International Journal of Design & Nature and Ecodynamics
Vol. 20, No. 2, February, 2025, pp. 401-409
Journal homepage: http://iieta.org/journals/ijdne

Rapid Differentiation of Sex in *Calamus johndransfieldii* Seedlings Using Near Infrared Spectroscopy

Muhammad Ikhsan Sulaiman[1], Rita Andini[2], Himmah Rustiami[3], Arief Arianto[4], Nurul Fitriah[5], Ahmad Zaelani[5], Muhammad Dani Supardan[6], Muhammad Hamsar Halomoan Sipayung[1], Agus Arip Munawar[7*]

[1] Department of Agricultural Product Technology, Faculty of Agriculture, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia
[2] Research Centre for Applied Botany, Research Organization for Life Sciences and Environment (ORHL), National Research and Innovation Agency (BRIN), Bogor 16911, Indonesia
[3] Herbarium Bogoriense, Research Center for Biosystematics and Evolution, Research Organization for Life Sciences and Environment (ORHL), National Research and Innovation Agency, Bogor 16911, Indonesia
[4] Research Centre for Agroindustry, Research Organization for Agriculture and Food, National Research and Innovation Agency (BRIN), Tangerang Selatan 15413, Indonesia
[5] Research Center for Genetic Engineering, Research Organization for Life Sciences and Environment (ORHL), National Research and Innovation Agency, Bogor 16911, Indonesia
[6] Department of Chemical Engineering, Faculty of Engineering, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia
[7] Department of Agricultural Engineering, Faculty of Agriculture, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia

*Corresponding Author Email: aamunawar@usk.ac.id

https://doi.org/10.18280/ijdne.200217

Received: 26 February 2024
Revised: 22 May 2024
Accepted: 5 June 2024
Available online: 28 February 2025

Keywords:
NIRS, dioecious plant, classification, NTFPs

ABSTRACT

Resiniferae rattans, or Dragon’s blood rattans (known locally in Indonesia as "Jernang"), are highly valuable Non-Timber Forest Products from Sumatra's deep rainforests. One of these species is *Calamus johndransfieldii*. Dragon’s blood rattans are dioecious, indicating that each individual plant is either male or female. Farmers prefer female plants as they produce economically valuable fruits. Traditionally, determining a plant's sex has been a lengthy process, requiring approximately 3–4 years from planting to fruit maturation. To address this, a novel and efficient method for sex determination using Near-Infrared Spectroscopy (NIRS) was investigated. Leaves collected from known male and female plants were analyzed using a Thermo Nicolet Antaris TM II MDS in conjunction with a portable sensing device (PSD NIRS i16) within the spectral range of 1,000–2,500 nm. Principal Component Analysis (PCA) was then applied to classify the spectral data, clearly distinguishing male from female plants. Further analysis utilized machine learning techniques, including Bootstrap Forest, Neural Booster, Support Vector Machines, and K-Nearest Neighbors, to refine predictive accuracy. The findings revealed marked differences in the spectral components of male and female plants, allowing for fast and reliable sex determination. The most effective method was identified as Bootstrap Forest, with an exceptionally low misclassification rate of 0.0811%. This rapid approach significantly reduces the waiting period for farmers, enhancing productivity and economic returns.

1. INTRODUCTION

In the diverse ecosystems of tropical rainforests, non-timber forest products (NTFPs) such as *Calamus johndransfieldii*, commonly known as dragon’s blood rattans or "Jernang," play an important role in both local economy and ecological balance. These resins are highly valued for their medicinal and dyeing properties, making them valuable in various industries [1]. However, the cultivation and harvesting practices for these products need to be sustainable to ensure long-term ecosystem stability and species survival. One aspect of this sustainability is recognizing and managing the biological differences between male and female plants.

The ability to differentiate between the sexes of *Calamus johndransfieldii* plants holds considerable significance in both agricultural production and biodiversity conservation. Agriculturally, the female plants are especially valuable because they bear fruits that are not only a source of the prized resin but also critical for seed production, which is vital for propagation and genetic diversity. Efficiently identifying and cultivating female plants can lead to more effective farming practices, increased yield, and higher economic returns for local communities that depend on these products for their livelihoods [2].

From a conservation standpoint, a balanced ratio of male to female plants is essential to maintain the population's genetic diversity and health. Over-harvesting one sex could lead to reduced reproductive success and long-term decline in population size, which can have cascading effects on the ecosystem. Furthermore, understanding the distribution and health of both male and female plants can inform more nuanced conservation strategies that support both the species and its habitat.

The fruits of dragon's blood produce red resin or exudate, which commands a relatively high price, ranging from IDR 800,000 (±US$53) and can reach up to 4,000,000 per kg at its peak [2, 3]. Indonesia is known as the biggest dragon's blood exporter in the world, although its total export values to foreign countries still remains unrecorded [2, 3].

The red resin is primarily exported to China, Taipei, and Hong Kong, with China's demand for this commodity ranging from 400 to 5,000 tons per year. It is primarily used in reverse glass paintings as well as in the health and pharmaceutical industries, especially in traditional Chinese medicine [3, 4]. A type of flavonoid compound known as dracorhodin, found in the red resin, is valued for its anti-haemostatic properties, enhancing blood circulation, healing internal ulcers, and serving as a painkiller. It also exhibits anti-microbial, anti-viral, anti-inflammatory, and anti-cancer properties, particularly against intestinal cancer [2, 5, 6].

Dragon's blood rattans or resiniferae rattans belong to the genus *Calamus*. *Calamus*, the largest rattan genus [7, 8] is a climbing palm which relies heavily on supporting trees to climb and obtain sunlight as they usually grow naturally in the deep tropical rainforests. Further, Henderson [7] explained that more than 400 species of *Calamus* have been known so far, and only nine species are known to produce the red resin; they are *C. confusus, C. dracunculus, C. gracilipes, C. johndransfieldii, C. latus, C. maculatus, C. micracanthus, C. propinquus, C. ruber* [9]. All those nine species were previously included in genus *Daemonorops* section Piptospatha. However, based on the current study, all species belonging to the genus *Daemonorops* are treated as synonymies of genus *Calamus* [7]. In regard to current systematic study [7], published species formerly known as *Daemonorops draco* (Wild.) Blume, will be treated as *Calamus draco* Willd.

