UNIVERSITY OF CAPE COAST

INNOVATIVE APPLICATIONS OF HANDHELD NEAR-INFRARED SPECTROSCOPIC TECHNOLOGY FOR QUALITY ASSESSMENT OF FRUITS AND FRUIT PRODUCTS IN GHANA

FRANCIS PADI LAMPTEY

2024

UNIVERSITY OF CAPE COAST

INNOVATIVE APPLICATIONS OF HANDHELD NEAR-INFRARED SPECTROSCOPIC TECHNOLOGY FOR QUALITY ASSESSMENT OF FRUITS AND FRUIT PRODUCTS IN GHANA

BY

FRANCIS PADI LAMPTEY

(AG/FPT/20/0004)

Thesis submitted to the Department of Agricultural Engineering of the School of Agriculture, College of Agriculture and Natural Sciences, University of Cape Coast, in partial fulfilment of the requirements for the award of Doctor of Philosophy degree in Food and Postharvest Technology.

OCTOBER, 2024

DECLARATION

Candidate's Declaration

I hereby declare that this thesis is the result of my own original research and that no part of it has been presented for another degree in this university or elsewhere.

Candidate's Signature …………………………… Date ……………

Name: Francis Padi Lamptey

Supervisors' Declaration

We hereby declare that the preparation and presentation of the thesis were supervised in accordance with the guidelines on supervision of thesis laid down by the University of Cape Coast.

Principal Supervisor's Signature ………………… Date ……………

Name: Rev. Engr. Prof. Ernest Teye

Co-Supervisor's Signature …………………… Date ……………

Name: Prof. Ernest Ekow Abano

ABSTRACT

Handheld near-infrared spectroscopy (NIRS) is emerging as a key technology for food analysis in Africa. This study explores its effectiveness, combined with chemometric techniques, for rapid and non-destructive evaluation of mango fruits and products. It focuses on developing predictive models for variety differentiation, classification of organic and inorganic samples, and assessing quality attributes such as total soluble solids (TSS) and pH. Additionally, it identifies ethephon residues and categorizes organic and inorganic pineapple juice. The study also examines the physicochemical and microbial changes in expired and unexpired commercial fruit juices using conventional laboratory methods. For mango variety identification, NIRS combined with multivariate algorithms achieved 97.44% accuracy. The synergy partial least squares model yielded r^2 values of 0.63 and 0.81 for TSS and pH predictions, with RMSEP values of 1.83 and 0.49, respectively. In detecting ethephon residues, the neural network model with multiplicative scatter correction reached 100% classification accuracy, while the partial least squares model demonstrated strong predictive performance (r^2 = 0.996, RMSEP = 0.068). The random forest algorithm classified organic and inorganic mango products with varying accuracy levels. When preprocessed using the second derivative, it achieved 88.76% accuracy for fresh fruit, 77.98% for chips, and 87.53% for juice without preprocessing. The combination of dual NIR spectrometers effectively distinguished organic and inorganic pineapple juice with 100% accuracy. Furthermore, a comparative assessment of expired and unexpired commercial fruit juices showed notable declines in titratable acidity (apple juice decreased from 0.60% to 0.12%) and vitamin C (a 57.6% reduction in pineapple juice), alongside an increase in microbial load. These findings highlight the potential of handheld NIRS as a reliable tool for quality control, product authentication, and food safety assurance. Its application could improve postharvest monitoring, mitigate food fraud, and enhance regulatory compliance within the fruit industry.

LIST OF PUBLICATIONS BY CANDIDATE

1. Lamptey, F. P., Teye, E., Abano, E. E., & Amuah, C. L. (2023). Application of handheld NIR spectrometer for simultaneous identification and quantification of quality parameters in intact mango fruits. *Smart Agricultural Technology*, *6*, 100357.

2. Lamptey, F. P., Amuah, C. L., Boadu, V. G., Abano, E. E., & Teye, E. (2024). Smart classification of organic and inorganic pineapple juice using dual NIR spectrometers combined with chemometric techniques. *Applied Food Research*, 100471.

3. Lamptey, F. P., Teye, E., Kaburi, S. A., Odoi-Yorke, F., Amuah, C. L. Y., Abano, E. E., & Otoo, G. S. (2025). Feasibility study on fingerprinting organic and conventional mango fruits, chips, and juice using portable near-infrared spectroscopy. *Analytical Methods*.

4. Lamptey, F. P., Teye, E., Odoi-Yorke F., Otoo G. S., Abano, E. E., Amuah C. L. Y., & Kaburi, S. A. (2024). A review of recent trends, advances, and future directions of near-infrared spectroscopy in fruit analysis. Infrared Physics and Technology (Under Review).

ACKNOWLEDGMENT

I am grateful to the Almighty God for the grace and good health He gave me each day to complete this study. I would like to express my sincere gratitude to Rev. Engr. Prof. Ernest Teye, who served as my main supervisor and advisor, for his positive feedback, outstanding direction, suggestions, and constant encouragement throughout my time at the University. He is very knowledgeable, organized, and competent, all of which have been helpful to me. I also want to thank my co-supervisor, Prof. Ernest Ekow Abano, for his support and encouragement.

I also would like to express my gratitude for the support and inspiration I received from Rev. Ing. Prof J. D. Owusu-Sekyere, Prof. L. K. Sam-Amoah, Prof. R. S. Amoah, Dr. Ransford Darko, Dr. Francis Kumi, and Dr. (Mrs) Rosemond Dadzie, as well as the entire staff of the Department of Agricultural Engineering. I appreciate the help of Dr. Charles L.Y. Amuah, Dr. Sandra Ama Kaburi, Dr. Mrs. Gifty Serwaa Otoo, Lady Winifred Teye, Regina Nyokeh, Gideon Kwesi Nakotey, Emmanuel Agyei, Joel Welbeck, Collins Asare, and all the Royal Basilica family for their valuable assistance throughout the research work. I acknowledge the support provided by the World Food Program, CSIR-Food Research Institute, Ghana, RUFORUM, UK Innovation and Research, and Cape Coast Technical University.

I would like to express my sincere gratitude to my wife, Judith Awudi (Mrs), for her constant encouragement, perseverance, prayers, and encouragement.

DEDICATION

I dedicate this work to my beloved Wife, Son, and Sister, Judith Awudi (Mrs), Jedidiah Royal Inie Lamptey, and Francisca Padikie Lamptey.

TABLE OF CONTENTS

DECLARATION	ii
ABSTRACT	iii
LIST OF PUBLICATIONS BY CANDIDATE	iv
ACKNOWLEDGMENT	v
DEDICATION	vi
TABLE OF CONTENT	vii
LIST OF TABLES	xiii
LIST OF FIGURES	xv
LIST OF ACRONYMS	xix
CHAPTER ONE: INTRODUCTION	
1.1 Background to the Study	1
1.2 Problem Statement	5
1.3 Justification	7
1.4 Objectives of the Study	9
1.4.1 General Objective	9
1.4.2 Specific Objectives	9
1.5 Research Questions	10
1.6 Linkages	10
1.7 Organization of the Thesis	10
CHAPTER TWO: LITERATURE REVIEW	
Abstract	14
2.1 Introduction	15
2.2. Methodology	20
2.3. Results and Discussion	23

