33
23	8.22	8.33	8.44	8.33	9.44	8.44	8.56	8.78	9.11	8.78	8.78	8.22
24	48.11	47	48.89	48.67	51.11	49.22	49.11	50.56	49.89	50.89	51.33	48.11
25	44.78	44.11	44.22	44.33	47.22	45.33	45.44	46.56	46.11	47.56	47.78	44.78
26	32.11	32.11	32.11	32.11	33.89	33.44	32.56	32.78	32.89	33.67	33.89	32.11

Appendix 2.: Pre-simulation dataset of sugarcane samples for model selection and evaluation