In its natural habitat, dragon's blood rattans can be found in hill slope and valley bottom on lower montane primary forest up to 1,000 m alt [7]. Furthermore, they are known as 'dioecious plant' meaning there are male and female plants [3]. The ratios between the male and female plants are different depending on what species. The sex ratios of male and female plants of *C. gracilipes* Miquel; *C. draco* Willd; and *C. longipes* were reported to be 2.2:1; 1:2; and 1:1.84, respectively. It seemed that higher female plants' ratio rather than the male one was existed within the *C. draco*. In general, the sex ratio of dioecious plants in their natural populations is expected to be 1:1 or being equal for the purpose of long-term sustainable breeding efforts.

However, environmental conditions and stress factors may influence variations in sex expression in dragon's blood plants. Under favorable conditions, female sex expression is more likely to occur, whereas under unfavorable conditions, male expression is more prevalent. Usually, the male and female sex organs would be obviously differentiated, with a lower percentage of male plants or less than 30% in their natural population. The male and female plants should be generatively propagated, with very low possibilities (± 1-2%) via vegetative propagation, e.g. plant tissue culture [8]. Unfortunately, at their juvenile stage, distinct morphological differentiation of their sex organs (male or female) is invisible or hardly recognizable. However, this differentiation can be clearly observed after 2 to 3 years, as *C. draco* fruits can be harvested starting from the fourth year. Usually, it takes 11 to 13 months during the process of fertilization until fruit ripening stage [4].

Some indigenous tribes (e.g. Suku Kubu, Talang Mamak, and Anak Dalam inhabiting the Southern part of Sumatra) and local farmers living at the periphery of Aceh's rainforests make a living from collecting these exotic plants. Due to this, the existence of *C. draco* and other dragon's blood rattans in their natural forests has been tremendously diminished over the past three decades as a result of quite massive deforestation. On the other hand, as a result of the increased selling price, many farmers are trying to domesticate, cultivate and even propagate it in their fields.

In response to this demand, farmers tend to select, find and purchase mostly female seedlings, as female seedlings are considered to be more productive and profitable because they can produce fruit. The price of one single female seedling costs ranging from IDR 200,000 until 350,000 depending on species; with the most expensive one is *C. draco*'s seedlings. Many farmers only realize that they have chosen the incorrect plants after 4 to 6 years, as reported in personal interviews conducted in 2022. Such experiences are common for many farmers, and unknown sex segregation has caused great losses to them if they bought the wrong seedlings. However, this is already too late. Therefore, there is an urgent need in the scientific community for a novel, non-invasive technique capable of rapid, real-time analyses, a criterion potentially met by near infrared spectroscopy [9, 10].

Recent advances in rapid gender differentiation, such as those discussed in the study, are therefore crucial. By cutting down the time to identify the plant sexes from several years to potentially just a few days or weeks, these techniques can not only boost agricultural efficiency but also enhance efforts to conserve these species in their natural habitats. This research not only supports the economic aspects of NTFPs cultivation but also aligns with broader environmental goals of maintaining biodiversity and sustainable ecosystem management.

Traditionally, sex differentiation in dioecious plants, including in *Calamus johndransfieldii*, has relied heavily on morphological observations and phenological cues, a method that is both time-consuming and often imprecise until the plants reach reproductive maturity. This typically means waiting for several years for the plants to exhibit their first flowering or fruiting phases, which are the visible indicators of their sex. Given the length of this wait, farmers and conservationists face significant challenges in planning and managing cultivation and conservation efforts effectively.

Other scientific methods previously employed include genetic testing and biochemical assays, which, while more accurate than visual assessment, often require sophisticated laboratory facilities, significant financial investment, and specialized technical expertise. These methods can be inaccessible for regular use in remote or under resourced areas where many NTFPs like *Calamus johndransfieldii* are harvested.

The development and application of Near Infrared Spectroscopy (NIRS) present a significant leap in this area. NIRS offers a non-destructive, rapid, and comparatively affordable means to determine the sex of plants long before they reach maturity. By analyzing the spectral data from the leaves of juvenile plants, the NIRS method bypasses the need for flowering or fruiting, enabling earlier and more efficient decision making in both agriculture and conservation contexts.

By using NIRS, researchers can quickly gather and analyze data on the ground, without the need for extensive laboratory infrastructure. This innovation not only speeds up the process of sex differentiation but also opens up new possibilities for large-scale screening efforts, which are critical for the effective management and sustainable use of biodiversity resources like *Calamus johndransfieldii*. The introduction of such advanced technology in the field of non-timber forest products represents a transformative approach that enhances both economic and ecological outcomes.

To date, NIRS has been increasingly adopted for both qualitative and quantitative analysis in agriculture and botany. Research compiled from numerous studies suggests that NIRS is emerging as an effective instrument for various purposes, including species identification, measurement of plant nutrient content, soil quality analysis, and even forecasting agricultural yields. NIRS has advantages over traditional methodologies, enabling on-site and real-time applications. However, the existing literature also highlights certain constraints, like the requirement for comprehensive calibration as well as its vulnerability to external factors including variations in sample temperature or granularity.

The application of NIRS in evaluating different characteristics of *Calamus johndransfieldii* has gained recognition. Research has investigated and verified its efficacy in sorting, recognizing, and profiling *Calamus johndransfieldii* based on its distinct spectral characteristics and features. These prior utilizations underscore the efficiency of NIRS as an innocuous, swift, and economical approach in the study and control of *Calamus johndransfieldii*. Yet, there is an ongoing necessity for additional research to optimize and corroborate these techniques over a wider array of *Calamus johndransfieldii* species and diverse environmental scenarios. The primary objective of this study is to evaluate the potential of NIRS technology as an approach and alternative method for rapid, non-destructive assessment to detect *Calamus johndransfieldii* gender and sex, focusing on male and female recognition and chemical property determination. We argue that each *Calamus johndransfieldii* plant's gender will exhibit a unique spectral pattern in the NIR wavelength region, allowing for classification based on spectral data. We further hypothesize that NIR spectra can be used for rapid detection and prediction of sex in Jernang plant samples.