2.3.1Summary of Bibliometric Data and Annual Article Production 23

2.4. Countries' Distribution and Collaborations 26

2.5. Analysis of Word Cloud and Trend Topics 32

2.6. Co-Occurrence Network, Thematic Map, and Thematic Evolution 37

2.7. Journals, Institutions, and Citations 45

2.8. Review of the Top 50 Most Relevant Papers on NIRS in Fruit Analysis 50

2.9 Advances in NIRS in Mango Fruit Analysis 58

2.9.1 Mango Quality Assessment and Prediction using NIRS 58

2.9.2. Non-Destructive Testing and Monitoring 61

2.9.3. Spectral Data Analysis and Model Optimisation 64

2.9.4. Technological Innovations and Instrumentation 67

2.9.5. Applications of Machine Learning and Artificial Intelligence 71

2.10. Summary and Directions for Future Research 73

2.11. Conclusion 78

CHAPTER THREE: APPLICATION OF HANDHELD NIR SPECTROMETER FOR SIMULTANEOUS IDENTIFICATION AND QUANTIFICATION OF QUALITY PARAMETERS IN INTACT MANGO FRUITS. 80

3.1 Introduction 84

3.2. Materials and Methods 86

3.2.1. Mango Fruit Samples 86

3.2.2. NIR Spectroscopy Measurement 87

3.2.3. Reference Method (TSS/°Brix and pH) 87

3.3. Statistical Analysis 87

3.3.1. Data Partition 87

3.4 Software Device 88

3.5 Spectral Data Preprocessing 88

3.6. Principal Component Analysis (PCA) 89

3.7. Identification Model 89

3.7.1 Quantification Model 91

3. 8 Results 92

3.8.1 Reference Measurement and Data Partitioning 92

3.8.2. Spectra Profile 93

3.8.3. Principal Component Analysis 93

3.8.4. Performance of the Identification Models. 95

3.8.5. Optical Variable Selection 96

3.8.6. Comparative Performance of TSS and pH Models 96

3.9 Discussion 97

3.10 Conclusion 115

CHAPTER FOUR: RAPID AND NON-INVASIVE PREDICTION OF ETHEPHON LEVELS IN MANGO FRUITS USING NEAR-INFRARED SPECTROSCOPY

4.1 Introduction 120

4.2 Methodology 122

4.2.1 NIR Spectra Acquisition 123

4.2.2 Statistical Analysis, Data Processing and Analysis 124

4.2.3 Chemometric Analysis 124

4.3 Results and Discussions 125

4.3.1 Spectral Analysis and Preprocessing 125

4.3.2 PCA (Principal Component Analysis) 127

4.3.3 Classification Models 129

4.3.4 Quantitative Prediction of Ethephon Concentration 130

4.4 Conclusion 131

CHAPTER FIVE: FEASIBILITY STUDY ON FINGERPRINTING ORGANIC AND CONVENTIONAL MANGO FRUITS, CHIPS, AND JUICE USING PORTABLE NEAR-INFRARED SPECTROSCOPY 135

5.1. Introduction 137

5.1.Methodology 140

5.2.1. Sample Collection 140

5.2.2. Spectra Collection 140

5.2.3. Initial Spectra Data Processing 141

5.2.4. Principal Component Analysis 142

5.2.5. Data Partitioning 142

5.2.6. Multivariate Classification Algorithms 143

5.3. Results and Discussions 144

5.3.1. Spectra Analysis 144

5.3.2. Principal Components 149

5.3.3. Classification Models for Organic and Conventional Mango Products 158

5.4. Conclusion 166

CHAPTER SIX: SMART CLASSIFICATION OF ORGANIC AND INORGANIC PINEAPPLE JUICE USING DUAL NIR SPECTROMETERS COMBINED WITH CHEMOMETRIC TECHNIQUES.

Abstract 170

6.1 Introduction 171

6.2. Materials and Methods 173

6.2.1 Pineapple Juice Preparation 173

6.2.2 Reference Measurement (TSS/°Brix and pH). 174

6.2.3 Instrumentation 174

6.2.4 Spectra Acquisition 176

6.2.5 Software Device, Data Partitioning, and Spectra Preprocessing 177

6.2.6 Partial Least Squares Discriminant Analysis (PLS-DA) for 178

6.2.7 Optimal Variable Selection 179

6.2.8 Evaluation of Model Performance 181

6.3. Results And Discussions 181

6.3.1 Reference Measurement 181

6.3.2 Analysis of NIR Spectra. 182

6.3.3 Classification using PLS-DA Model 189

6.3.4 Variable Importance Projection Model 200

6.4. Conclusion 202

CHAPTER SEVEN: COMPARATIVE ANALYSIS OF EXPIREDAND UNEXPIRED COMMERCIAL FRUIT JUICES: PHYSICOCHEMICAL AND MICROBIAL PROPERTIES. 204

Abstract 206

7.1 Introduction 207

7.2 Methodology 208

7.2.1 Sample Collection 208

7.2.2 Total Soluble Solids, pH, and Titratable Acidity Determination 209

7.2.3 Determination of Vitamin C 210

7.2.4 Determination of Colour 210

7.2.5 Microbial Analysis 211

7.3. Results and Discussions 212

7.3.1 Physicochemical Changes 213

7.3.1.1 Titratable Acidity (TA) and pH 213

7.3.1.2 Total Soluble Solids (TSS) 214

7.3.1.3 Colour Parameters (L*, a*, b*) 214

7.3.1.4 Vitamin C Content 214

7.3.1.5 Microbial Changes 215

7.4 Conclusion 219

CHAPTER EIGHT: SUMMARY, GENERAL CONCLUSIONS AND RECOMMENDATION 220

8.1 Summary 220

8.1.1 Key Findings 221

8.2 Conclusion 223

8.3 Recommendation 225

REFERENCE LIST 227

LIST OF TABLES

Table		Page
3.1	Reference Measurements of TSS and pH	101
3.2	Performance of Identification Model for Mango Varieties at 3PCs	109
3.3	Performance of SiPLS Model with some Selected Optimal Spectra Regions	110
3.4	Efficient Variable Selection Performance Comparative Models for TSS and pH	112
4.1	Performance of Spectral Pre-Processing Techniques During Model Development and Cross Validation	132
4.2	Prediction of Concentration of the Ethephon	133
5.1	Performance of Classification Models in Differentiating Organic Mango Fruits from Conventional Ones using	161
5.2	Performance of Classification Models in Differentiating Organic Mango Chips from Conventional Ones using	162
5.3	Performance of Classification Models in Differentiating Organic Mango Juice from Conventional Ones using Pre-Processing Techniques with 10-Fold Cross-Validation for both Training and Test Sets	163
6.1	Reference Measurement of pH and TSS	191
6.2	Training and Testing Results of the PLS-DA Model Using Spectral Data from the Handheld NIR Spectrometer (SCIO^{TM}) with Various Preprocessing Methods.	192

6.3 Training and Testing Results of the PLS-DA Model using Spectral Data from the Handheld NIR Spectrometer (Tellspec) with Various Preprocessing Methods. 193

6.4 Training and Testing Results of the PLS-DA Model using Combined Spectral Data from Handheld NIR Spectrometers (Tellspec and SCiO^{TM}) for Classification. 194