To test these hypotheses, we conducted an extensive analysis of 200 *Calamus johndransfieldii* samples using a benchtop NIRS instrument within the 1,000–2,500 nm wavelength range. We collected reflectance data in the near infrared region for intact *Calamus johndransfieldii* samples, representing a simplification compared to traditional methodologies that generally require complex sample preparation stages. Following data acquisition, spectral features were identified, and multivariate statistical analyses were conducted to categorize the sex of *Calamus johndransfieldii* plants based on their spectral properties.

2. MATERIALS AND METHODS

2.1 *Calamus johndransfieldii* samples

A total of 259 samples representing male, female and neutral sex gender *Calamus johndransfieldii* species were collected from different locations in *Bener Meriah* (4,75 °N; 96,95 °E) in Aceh province region during the peak growth season. The young, fully developed leaf stems of the Jernang plant, ranging from 30 to 50 cm in height, are selected for analysis as shown in Figure 1. A leaf from the middle of the stem is carefully snipped with scissors and placed inside a paper envelope. After 1 to 3 hours, the fresh leaf's NIRS spectral absorption is measured across wavelengths of 1,100 to 2,498 nm at 2 nm intervals.

Calamus johndransfieldii specimens were chosen due to their abundance, ecological relevance, economical value, as well as practical farming practices. To capture a broad representation of the population, a stratified random sampling technique was used. Various ecosystems, such as farmlands, woodlands, and wildlands, were pinpointed for gathering these samples. Individual specimens of *Calamus johndransfieldii* were cataloged with a distinct identifier and tagged with details including species, place of collection, and date of acquisition. Once in the laboratory, the samples were kept in a chilled, dim environment at roughly 4°C to maintain their original state prior to subsequent examination.

Figure 1. *Calamus johndransfieldii* (Jernang) plants: (a) Jernang tree grows naturally in the tropical forest, (b) Jernang fruit matured optimally, (c) Jernang resin contains more antioxidant properties especially the dracorhodin, (d) Jernang fruits of *Calamus johndransfieldii*

2.2 NIR spectral acquisition

Spectral information within the near infrared (NIR) spectrum was collected using a benchtop NIRS instrument (Thermo Nicolet Antaris TM II) aided with portable sensing device PSD NIRS i16 with NIR sensor capable of capturing wavelengths spanning 1,000 to 2,500 nm as shown in Figure 2. Each sample was placed upon sample holder to fully occupy the measurement space of the NIR device. The unprocessed spectral data retrieved from each sample through the device was systematically gathered and documented [11].

2.3 Spectral data analysis

The spectral data analysis of Jernang leaf samples was conducted to develop classification models using principal component analysis (PCA), Bootstrap Forest, K-Nearest Neighbors, and Support Vector Machine Classification to enable classification of the leaf samples based on sex. PCA, a dimensionality reduction technique, was applied to identify the underlying patterns and principal components within the high dimensional spectral data [12]. Subsequently, PCA was used as an unsupervised classification technique to maximize the separation between different classes of weed samples using the extracted principal components.

PCA utilized the information from the class labels to define linear discriminant functions that optimally distinguished between leaf samples. By projecting the transformed data onto these discriminant functions, the weed samples were classified into specific classes based on their spectral characteristics.

Bootstrap Forest, also known as Random Forest, is an ensemble learning method used for classification and regression. It works by constructing multiple decision trees during the training process and outputting the mode of the classes for classification or mean prediction for regression of the individual trees. Bootstrap Forest enhances accuracy and controls over-fitting by averaging multiple deep decision trees, each trained on different parts of the same training set.

On the other hand, K-Nearest Neighbors is a simple, instance-based learning algorithm that is used for classification and regression. In K-NN classification, the output is a class membership. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors (k is a positive integer, typically small). If k = 1, then the object is simply assigned to the class of its nearest neighbor.

For statistical validation of the models, confusion matrices, accuracy, sensitivity, specificity, and other relevant metrics were calculated and compared. The confusion matrices provided insight into the model's performance in classifying male and female plants. Metrics such as accuracy, sensitivity (True Positive Rate), specificity (True Negative Rate), precision, and F1 score were calculated for each model tested.

Furthermore, cross-validation techniques were utilized to assess the robustness and generalization of the models. The statistical analyses aimed to provide a comprehensive evaluation of the model's performance, ensuring reproducibility and reliability of the results.

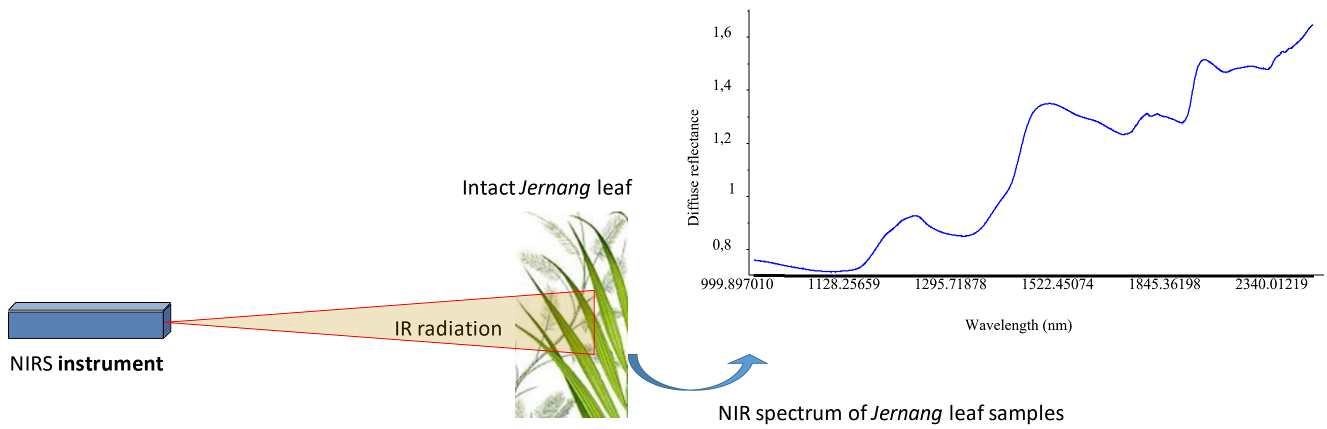

Figure 2. NIR spectral data acquisition of Jernang leaf samples in the near infrared region

3. RESULTS AND DISCUSSION

3.1 Spectra features

Typical spectrum of all *Calamus johndransfieldii* (Jernang) leaf samples from different weed plants species is presented in Figure 3. Describing the spectral features and band assignment of *Calamus johndransfieldii* leaves in the near infrared wavelength region involves analysing the specific absorbance peaks and their corresponding molecular bonds, as well as discussing the potential antioxidant properties that can be inferred from these spectral data.

In the NIR region, typically ranging from 780 nm to 2,500 nm, different overtones and combination bands of fundamental molecular vibrations occur. For plant leaves like *Calamus johndransfieldii*, commonly observed absorptions in the NIR spectrum are related to vibrations of hydrogen-containing molecular groups such as O-H (hydroxyl), N-H (amine), and C-H (alkyl groups) as presented in Figure 4. The spectral features of plants in the NIR spectrum range are primarily linked to chemical bonds such as O-H, N-H, and C-H present in plant biochemical component, which include water, cellulose, proteins, and other organic compounds [10, 13]. The combination of these components within a plant result in specific absorption features that are unique to that plant species.

The unique absorbance peaks correspond to molecular vibrations involving hydrogen-containing groups. For *Calamus johndransfieldii* leaves, key NIR spectral features might include the first overtone of O-H stretching in the 1400 to 1500 nm range often linked to water and plant tissue hydroxyl components, including those related to antioxidant polyphenolic compounds. The 1700 to 1800 nm window can suggest combination bands of O-H bending and C-H stretching, indicative of carbohydrates like cellulose or sugars within the leaves.

Absorptions perceived between 1,900 to 2,000 nm are likely to correlate with the second overtone of C-H stretching, hinting at fatty acid presence or waxy substances in the leaf surface. Additionally, the 2,100 to 2,300 nm absorption could potentially be associated with C-H combination bands, suggesting lignin and complex carbohydrates. [9, 13].

For antioxidants specifically, NIR spectroscopy is not the primary method for detailed analysis; however, the general presence of compounds with antioxidant properties can potentially be indicated through characteristic absorptions in the NIR spectrum. Polyphenols are an important group of antioxidants, and they usually display spectral features in certain regions due to their aromatic structures and hydroxyl groups.

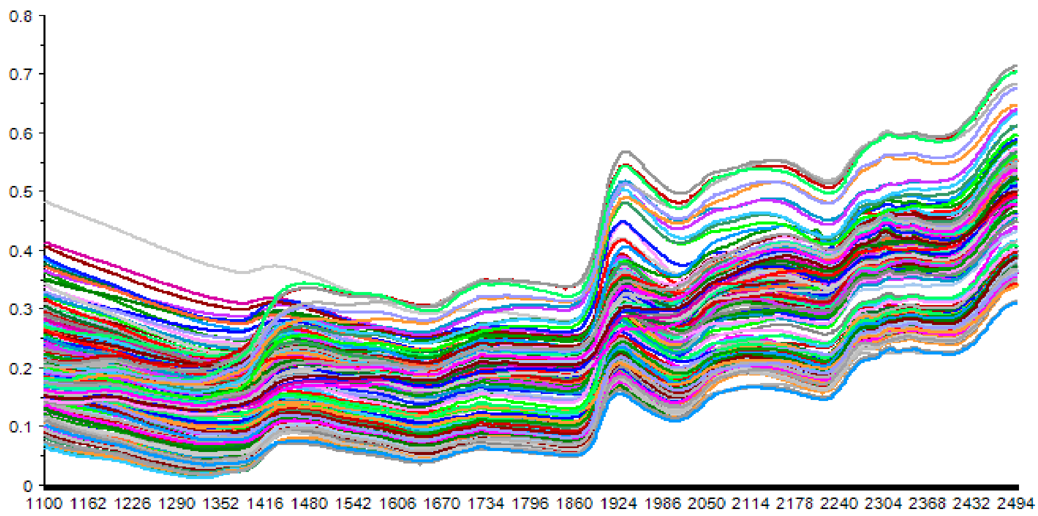

Figure 3. NIR spectrum of *Calamus johndransfieldii* plants in NIR region

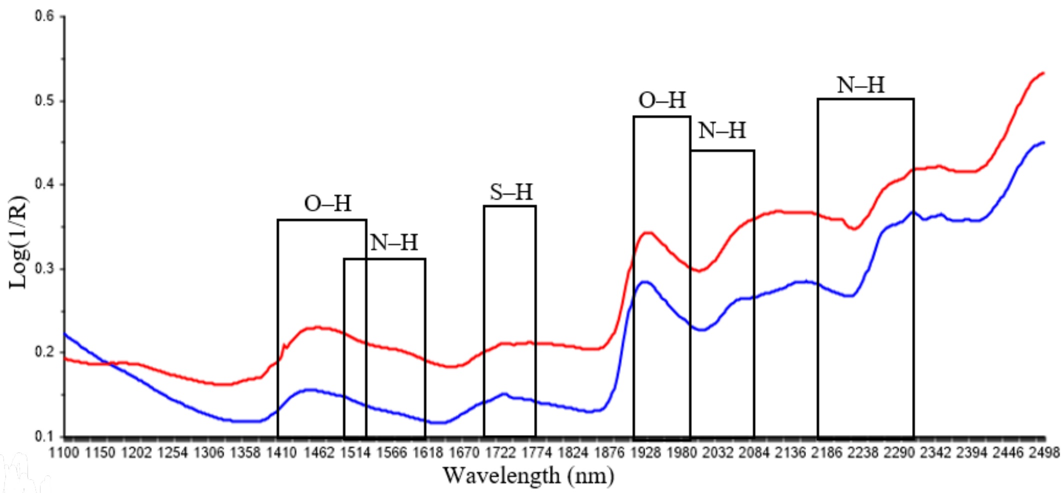

Figure 4. Bands assignment in NIR wavelength region of *Calamus johndransfieldii* plants