7.1 Details of Expired and Unexpired Commercial Fruit Juice 209

7.2 Physicochemical Properties and Microbial Load of Expired and Unexpired Apple Juice Samples (Sample 1) 216

7.3 Physicochemical Properties and Microbial Load of Expired and Unexpired Grape Juice Samples (Sample 2) 216

7.4 Physicochemical Properties and Microbial Load of Expired and Unexpired Pineapple Juice Samples (Sample 3) 217

7.5 Physicochemical Properties and Microbial Load of Expired and Unexpired Apple Juice Samples (Sample 4) 218

LIST OF FIGURES

Figure		Page
2.1	Study Framework Employed for Bibliometric Analysis	23
2.2	Summary of Bibliometric Data	24
2.3	Annual Scientific Production	26
2.4	Country Scientific Production	27
2.5	Country Collaboration Map	30
2.6	Corresponding Author's Countries	32
2.7	Word Cloud (50 words)	35
2.8	Trend Topics from 2003 - 2023	37
2.9	Keywords Co-occurrence Network	40
2.10	Thematic Map	43
2.11	Thematic evolution	45
2.12	Top 20 Most Relevant Journals Based on Total Publications	47
2.13	Top 20 Most Relevant Institutions Based on Total Publications	48
2.14	Total Citations per Country	50
2.15	(a) NIRQuest spectrometer (b) SCIO NIR spectrometer with the three varieties of mango fruits	57
3.1	General set up for Scanning and the Three Mango Varieties.	101
3.2	Spectra of RAW (a), MC (b), MSC (c), SNV (d), FD (e) and SD (f)	103
3.3	Mean spectra profile of Haden, Keitt and Local fruits.	104
3.4	PCA Score Plot of Spectra Data. RAW (a), Preprocessed with MC (b), Preprocessed with MSC (c), Preprocessed with SNV (d), Preprocessed with FD (e), and Preprocessed with SD (f), SD	

with PCA Technique Gave a Separation with Clear Cluster Trend. 107

3.5 PC Loadings with Three Principal Components 108

3.6 Optimal Selection and Combination of Spectra Region by SiPLS. 111

3.7 Scatter Plot of Reference Measured Versus Handheld NIR Predicted (a) TSS and (b) pH by SiPLS Model. 115

4.1 Experimental set-up 123

4.2 Raw and mean Spectral profile of naturally ripened mangoes and artificially treated mangoes with 250, 750, and 1000 ppm ethephon. 127

4.3 PCA plot of the topmost three PCs for naturally ripened mangoes and artificially treated mangoes with 250, 750, and 1000 ppm ethephon 129

5.1 Flowchart illustrating the classification process for distinguishing 144

5.2 Raw Spectral Profile of Organic and Conventional Mango (A) Fruits 147

5.3 Mean Spectral Profile of Organic and Conventional Mango (A) Fruits 148

5.4 Comparison of Mean Spectra Profile of Organic and Conventional 149

5.5 PCA Score Plot of the First Three PCs of Organic and Conventional 152

5.6 PCA Score Plot of the First Three PCs of Organic and Conventional Mango Chips Pre-Processed- a) RAW b) MSC c) SNV d) FD e) SD. 153

5.7 PCA Score Plot of the First Three PCs of Organic and Conventional Mango Juice Pre-Processed- a) RAW b) MSC c) SNV d) FD e) SD. 155

5.8 (A). Principal Component Loadings with Three Principal Components for Mango Fruits 155

5.8 (B). Principal Component Loadings with Three Principal Components for Mango Chips 156

5.8 (C). Principal Component Loadings with Three Principal Components for Mango Juice 157

6.1 Schematic of the Procedure for Obtaining Spectra of Juice (from 176

6.2 Spectra of Pineapple Juice Based on Handheld NIR SCiO ™ (740 - 1070 nm): (a) Raw, (b) Mean, (c) MC (d) MSC, (e) SNV, (f) FD, and (g) SD. 186

6.3 Spectra of Pineapple Juice Based on Handheld NIR Tellspec (900–1700 nm): (a) Raw, (b) Mean, (c) MC, (d) MSC, (e) SNV, (f) FD, and (g) SD. 189

6.4 Classification of Organic and Inorganic Pineapple Juice Utilizing PLS-DA with FD Preprocessing Based on Handheld SCiO ™ (740 - 1070 nm): (a) Training and (b) Testing. 195

6.5 Classification of Organic and Inorganic Pineapple Juice Utilizing PLS-DA with SD Preprocessing Based on Tellspec: (a) Training (b) Testing. 196

6.6 Classification Results for Organic and Inorganic Pineapple Juice Utilizing PLS-DA with RAW Preprocessing Based on Combining Handheld Tellspec and SCiO™: (a) Training and (b) Testing. 197

6.7 ROC Curve for Pineapple Juice with FD Preprocessing of NIR Spectra Based on Handheld NIR SCiO ^{TM} (740 - 1070 nm (a) Training and (b) Testing. 198

6.8 ROC Curve for Pineapple Juice with SD Preprocessing of NIR Spectra Based on Tellspec: (a) Training (b) Testing. 199

6.9 ROC Curve for Pineapple Juice with Raw Preprocessing of NIR Spectra Based on the Combination of Handheld Tellspec and SCiO^{TM}: (a) Training and (b) Testing. 199

6.10 VIP Calculated Utilizing SCiO^{TM}. The Line Indicates Threshold 201

6.11 VIP Calculated Utilizing Tellspec. The Line Indicates the Threshold 201

6.12 VIP Calculated Utilizing a Combination of both Tellspec and 202

7.1 Samples of fruit juice: A. Expired apple juice, B. Unexpired apple juice, C. Unexpired grape juice D. Expired grape juice E. Unexpired Pineapple juice, F. Expired Pineapple juice, G. Unexpired Apple juice, and H. Expired Apple juice 212

LIST OF ACRONYMS

ANN	-	Artificial Neural Networks
APC	-	Aerobic Plate Count
AUC	-	Area Under Curve
Bi-PLS	-	Back Interval Partial Least Squares
CFU/mL	-	Colony Forming Units per Millilitre
FD	-	First Derivative
Ga- PLS	-	Genetic Algorithm Interval Partial Least Square
IPLS	-	Interval Partial Least Squares
KNN	-	K-Nearest Neighbors
LDA	-	Linear Discriminant Analysis
LDA-SVM	-	Linear Discriminant Analysis - Support Vector Machine
LS-SVM	-	Least squares - support vector machines
LSSVR	-	Least-squares support vector regression
RBF	-	Radial basis function
MC	-	Mean centering
MSC	-	Multiplicative Scatter Correction
NIR	-	Near Infrared
NN	-	Neural Network
PCA	-	Principal Component Analysis
pH	-	Potential of Hydrogen
PLS	-	Partial Least Squares
PLSR	-	Partial Least Square Regression
PLS-DA	-	Partial Least Squares Discriminant Analysis
r^2	-	Coefficient of Determination

RF	-	Random Forest
ROC	-	Receiver Operating Characteristic
RPD	-	Residual Prediction Deviation
RMSEP	-	Root Mean Square Error of Prediction
SD	-	Second Derivative
SMOTE	-	Synthetic minority oversampling technique
SWNIR	-	Shortwave near infrared
Si-PLS	-	Synergy Partial Least Squares
SNV	-	Standard Normal Variate
SVM	-	Support Vector Machine
TSS	-	Total Soluble Solids
VIP	-	Variable Importance in Projection

CHAPTER ONE

INTRODUCTION

1.1 Background to the Study

Fresh fruits are rich in water, minerals, vitamins, and fibre, all contributing significantly to human and animal health. Fruit consumption reduces the risk of various illnesses and functional decline linked with aging (Emelike & Akusu, 2019). These fruits can be processed into value-added products, including fresh fruit, puree, slices in syrup, leather, canned slices, chutney, juice concentrate, ready-to-drink juice, wine, jams, jellies, pickles, smoothies, chips, and powder (Jahurul et al., 2015; Owino & Ambuko, 2021). Fruit juice is widely regarded and promoted as a low-fat, often natural, healthful beverage. To assure quality and safety, quality evaluation and assurance of fruits are essential before export, during processing, and in the fresh market. Internal quality characteristics such as total soluble solids (TSS), firmness, and acidity are commonly used. However, TSS (°Brix) has been identified as the most important internal quality indicator. TSS, for example, is a critical internal quality criterion for predicting fruit maturity and harvesting time and analyzing and grading postharvest quality fruits. Mango fruits are noted to be one of the important tropical fruits grown and consumed worldwide and a rich source of nutrients and phytochemicals, including vitamin C and β- carotene.