In terms of antioxidant properties in the NIR wavelength region, there can be an indirect analysis. For instance, if a particular wavelength shows a strong correlation with known antioxidant activity through chemometric analysis, then it's plausible to predict antioxidant content. However, it should be noted that NIR is typically used for a quick, non-destructive assessment, and detailed quantification of specific compounds like antioxidants would require more targeted analytical techniques, such as high-performance liquid chromatography or abbreviated as HPLC paired with a mass spectrometer or UV-visible spectroscopy [14].

NIRS has proven effective in distinguishing sex gender of *Calamus johndransfieldii* plants through their unique spectral signatures. It operates by examining how plant tissues reflect or absorb near infrared light, pinpointing particular molecular formations and chemical components. This spectral data serves as a benchmark for constructing classification models, allowing for the swift and non-invasive recognition of *Calamus johndransfieldii* plants directly within their sex gender as shown in Figure 5.

PCA is a powerful statistical technique that simplifies the complexity in high-dimensional data while retaining trends and patterns. It does this by transforming the data into a new set of orthogonal variables, the principal components, which are linear combinations of the original variables. Each principal component captures a portion of the total variance in the data, with the first principal component accounting for the largest variance and each subsequent component having the lesser variance.

Applied to plant classification, PCA can condense a complex dataset, consisting of various plant characteristics, into key factors that reflect the most significant variation among the data. This could include a wide array of plant traits measured, such as morphological attributes (height, leaf length, or width, petal size), chemical properties (concentrations of specific metabolites, nutrients, or toxins), or even ecological preferences like optimal light conditions, soil and pH Once these characteristics are translated into principal components, they can be plotted on a scatter plot to visually segregate and distinguish different plant species based on their unique characteristics.

In conducting PCA for plant classification as shown in Figure 6, the initial step entails a thorough identification and measurement of these characteristics to ensure the analysis effectively encapsulates the variation relevant to distinguishing between species. By focusing on this refined set of variables, researchers can enhance the efficiency and accuracy of their classification models, expediting the process of species identification and aiding in the comprehensive study of plant biodiversity. other traits.

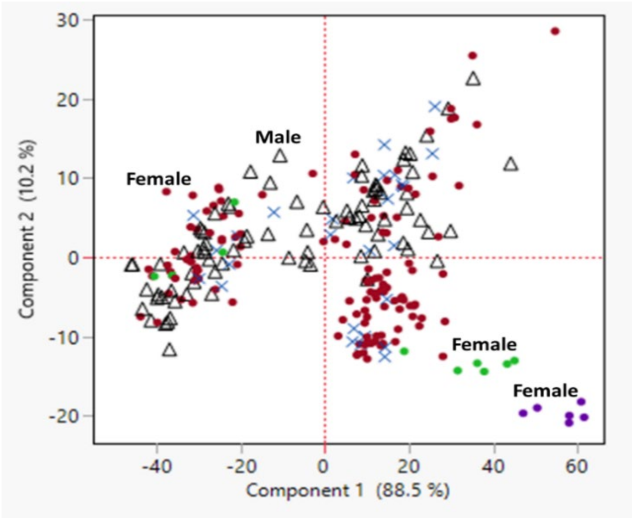

Figure 5. Classification based on NIR spectral data of *Calamus johndransfieldii* leaves samples using PCA method

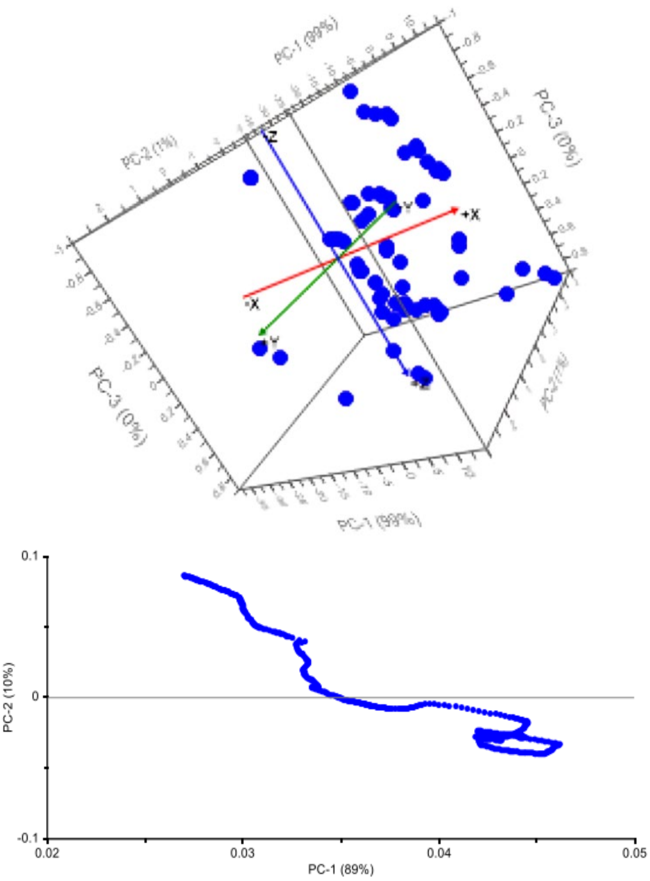

Figure 6. *Calamus johndransfieldii* plants distribution based on NIR spectral data in PCA scores and loadings

When applying PCA to the classification of *Calamus johndransfieldii* plants based on sex, the analysis has revealed a highly effective dimensional reduction, encapsulating a majority of the variance within just two principal components. Specifically, PC1 accounts for an impressive 88.5% of the variance, suggesting that this component captures the primary differences between the male and female plants. This substantial percentage indicates that the largest share of the detectable variation is attributable to a defining set of characteristics which might include a range of attributes such as physical traits, biochemical markers, or other sex-specific indicators. PC2 adds another 10.2% to the variance explanation, capturing secondary but still significant variations between the sexes that PC1 does not.