The deficiency of β- carotene, a precursor to the biosynthesis of vitamin A, is a major challenge faced by Sub-Saharan African countries. Therefore, mango consumption could relatively be an affordable strategy to supply β- carotene to alleviate vitamin A deficiency in Sub-Saharan Africa

(Ntsoane, Zude-Sasse, Mahajan, & Sivakumar, 2019). Consumption of Mango fruits and juice could promote the attainment of sustainable development goals 2 and 3, zero hunger, and good health and well-being. However, safety measures such as checking the expiration dates of fruit juice are crucial.

The chemical composition of mangoes varies depending on the location of cultivation, variety, and stage of ripeness (Tharanathan, Yashoda, & Prabha, 2006). A key factor influencing mango quality and shelf life is maturity at harvest, which significantly affects flavor, texture, and susceptibility to postharvest physiological disorders (Kader, 1997). Mangoes must be harvested at the appropriate maturity stage to ensure optimal fruit quality. Immature fruits tend to have a lower-quality flavor, are more prone to mechanical damage, and often do not ripen evenly. Overripe fruits, on the other hand, have a shorter shelf life and are more susceptible to microbial spoilage (Jha, Chopra, & Kingsly, 2007).

Some popular varieties of mangoes include Keitt, Haden, Palmer, Parwin, Tommy Atkins, Kent, Alphonso, Benishaan, Kesar, Chausa, Dasheri, Langra, Malda, and local (Abu, Abbey, & Amey, 2021; Jha et al., 2013; Souza, Leonel, Modesto, Ferraz, & Gonçalves, 2018). These varieties are available in various colours, sizes, and shapes (Jha et al., 2013). Methods used to determine the internal quality of mangoes and varieties are destructive, labor-intensive, and often expensive, necessitating the development of faster, non-destructive alternatives (Cortés et al., 2016; Jha et al., 2013).

Non-destructive methods such as near-infrared spectroscopy (NIRS) have gained traction in recent years for assessing fruit quality. NIRS is a rapid, non-invasive, and reliable technology successfully applied to evaluate various internal quality parameters of fruits, such as TSS and pH. This technology measures light absorption in the near-infrared region (700–2500 nm), providing valuable information about the fruit's internal composition without damaging the sample. It has been employed in numerous studies for quality assessment of pineapples, cocoa, strawberries, and mangoes, making it a promising tool for the food industry (Amodio, Ceglie, Chaudhry, Piazzolla, & Colelli, 2017; Amuah et al., 2019; Cortés et al., 2016; Teye, Anyidoho, Agbemafle, Sam-Amoah, & Elliott, 2020).

Mango, a climacteric fruit, undergoes significant physiological changes during ripening. Artificial ripening agents like ethephon and calcium carbide are widely used to induce fruit ripening. However, these chemicals pose health risks, particularly calcium carbide, which has been linked to neurological disorders and seizures due to contamination with arsenic and phosphorus (Siddiqui & Dhua, 2010). While ethephon is considered safer when used within permissible limits, regulatory bodies in many African countries struggle to enforce proper usage standards, leading to potential food safety concerns (Islam, Mursalat, & Khan, 2016; Ruwali, Thakuri, Pandey, Mahat, & Shrestha, 2022). Traditional methods for detecting artificial ripening agents like ethephon are destructive and require complex laboratory setups, making them unsuitable for large-scale or field-based applications. NIRS presents a solution to this problem by providing a rapid, non-destructive means of predicting ethephon concentrations in fruits. Recent advancements in NIRS technology have led to the development of portable, battery-powered devices that can be used on-site for quality control and regulatory compliance (Mahanti & Chakraborty, 2020).

In addition to detecting ripening agents, NIRS has proven effective in authenticating organic products, which are gaining popularity due to their perceived health benefits and environmental sustainability (Andika & Bidayati, 2024). Organic farming practices avoid synthetic pesticides and fertilizers, contributing to higher food safety standards and promoting sustainability (Andika & Bidayati, 2024). Differentiating between organic and inorganic products, however, is often labor-intensive, expensive, and time-consuming when using traditional methods. Portable NIR spectrometers offer a cost-effective and non-invasive alternative for verifying organic authenticity, ensuring consumer confidence in product labeling (Leitner & Vogl, 2020).

NIRS has been successfully employed in various studies to differentiate organic and inorganic products across a range of fruits, like apples and pineapples (Amuah et al., 2019; Song, Wang, Maguire, & Nibouche, 2016). By analyzing the spectral data generated from the internal vibrations of chemical bonds, NIRS can reliably classify fruits based on their organic status, enhancing the integrity of organic certification processes (Song et al., 2016).

Beyond fresh fruits, NIRS is also applicable to processed fruit products like juices (Šnurkovič, 2013). Pineapple juice, a major export product from countries like Ghana, is highly susceptible to adulteration due to the high value associated with organic labeling. Fraudulent practices such as falsely labeling inorganic juice as organic (Amuah et al., 2019) have become a significant issue, necessitating reliable authentication methods. NIRS offers a solution by providing a quick and non-destructive means of differentiating organic and inorganic juices, ensuring consumer protection and product integrity.

This study addresses the limited research on near-infrared spectroscopy (NIRS) for fruit quality assessment in Ghana. While several studies have demonstrated the effectiveness of handheld NIR spectroscopy in assessing fruit quality in other regions, most of these studies have focused on temperate fruits such as apples, grapes, and citrus (Grabska et al., 2023; Santos et al., 2021; Vallone et al., 2019). This study comprehensively evaluates the use of handheld NIRS to distinguish organic and inorganic mangoes and their processed products. By demonstrating its feasibility for rapid, non-destructive classification, this research provides valuable insights for quality control and postharvest loss reduction in Ghana's mango industry, filling a critical knowledge gap in Sub-Saharan Africa.

1.2 Problem Statement

Despite the increasing demand for mangoes and products, traditional methods to assess fruit quality remain labor-intensive, costly, and destructive, requiring extensive sample preparation. These methods, which involve measuring internal quality parameters such as total soluble solids (TSS), titratable acidity (TA), and pH, often result in errors due to their reliance on representative sampling, which may not accurately reflect the quality of the entire crop (Ncama, Magwaza, Mditshwa, & Tesfay, 2018). Furthermore, these techniques necessitate trained personnel and specialized equipment, making them impractical for large-scale or real-time applications (Cortés et al., 2016). As a result, there is a growing need for faster, non-destructive techniques capable of reliably assessing fruit quality without damaging the samples (Arendse, Fawole, Magwaza, & Opara, 2018; Shah et al., 2021).

One of the major challenges in mango production is ensuring optimal harvest time to achieve the best quality and shelf life. Immature fruits harvested too early may not ripen properly, while overripe fruits are more prone to spoilage and have a shorter shelf life. Chemical indicators like TSS and pH are essential for determining fruit ripeness and postharvest quality. However, current methods of assessing these indicators, such as chemical assays, are time-consuming and destructive, limiting their applicability for continuous or large-scale monitoring (He et al., 2022).