While this is a smaller proportion in comparison to PC1, the characteristics associated with PC2 provide additional nuanced details that further isolate the differences between the sexes. When these two components are taken together, they account for a staggering 98.7% of the total variation within the data, indicating that PC1 and PC2 combined effectively differentiate between male and female *Calamus johndransfieldii* with high clarity.

In a detailed PCA scatter plot as presented in Figure 5, this pronounced differentiation likely manifests as two distinct clusters along the PC1 and PC2 axes. Male and female plants would separate into these clusters, with the axis of PC1 delineating the primary and most pronounced sexual differences, and PC2 refining the distinction with more subtle traits. The distinction is so clear that, within the graph, minimal overlap would be expected between the clusters representing each sex. This PCA result is particularly valuable in areas such as selective breeding, horticultural studies, and conservation biology, offering a fast and non-destructive means of sex differentiation that can significantly streamline these plant-centric domains. The power of PCA in this context hinges upon its ability to convert a multivariate dataset into a visual snapshot where the complexities of plant physiology and gender specific attributes are reduced into a comprehensible form, vastly simplifying the task of biological classification.

Furthermore, sex gender classification of *Calamus johndransfieldii* plants using 4 different machine learning approaches is presented in Table 1. It provides the error rates for classifying male and female *Calamus johndransfieldii* plants using various classification methods. The error rates are as follows: Bootstrap Forest 0.0811, Neural Boosted 0.2047, Support Vector Machines 0.3243, and K Nearest Neighbors 0.3591 respectively.

Table 1. Classification error of *Calamus johndransfieldii* plants based on NIR spectral data using various machine learning approaches

Classification Method	Error Classification
Bootstrap Forest	0.0811
Neural Boosted	0.2047
Support Vector Machines	0.3243
K-Nearest Neighbors	0.3591

These error rates indicate the accuracy and reliability of each classification method in distinguishing between male and female plants. A lower error rate signifies a higher level of accuracy in classification. From the table, it is evident that the Bootstrap Forest method has the lowest error rate at 0.0811, indicating that it is the most accurate method for classifying male and female *Calamus johndransfieldii* plants in this context.

Conversely, the Support Vector Machines and K Nearest Neighbors methods have higher error rates, implying a lower level of precision in their classifications. Neural Boosted falls between these, showing moderate accuracy in comparison to the others.

The significance of these results lies in the practical application of these classification methods. The Bootstrap Forest method, with its lower error rate, would be preferred when precision in classifying male and female *Calamus johndransfieldii* plants is crucial [15]. In contrast, the Support Vector Machines and K Nearest Neighbors methods may be less suitable for such precise classifications based on the higher error rates.

Overall, this table aids in the selection of the most appropriate classification method for distinguishing between male and female *Calamus johndransfieldii* plants, providing guidance in choosing the method that best suits the desired level of accuracy for the specific classification task at hand.

This classification table provides essential insights into the practical application of NIRS technology for distinguishing and classifying the sexes of *Calamus johndransfieldii* plants. The varying error rates elucidate the effectiveness of different classification methods when utilizing NIRS data for this specific task.

The substantial superiority of the Bootstrap Forest classification method, indicated by its remarkably low error rate of 0.0811, underscores the potency and precision of NIRS data in differentiating between male and female plants. This suggests that NIRS holds significant discriminatory potential, enabling accurate gender classification, particularly when paired with the Bootstrap Forest classification algorithm as shown in Table 2 and Table 3.

Table 2. Classification result and accuracy based on Bootstrap Forest method

Samples	Actual	Predicted		
		Neutral	Female	Male
259 samples	Neutral	16	8	4
	Female	0	143	5
	Male	0	4	79

Table 3. Classification error based on Bootstrap Forest method

Error Classification		
Male	Neutral	Female
0.571	0.286	0.143
0.000	0.966	0.034
0.000	0.048	0.952

The provided table offers insights into the efficacy of a classification model in accurately predicting the gender of samples within three distinct categories: Neutral, Female, and Male. The majority of the samples are correctly classified, reflecting the model's overall strong performance. It is important to note that misclassifications exist within the data, suggesting areas where the model may face challenges in differentiation. Specifically, the model shows notable success in accurately predicting Female and Male samples, with few erroneous classifications, indicating high precision in these categories.

The misclassifications within the Neutral category, while present, appear relatively limited when considering the total number of neutral samples, suggesting a balanced level of tolerance for errors in this category. As such, while the model exhibits strong performance overall, understanding the patterns of misclassification will be crucial for refining the model and enhancing its accuracy in future predictions [16].

Conversely, the relatively higher error rates associated with other classification methods could indicate a varying degree of effectiveness when interpreting NIRS data for differentiating between the sexes of *Calamus johndransfieldii* plants. This variation may stem from the ability of each method to effectively capture and interpret the subtle spectral differences related to gender specific biological traits within the NIRS data.

Ultimately, these findings essentially highlight the accuracy and relevance of NIRS in enabling sex gender classification of *Calamus johndransfieldii* plants. The results findings emphasize the potential of NIRS technology application when combined with an appropriate classification method, such as Bootstrap Forest, to deliver precise, non-destructive, and rapid gender discrimination within botany, conservation efforts, and agricultural studies. This information is crucial for researchers and practitioners, offering a clearer understanding of the most effective approach when utilizing NIRS technology for plant gender classification [17, 18], which can consequently aid in improving plant breeding strategies, conservation efforts, and broader ecological studies.

In comparing the results of our study on plant sex differentiation using NIRS with existing research in this field, several key advantages and limitations emerge. While traditional methods such as morphological observations and genetic testing have been the standard approaches, the use of NIRS presents a novel and efficient alternative for early sex determination in dioecious plants like *Calamus johndransfieldii*. Our study demonstrated the effectiveness of this technique in accurately distinguishing between male and female plants at early growth stages, significantly reducing the time and resources required for sex differentiation. Similar attempts had been already applied in the utilization of NIRS spectral data to distinguish the male and female plants in date palm (*Phoenix dactylifera* L.). Date palm is a diploid plant (2x= 36) that has an extremely high divergence in sex at the flowering level. However, it is still classified in Arecaceae. The plant itself has been regarded as one of the most economically important trees of the arid regions, for example in the Middle East and North [19].

Over the past few years, the genetic diversity of date palm trees as well as their breeding methodologies are being threatened by exposure to a wide array of environmental and biological pressures. Thus, such challenges have led to investment in highly tolerant varieties through various intensive breeding programs. However, such programs are mostly time-consuming, not only in terms of plant growth, but also in the selection process as well as differentiation of male and female germplasms for further cultivation and propagation. Therefore, identifying female plants, which are responsible for fruit production at an immature stage, would certainly enhance the uniformity in crossing, The Near-Infrared Reflectance Spectroscopy (NIRS) has been applied in discriminating between sexes in date palm trees through vibrational and nuclear magnetic resonance spectroscopy. In general, NIRS could measure the chemical structure and composition of a sample based on the absorption of near-infrared radiation by CO-H and N-H bounds, resulting in overtones and variations on bands that can be observed at wavelengths of 780-2,500 nm [20]. Being a robust tool, there were already few examples of the use of NIRS to carry out phenotypic, metabolomic, and physiological assessments of trees [21].

Advantages observed in our study include the non-destructive and rapid nature of NIRS analysis, enabling early identification of plant sexes without the need to wait for reproductive maturity. The utilization of advanced data analysis methods such as Principal Component Analysis (PCA) and Bootstrap Forest contributed to robust and reliable sex classification models. However, limitations such as variability in spectral data interpretation and potential environmental factors influencing spectral readings should be considered in future research and application.

Discussing the environmental and ecological considerations of widespread adoption of NIRS technology in forestry and agriculture is crucial. Positive impacts could include increased efficiency in plant breeding programs, enhanced biodiversity conservation efforts through accurate sex ratio monitoring, and improved sustainability in non-timber forest product industries [20, 21]. By facilitating targeted cultivation of female plants, the technology could contribute to better seed production and genetic diversity conservation.

On the other hand, negative consequences may arise if the technology leads to monoculture practices or neglect of male plants, potentially disrupting natural pollination dynamics and ecosystem balance. Careful monitoring and management strategies should be implemented to mitigate any adverse effects on plant populations and their habitats. It is essential to conduct further research on the long-term ecological implications of widespread NIRS application to ensure that the benefits of enhanced productivity do not compromise the overall health and resilience of forest ecosystems. Balancing technological advancements with environmental stewardship is key to realizing the full potential of innovative tools like NIRS in sustainable forestry and agriculture practices.

4. CONCLUSION

Based on all the study findings, it is clear that there are various methods available for differentiating between male and female *Calamus johndransfieldii* plants. Near infrared spectroscopy (NIRS) has shown promise in accurately classifying different plant species based on their spectral patterns. Additionally, principal component analysis (PCA) has proven to be effective in condensing complex datasets, spotlighting the most significant variation among plant data. The error rates provided illustrate the efficacy of different classification models, with some demonstrating a higher accuracy in distinguishing between male and female specimens.

These research findings provide a comprehensive understanding of the potential for accurate gender classification within *Calamus johndransfieldii* plants. The utilization of NIRS technology, PCA, and various classification methods like bootstrap forest method enables a multifaceted and precise approach to differentiating between male and female specimens, a capability that can significantly impact botanical research, conservation efforts, and agricultural practices.

Through this holistic approach, accurate gender classification can support the development of tailored breeding programs, the preservation of plant biodiversity, and the enhancement of ecosystem health. These findings collectively signify an important advancement in plant classification methodologies and lay a foundation for informed decision-making in botany, conservation, and ecology.

From a practical standpoint, the implications of our findings are profound. The ability to quickly and accurately differentiate between male and female plants has direct implications for enhancing agricultural productivity, improving breeding programs, and optimizing conservation efforts for biodiversity. The rapid and cost-effective nature of the NIRS technique can empower farmers, conservationists, and forest resource managers to make informed decisions that promote sustainable land use practices and ensure the long-term health of plant populations. As we look to the future, the widespread adoption of NIRS technology holds promise for transforming not only how we manage and cultivate dioecious plants but also how we approach biodiversity conservation and ecological sustainability.

By integrating innovative scientific methodologies with practical applications, our study paves the way for a more efficient and effective approach to plant sex determination, with implications that extend beyond the boundaries of our research to shape the future of agriculture, forestry, and environmental stewardship.

ACKNOWLEDGEMENTS

Funding for this article is the Indonesia Endowment Fund for Education or eRISPRO LPDP with the National Research and Innovation Agency (BRIN). LPDP is an abbreviation from "*Lembaga Pengelola Dana Pendidikan*" and is a public service agency of the Ministry of Finance of Indonesia (https://risprolpdp.kemenkeu.go.id/) under the scheme "*Riset dan Innovasi untuk Indonesia Maju* (RIIM)" for the 2nd Batch, for the fiscal year fund: 2022-23; ID Contract No.: B-1738/II.7.5/FR/11/2022 and No.: B-2699/III.5/PR.03.08/11/2022. The grant title was "Pengembangan Industri Biofarmaka Berbasis Resis Rotan Jernang (*Dragon blood*) Tanaman Endemik Sumatera."

REFERENCES

[1] Wahyudi, A., Jannetta, S. (2011). Potensi dan permudaan alam rotan penghasil jernang di kawasan Taman Nasional Bukit Tigapuluh, Riau. Jurnal Penelitian Hutan dan Konservasi Alam, 8(3): 237-243.