Additionally, the increasing use of artificial ripening agents, such as ethephon and calcium carbide, introduces safety concerns. Although ethephon is considered safer than calcium carbide, its improper use can result in harmful concentrations that pose health risks to consumers. Current detection methods for artificial ripening agents are complex, destructive, and require specialized laboratory setups, making them unsuitable for large-scale field applications (Mahanti & Chakraborty, 2020). This situation is further exacerbated by the lack of stringent regulatory enforcement in many developing countries, leading to widespread misuse of ripening agents (Islam et al., 2016).

Another significant challenge lies in the authentication of organic products. With the rising demand for organic mangoes and other fruit products, ensuring the integrity of organic labels has become a critical issue. Traditional certification processes are often time-consuming and expensive, relying heavily on paperwork rather than product analysis (Leitner & Vogl, 2020; Zorn, Lippert, & Dabbert, 2009). Additionally, detecting unwanted contaminants in inorganic products requires sophisticated laboratory equipment, which is neither readily available nor feasible for use in the food distribution chain (López, Arazuri, García, Mangado, & Jarén, 2013).

In light of these challenges, portable near-infrared (NIR) spectroscopy presents a promising solution. NIR spectroscopy provides a fast, non-destructive, portable, and cost-effective approach for evaluating fruit quality parameters (Amuah et al., 2019), detecting ripening agents (Lakade, V, Ramasamy, & Shetty, 2019), and authenticating organic products (Anyidoho, Teye, & Agbemafle, 2021). However, despite the proven utility of NIR in other areas of food quality assessment, there is limited research on its application in distinguishing between organic and inorganic mango products and detecting chemical ripening agents in mangoes. Moreover, using portable NIR devices combined with chemometric techniques to classify mangoes based on internal quality parameters such as TSS and pH remains underexplored.

This study seeks to address these gaps by investigating the potential of portable NIR spectroscopy, combined with advanced chemometric models, to provide a reliable, non-destructive method for assessing mango quality, detecting artificial ripening agents, and differentiating between organic and inorganic mango products. By developing robust models for on-site applications, this research aims to contribute to the broader goal of improving food quality, safety, and labeling integrity in the global mango industry.

1.3 Justification

The application of near-infrared (NIR) spectroscopy in food quality assessment has gained significant traction due to its numerous advantages, such as being a rapid, non-destructive, non-invasive, and environmentally friendly technique (Gullifa et al., 2023; Yin et al., 2019). NIR spectroscopy offers chemical-free analysis, making it ideal for ensuring product integrity while minimizing waste and environmental impact. It has been widely used to assess various food products, including grains (Egesel & Kahrıman, 2012), tea (Sun et al., 2020), honey (Cozzolino & Corbella, 2003), and coffee (Santos, Sarraguça, Rangel, & Lopes, 2012), with successful applications in both qualitative and quantitative evaluations.

In recent years, NIR spectroscopy has been applied to distinguish between different fruit cultivars and assess key quality parameters such as total soluble solids (TSS) and acidity in fruits like mangoes and pineapples (Amuah et al., 2019; Schmilovitch, Mizrach, Hoffman, Egozi, & Fuchs, 2000). Its ability to predict internal quality parameters without damaging the sample makes it valuable for non-invasive food quality control (Anyidoho, Teye, Agbemafle, Amuah, & Boadu, 2021). This technology's fast and precise results significantly reduce analysis time and costs, making it suitable for real-time and large-scale quality assessments.

Despite the extensive use of NIR spectroscopy in the food industry, limited research has been conducted on its application to differentiate between organic and inorganic mango products or to detect chemical ripening agents such as ethephon. This study aims to fill that gap by employing portable NIR devices in combination with chemometrics to non-invasively classify organic and inorganic mangoes, chips, and juices and to predict the concentration of artificial ripening agents. By developing robust, non-destructive models, this research will improve food safety, product authentication, and the integrity of food labeling systems. The results could offer a practical and environmentally sustainable alternative to destructive and labor-intensive fruit quality assessment methods.

The ability of NIR spectroscopy to provide rapid, accurate, and non-destructive analysis with minimal sample preparation justifies its use in this study, particularly in addressing challenges related to food safety and fraud detection in the fruit industry. This work can enhance consumer confidence in organic products and help regulatory bodies ensure compliance with food safety standards.

1.4 Objectives of the Study

1.4.1 General Objective

The general objective of this study is to develop rapid identification, quantification, and quality assessment techniques for fresh and processed fruit products using near-infrared spectroscopy (NIRS) and wet chemistry.

1.4.2 Specific Objectives

In order to achieve the general objective, the following were the specific objectives:

1. To non-destructively classify mango varieties and detect TSS and pH in intact mangoes using NIRS.
2. To measure ethephon qualitatively and quantitatively in intact mango fruits.
3. To classify organic and inorganic mangoes to detect fraud in;
   i. Intact fruits.
   ii. Dried fruits (chips).
   iii. Juice.

4. To authenticate organic pineapple juice from inorganic ones.
5. To analyse the physicochemical and microbial properties of expired and unexpired commercial fruit juices.

1.5 Research Questions

1. What quality parameters of mango fruits can be quantified using a handheld NIR spectrometer?
2. Can NIRS accurately predict ethephon levels in mango fruits?
3. Is fingerprinting organic and inorganic mango products feasible using portable NIRS technology?
4. How effectively can dual NIR spectrometers distinguish between organic and inorganic pineapple juice?
5. What are the physicochemical and microbial properties of expired and unexpired fruit juices?

1.6 Linkages

The five research topics in this study are interconnected through their reliance on NIRS for non-destructive analysis of fruit quality. Each topic builds on the previous, expanding the application of NIRS from fresh fruit assessment to processed products and food safety. The classification of organic and inorganic products, detection of chemical residues, and analysis of expired fruit products contribute to a comprehensive understanding of how NIRS can enhance quality control across various fruit production and processing stages.

1.7 Organization of the Thesis

This thesis consists of eight (8) chapters. Chapter one (1) highlights the background of the study, the main and specific objectives, and the justification of the study. Chapter two (2) overviews relevant literature and theoretical foundations on the research subject. Chapters 3, 4, 5, 6, and 7 are dedicated to the research articles specific to the abovementioned objectives. The summary, key findings, conclusion, and recommendations are presented in Chapter 8.

CHAPTER TWO

LITERATURE REVIEW

A review of recent trends, advances, and future directions of near-infrared spectroscopy in fruit analysis.

*Francis Padi Lamptey[1,2,5], *Ernest Teye[1,5], Flavio Odoi-Yorke[3], Gifty Serwaa Otoo[2], Ernest Ekow Abano[1], Charles L. Y. Amuah[4,5], Sandra Ama Kaburi[2]

[1] University of Cape Coast, College of Agriculture and Natural Sciences, School of Agriculture, Department of Agricultural Engineering, Cape Coast, Ghana

[2] Cape Coast Technical University, School of Applied Sciences and Technology, Department of Food Science and Postharvest Technology, Cape Coast, Ghana

[3] Cape Coast Technical University, School of Engineering, Department of Renewable Energy Technology, Cape Coast, Ghana

[4] University of Cape Coast, College of Agriculture and Natural Sciences, School of Physical Sciences, Department of Physics, Laser and Fibre Optics Centre, Cape Coast, Ghana

[5] Africa Centre of Excellence for Food Fraud and Safety, AfriFoodinTegrity Centre, University of Cape Coast, Cape Coast, Ghana

Tel: 233-243170302/+233-243514504

*Email:ernest.teye@ucc.edu.gh / francis.lamptey@cctu.edu.gh

Under Review, Infrared Physics and Technology (Scopus / Sci, IF =3.1)

Statement of contribution of Joint Authorship

Francis Padi Lamptey: (Candidate)

Conceptualised the topic, established methodology, data collection and analysis, preparation of tables and figures, writing and compilation of the original manuscript.

Teye, E.: (Principal Supervisor)

Conceptualised the topic, established methodology, supervised and edited the manuscript and co-author of manuscript.

Ernest Ekow Abano: (Co-Supervisor)

Conceptualised the topic, supervised and edited the manuscript and co-author of manuscript.

Charles L. Y Amuah: (Research Team member)

Validation, visualization and data curation and co-author of manuscript

Flavio Odoi-Yorke (Research Team Member)

Validation, visualization and data curation and co-author of manuscript

Gifty Serwaa Otoo (Research Team Member)

Validation, visualization and data curation and co-author of manuscript

Sandra Ama Kaburi (Research Team Member)

Validation, visualization and data curation and co-author of manuscript.

Abstract

The growing global demand for fresh fruits necessitates efficient, non-destructive methods for assessing fruit quality, especially for export. Traditional fruit quality assessment techniques are often labor-intensive, time-consuming, and destructive, making them unsuitable for large-scale or real-time analysis. This study provides a comprehensive bibliometric analysis of near-infrared spectroscopy (NIRS) in fruit analysis to address these limitations. The study adhered to the PRISMA guidelines to extract peer-reviewed papers from 2003 – 2023 from the Scopus database. Thereafter, the bibliometric analysis was conducted using R software's Bibliometrix package to evaluate global trends, key contributors, and emerging themes in the field. The results show that NIRS has become an essential tool for non-destructive quality assessment in fruits, accurately predicting attributes such as total acidity, soluble solids content, and internal disorders. Integrating machine learning and artificial intelligence models, particularly artificial neural networks and deep learning has further enhanced the predictive capabilities of NIRS. Additionally, technological innovations such as portable spectrometers and hyperspectral imaging have expanded the applicability of NIRS beyond laboratory settings to in-field assessments. The findings highlight the ongoing evolution of NIRS technology, its significant impact on fruit quality evaluation, and the potential for future advancements in this field. Future research should focus on improving the adaptability of NIRS to diverse fruit types and production environments and exploring the use of artificial intelligence and machine learning to further enhance data interpretation and predictive accuracy. Such innovations could significantly broaden the scope of NIRS applications, making it a critical tool for sustainable agriculture and global food security.

Keywords: Bibliometric analysis; Fruit; Near-infrared spectroscopy; Non destructive testing; Quality

2.1 Introduction

The consumer demand for fresh fruits has significantly increased in recent years due to the growing awareness of their consumption's health and nutritional benefits (Ramos, Miller, Brandão, Teixeira, & Silva, 2013). Global consumption levels are influenced by various factors, including agronomical conditions like climate and seasonality, economic aspects shaped by local policies such as affordability and infrastructure, and socio-cultural factors like food habits, education, and availability of alternatives (Liu et al., 2022). The quality of fruit is a critical determinant of its market value. High-quality fruit is associated with enhanced antioxidant properties and a rich concentration of bioactive compounds such as flavonoids, anthocyanins, polyphenols, and ascorbic acid. Therefore, giving enough attention to both the maintenance and assessment of fruit quality is important for ensuring high quality fruits (Wang, Sun, Yang, Pu, & Zhu, 2016). Traditional methods of fruit quality assessment (e.g., visual inspection, sensory evaluation, instrumental, and physicochemical analysis) are well established. However, these methods are often sample-destructive, labor-intensive, and time-consuming, which reduces their effectiveness for online or in-line quality monitoring (Wang et al., 2016).

Near-infrared spectroscopy (NIRs), which operates within the 700 to 2500 nm range of the electromagnetic spectrum (Salam, Saad, Manap, Salehuddin, & Karim, 2018), primarily derives its spectral information from the internal vibration and absorption of O–H, C–H, N–H, and other hydrogen-containing groups in overtone and combination bands (Teye & Amuah, 2022). This allows for the detection of specific quality attributes such as sugar content (Subedi & Walsh, 2011), acidity (Purwanto, Sari, & Budiastra, 2015), moisture (Malvandi, Feng, & Kamruzzaman, 2022), and even internal defects in fruits (Raghavendra, Guru, & Rao, 2021). This non-invasive technology has revolutionized monitoring of fruit quality, providing a rapid, accurate, and efficient alternative to traditional methods. NIRS has found widespread applications across different stages of fruit production—from monitoring ripeness during growth (González-Caballero, Sánchez, Fernández-Novales, López, & Pérez-Marín, 2012), estimating maturity (Shah et al., 2020), to quality control during post-harvest handling and retail (Theanjumpol, Self, Rittiron, Pankasemsuk, & Sardsud, 2014).

Over the past few decades, NIRS technology has evolved remarkably. Initially, benchtop NIR spectrometers were confined to laboratory settings. Today, portable and handheld NIRS devices have emerged, providing significant advantages, including small size, low cost, robustness, a user-friendly interface, and portability. These advancements facilitate in-field assessments and enable real-time decision-making in the agricultural industry (Escribano, Biasi, Lerud, Slaughter, & Mitcham, 2017). As a result, farmers, producers, and retailers can now assess the quality of their products on-site, reducing post-harvest losses. Adopting portable NIRS has been particularly valuable in industries dealing with highly perishable fruits like pineapples, bananas, and mangoes (Yahia, Ornelas-Paz, & Elansari, 2011), where quality attributes must be monitored continuously to meet market standards.

A key driver behind the success of NIRS in fruit quality assessment is the integration of chemometrics—a fusion of mathematics, statistics, and chemistry, which has led to significant advances in analytical data assessment (Molognoni et al., 2020). By applying chemometric techniques such as principal component analysis (PCA), linear discriminant analysis (LDA), k-nearest neighbors (kNN), partial least squares regression (PLSR), partial least square discriminant analysis (PLS-DA), support vector machine (SVM), researchers have been able to extract meaningful information from NIR spectra, building predictive models that can accurately assess fruit quality based on spectral signatures (Hidalgo, Fechner, Marchevsky, & Pellerano, 2016; Mishra, Woltering, & Harchioui, 2020). The combination of NIRS and chemometrics has not only enhanced the accuracy of quality assessments but has also opened the door to more sophisticated applications such as the detection of adulteration (Alamar, Caramês, Poppi, & Pallone, 2020), classification of organic versus conventionally grown fruits (Amuah et al., 2019), and the prediction of storage life (Beghi, Giovanelli, Malegori, Giovenzana, & Guidetti, 2014).

Despite the extensive research and growing body of literature on the applications of NIRS in fruit quality assessment, there has been a lack of comprehensive studies that analyse the evolution of the technology, current trends, and future directions. As the field expands, understanding how NIRS has been applied, which fruits have been the research focus, and what gaps exist is crucial for driving future innovation. Bibliometric analysis is a quantitative method used to determine the volume and growth patterns of literature in a particular emerging field (Wang & Si, 2023).

Bibliometric analysis provides a robust framework for achieving this, allowing researchers to quantitatively assess trends, collaborations, and the growth of scientific knowledge over time (Odoi-Yorke, 2024). Bibliometric tools have been widely used in other domains of food science to explore research productivity, reveal collaboration networks, and identify emerging research themes. However, such an analysis focused on NIRS applications in fruit quality is still lacking. The review by Aleixandre-Tudó, Castelló-Cogollos, Aleixandre, and Aleixandre-Benavent (2020) emphasizes spectroscopy's role as a cost-effective, efficient, and non-destructive method for evaluating food quality. Its application spans various food products, from fruits and vegetables to meat and olive oil. NIRS appears as a dominant technique, mainly due to its accuracy in food quality assessment without causing damage. The growing interest in Raman and fluorescence spectroscopy further highlights the field's adaptability, where different techniques cater to specific food safety needs, such as adulteration detection and authentication. Liu et al. (2023) highlight AI's significant influence on food safety, particularly in improving yield, quality, and traceability.