[2] Asra, R., Syamsuardi, S., Mansyurdin, M., Witono, J.R. (2014). The study of genetic diversity of Daemonorops draco (Palmae) using ISSR markers. Biodiversitas Journal of Biological Diversity, 15(2): 109-114. https://doi.org/10.13057/biodiv/d150201

[3] Matangaran, J.R., Puspitasari, L. (2012). Potential and harvesting of Jernang Rattan fruit. Jurnal Silvikultur Tropika, 3(1): 65-70.

[4] Baumer, U., Dietemann, P. (2010). Identification and differentiation of dragon's blood in works of art using gas chromatography/mass spectrometry. Analytical and Bioanalytical Chemistry, 397: 1363-1376. https://doi.org/10.1007/s00216-010-3620-0

[5] Shi, J., Hu, R., Lu, Y., Sun, C., Wu, T. (2009). Single-step purification of dracorhodin from dragon's blood resin of Daemonorops draco using high-speed counter-current chromatography combined with pH modulation. Journal of Separation Science, 32(23-24): 4040-4047. https://doi.org/10.1002/jssc.20090039

[6] Sari, R.W., Hikmat, A., Santosa, Y. (2015). Production estimation dragon's blood (*Daemonorops didymophylla* Becc.) based morphometric characteristics rattan. Media

Konservasi, 20(2): 140-148.
[7] Henderson, A. (2020). A revision of *Calamus* (Arecaceae, Calamoideae, Calameae, Calaminae). Phytotaxa, 445(1): 1-656. https://doi.org/10.11646/phytotaxa.445.1.1
[8] Rustiami, H., Setyowati, F.M., Kartawinata, K. (2004). Taxonomy and uses of Daemonorops draco (Willd.) Blume. Journal of Tropical Ethnobiology, 1(2): 65-75.
[9] Sahu, A., Sen, S., Mishra, S.C. (2020). Processing and properties of Calotropis gigantea bio-char: A wasteland weed. Materials Today: Proceedings, 33: 5334-5340. https://doi.org/10.1016/j.matpr.2020.03.024
[10] Zhao, H.T., Feng, Y.Z., Chen, W., Jia, G.F. (2019). Application of invasive weed optimization and least square support vector machine for prediction of beef adulteration with spoiled beef based on visible near-infrared (Vis-NIR) hyperspectral imaging. Meat Science, 151: 75-81.
[11] Liang, D., Xia, C., Huang, H., Liu, Y., Ma, Z., Li, S., Zhang, Q., Meng, Z. (2023). Weed control and slow-release behavior of 2-methyl-4-chlorophenoxyacetate intercalated layered double hydroxide. Colloids and Surfaces A: Physicochemical and Engineering Aspects, 658: 130661. https://doi.org/10.1016/j.colsurfa.2022.130661
[12] Milojković, J.V., Lopičić, Z.R., Anastopoulos, I.P., Petrović, J.T., Milićević, S.Z., Petrović, M.S., Stojanović, M.D. (2019). Performance of aquatic weed-Waste Myriophyllum spicatum immobilized in alginate beads for the removal of Pb (II). Journal of Environmental Management, 232: 97-109. https://doi.org/10.1016/j.jenvman.2018.10.075
[13] Su, W.H. (2020). Crop plant signaling for real-time plant identification in smart farm: A systematic review and new concept in artificial intelligence for automated weed control. Artificial Intelligence in Agriculture, 4: 262-271. https://doi.org/10.1016/j.aiia.2020.11.001
[14] González-Martín, I., Hernández-Hierro, J.M., Revilla, I., Vivar-Quintana, A., Ortega, I.L. (2011). The mineral composition (Ca, P, Mg, K, Na) in cheeses (cow's, ewe's and goat's) with different ripening times using near infrared spectroscopy with a fibre-optic probe. Food Chemistry, 127(1): 147-152. https://doi.org/10.1016/j.foodchem.2010.12.114
[15] Sun, W., Zhang, X., Sun, X., Sun, Y., Cen, Y. (2018). Predicting nickel concentration in soil using reflectance spectroscopy associated with organic matter and clay minerals. Geoderma, 327: 25-35. https://doi.org/10.1016/j.geoderma.2018.04.019
[16] Pasquini, C. (2018). Near infrared spectroscopy: A mature analytical technique with new perspectives–A review. Analytica Chimica Acta, 1026: 8-36. https://doi.org/10.1016/j.aca.2018.04.004
[17] Duncan, L., Miller, B., Shaw, C., Graebner, R., Moretti, M.L., Walter, C., Selker, J., Udell, C. (2022). Weed warden: A low-cost weed detection device implemented with spectral triad sensor for agricultural applications. HardwareX, 11: e00303. https://doi.org/10.1016/j.ohx.2022.e00303
[18] Devi, C., Khwairakpam, M. (2023). Weed biomass: Bioconversion through composting followed by vermicomposting to optimize time required. Bioresource Technology Reports, 21: 101326. https://doi.org/10.1016/j.biteb.2022.101326
[19] Ventura, M., Silva, J.R., Andrade, L.H.C., Junior, R.S., Lima, S.M. (2018). Near-near-infrared thermal lens spectroscopy to assess overtones and combination bands of sulfentrazone pesticide. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 188: 32-36. https://doi.org/10.1016/j.saa.2017.06.043
[20] Khan, A.L., Al-Harrasi, A., Numan, M., AbdulKareem, N.M., Mabood, F., Al-Rawahi, A. (2021). Spectroscopic and molecular methods to differentiate gender in immature date palm (Phoenix dactylifera L.). Plants, 10(3): 536. https://doi.org/10.3390/plants10030536
[21] Warton, D.I., Blanchet, F.G., O'Hara, R.B., Ovaskainen, O., Taskinen, S., Walker, S.C., Hui, F.K. (2015). So many variables: Joint modeling in community ecology. Trends in Ecology & Evolution, 30(12): 766-779. https://doi.org/10.1016/j.tree.2015.09.007