The growing collaboration between the U.S. and China and the increasing contributions from institutions like the Chinese Academy of Sciences reflect AI's critical role in modern food safety protocols. Key AI applications span precision agriculture and nutrition, demonstrating its importance across the food chain. The potential of AI to reduce food waste and enhance sustainability also aligns with broader global priorities for food security and resource efficiency. The study on chemometrics by Aleixandre-Tudó, Castelló-Cogollos, Aleixandre, and Aleixandre-Benavent (2022) highlights its importance in data-driven decision-making for food quality evaluation. Principal component analysis (PCA), partial least squares (PLS), and discriminant analysis (DA) are prominent techniques that drive advancements in this area. Integrating chemometrics with spectroscopy allows for more accurate assessments of food products, making these tools indispensable for food science research. Bannor, Arthur, Oppong, and Oppong-Kyeremeh (2023) focus on food fraud, identifying organic foods, seafood, and olive oil as high-risk products. The study suggests that technological gaps, particularly in real-time authentication, contribute to the increasing global prevalence of food fraud. Addressing these gaps through international collaboration and enhanced regulatory frameworks could significantly reduce instances of fraud and increase consumer trust in food markets. The research by (Ma, Luo, Zhang, & Gao, 2022; Ni et al., 2023) on non-destructive testing technologies for fruit quality assessment highlights the rapid advancements in electronic nose technology, machine vision, and spectral detection techniques. These technologies offer non-invasive monitoring of key fruit quality parameters, such as sugar content and hardness, promoting more efficient and sustainable fruit production practices. The use of NIR spectroscopy, specifically in ripeness detection, aligns with the broader trend of adopting non-destructive techniques in the food industry, further enhancing production efficiency and reducing food waste.

This paper addresses this gap by conducting a comprehensive bibliometric analysis of NIRS applications in fruit quality assessment. Through this approach, the study seeks to examine the evolution of NIRS technology, identify emerging trends, and provide insights into potential future research directions. The bibliometric analysis explores global research productivity, highlights key contributors, and assesses the intellectual and social structures underpinning the field. The findings will contribute to a deeper understanding of the technological progress, the remaining challenges, and the opportunities for future innovation. Finally, this study aims to provide a roadmap for researchers and industry stakeholders to harness the full potential of NIRS in ensuring the quality, safety, and sustainability of fruits in the global market.

2.2. Methodology

This study employed bibliometric analysis to investigate recent trends, advances, and future directions of NIRS in fruit analysis. The methodology was designed to ensure a thorough and unbiased examination of the existing literature, following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher, Liberati, Tetzlaff, Altman, & Group, 2010; Odoi-Yorke, 2024). PRISMA is a structured approach for conducting and reporting systematic reviews, enhancing transparency, reproducibility, and quality of research synthesis in various fields (Abelha, Fernandes, Mesquita, Seabra, & Ferreira-Oliveira, 2020; Odoi-Yorke, 2024; Sohrabi et al., 2021).

We conducted data extraction on August 13, 2024, using the Scopus database, which was selected for its comprehensive coverage of peer-reviewed literature, particularly in science, technology, engineering, and mathematics (Baas, Schotten, Plume, Côté, & Karimi, 2020; Le et al., 2021). Scopus was chosen over Web of Science due to its broader journal coverage and more extensive citation tracking capabilities, especially for recent publications (Falagas, Pitsouni, Malietzis, & Pappas, 2008; Joshi, 2016; Le et al., 2021; Mongeon & Paul-Hus, 2016). Although offering a wide range of sources, Google Scholar was not selected due to its inclusion of non-peer-reviewed content and the potential for duplicate entries, which could compromise the reliability of our analysis (Adriaanse & Rensleigh, 2013).

The search strategy employed a combination of terms related to NIRS and fruit, with the exact Boolean operators and search strings as follows: ("near-infrared spectroscopy" OR "NIR spectroscopy" OR "NIRS") AND ("fruit" OR "fruits"). This search strategy was designed to capture a wide range of relevant literature while minimizing the inclusion of irrelevant studies. The search period from 2003 to 2023 comprises two decades of research to provide a comprehensive overview of the field's evolution. As shown in Figure 2.1, this search initially yielded 1,632 documents. Inclusion criteria focused on peer-reviewed research outputs, specifically articles, conference papers, review papers, and book chapters. Exclusion criteria were implemented to remove less substantive document types such as notes, errata, and editorials, reducing the dataset to 1,621 papers. A further refinement to include only English-language publications resulted in a final dataset of 1,348 documents, which formed the basis for our bibliometric analysis.

The bibliometric analysis was conducted using the Bibliometrix package in R software, facilitated by the biblioshiny library. Bibliometrix is an R-based tool for comprehensive science mapping analysis (Aria & Cuccurullo, 2017; Derviş, 2019; Wang et al., 2024). It enables researchers to perform quantitative analysis of scientific literature, including citation analysis, co-citation networks, and bibliographic coupling (Agyekum, Odoi-Yorke, Abbey, & Ayetor, 2024; Odoi-Yorke, 2024). Bibliometrix facilitates the exploration of research trends, collaboration patterns, and the impact of publications across various scientific fields, aiding in literature reviews and research planning (Ma et al., 2022; Ni et al., 2023; Odoi-Yorke, 2024). The bibliometric analysis performed in this study is as follows: (i) publication trends over the 20 years were examined to identify patterns in research output and growth of the field; (ii) country production and collaboration networks were mapped to understand the global distribution of research efforts and international partnerships; (iii) keyword analysis and trend topic identification were performed to discern the most prominent themes and emerging areas of interest within NIRS fruit analysis; (iv) thematic mapping and evolution were conducted to visualize the development of research themes over time; (v) factorial analysis was employed to uncover latent structures in the research landscape; (vi) co-occurrence network analysis of keywords was utilized to reveal interconnections between different research topics and methodologies and (v) examining journals, institutions and total citations received per country. Furthermore, a systematic review was conducted on the top 50 most relevant papers. This analysis provides important insights into the current research areas, highlighting key findings and recommendations and suggesting possible future research directions.

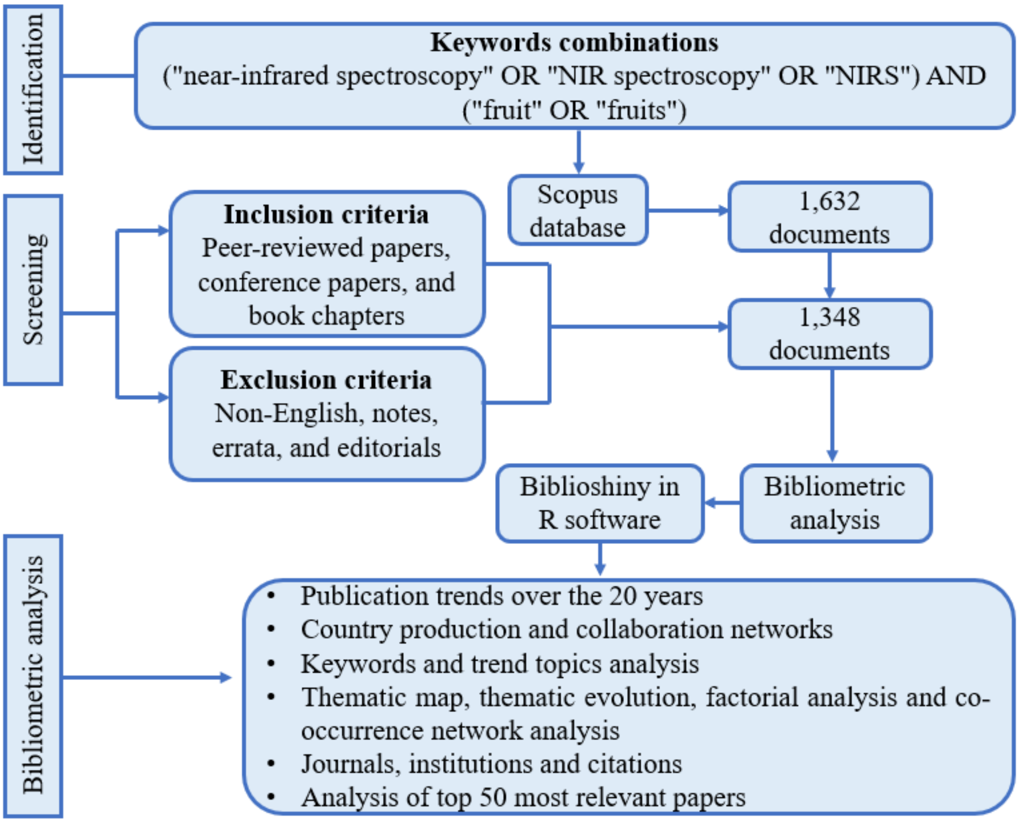


Figure 2.1: Study Framework Employed for Bibliometric Analysis

2.3. Results and Discussion

2.3.1Summary of Bibliometric Data and Annual Article Production

The bibliometric data in Figure 2.2 shows a significant body of research comprised of two decades from 2003 to 2023. During this period, the field experienced substantial growth, with an average annual growth rate of 11.85%. This indicates a rapidly expanding area of study with increasing interest and investment. The analysis covers 1348 documents from 403 sources, suggesting a diverse range of publications and research outlets contributing to this field. The many authors (3366) involved demonstrate widespread engagement from the scientific community, with an average of 4.63 co-authors per document. This level of collaboration is further emphasized by the 19.44% rate of international co-authorship, indicating global interest and cross-border cooperation in advancing this area of research.

Remarkably, only 36 authors produced single-authored documents, highlighting the predominantly collaborative nature of work in this domain. The research appears well-established and influential, evidenced by the average of 28.42 citations per document. This suggests that the published works have significantly impacted subsequent research. The document's average age of 7.78 indicates a balance between newer and more established research, allowing for fresh perspectives and long-term studies. The extensive use of references (43,205) points to a thorough grounding in existing literature and a commitment to building upon prior knowledge.

Figure 2.2: Summary of Bibliometric Data

Figure 2.3 illustrates the annual article production from 2003 to 2023, showing a clear upward trend in research output over the two-decade period. Starting with a modest 13 articles in 2003, the field has grown substantially, culminating in 122 articles published in 2023. This growth trajectory implies increasing interest and investment in applying near-infrared spectroscopy to fruit analysis. The early years (2003-2010) show a steady but gradual increase, with article production roughly tripling from 13 to 41 annually. This initial phase likely represents the establishment of foundational research and methodologies. A notable acceleration occurred from 2011 onwards, with production nearly doubling from 50 articles in 2011 to 111 in 2019. This rapid expansion suggests a maturation of the field, possibly driven by technological advancements in spectroscopy equipment, increased recognition of the technique's potential in fruit quality assessment, and growing demand for non-destructive analytical methods in the food industry. The years 2019-2023 mark a period of sustained high productivity, with over 100 articles published annually. The peak of 135 articles in 2022 indicates the field's current vitality and relevance. The slight decrease to 122 articles in 2023 may indicate natural fluctuations or signal a stabilization of research output. This overall trend has significant implications for the fruit industry and food science. Consistent publication growth suggests improvements in near-infrared spectroscopy techniques, potentially leading to more accurate, efficient, and cost-effective methods for assessing fruit quality, ripeness, and composition. The volume of recent research also implies a diverse range of applications being explored, from sorting and grading fruits to monitoring nutritional content and detecting defects or diseases.

Furthermore, this sustained research interest likely correlates with the increased adoption of near-infrared spectroscopy in commercial and industrial settings, potentially revolutionizing quality control processes in the fruit supply chain. The results also hint at this research's interdisciplinary nature, potentially involving collaborations between spectroscopists, food scientists, agronomists, and data analysts, especially as the field may incorporate advanced data processing techniques and machine learning to interpret spectral data.

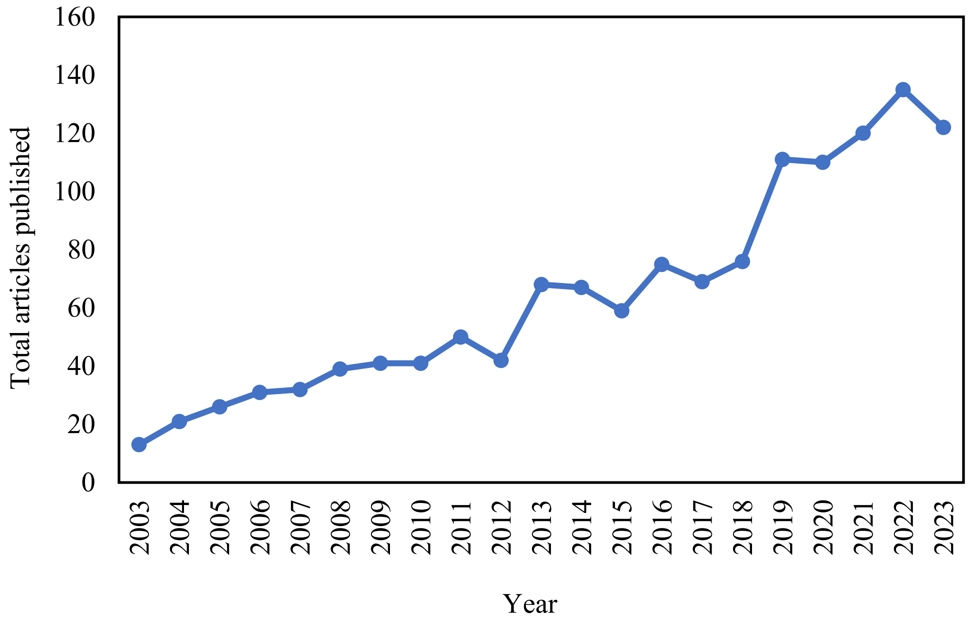


Figure 2.3: Annual Scientific Production

2.4. Countries' Distribution and Collaborations

Figure 2.4 presents a comprehensive overview of country-specific production in near-infrared spectroscopy and fruit studies, revealing significant disparities in research output across different regions. China has the highest publications of 2,189 articles, far surpassing other countries and accounting for a substantial portion of the global research in this field. This dominance suggests China's strong emphasis on agricultural technology and food science research, potentially driven by its large population and the need for efficient food production and quality control (Sandrey & Edinger, 2009; Xu, Li, Qi, Tang, & Mukwereza, 2016; Zhao & Huang, 2011). Following China, European countries collectively show a strong presence, with Italy and Spain leading the continent. This indicates Europe's continued investment in agricultural research and technology, likely influenced by its diverse climate zones and rich agricultural traditions. The United States ranks fourth with 298 articles, highlighting its significant but comparatively smaller contribution to this research area.

However, several developing and emerging economies feature prominently in the list, including Brazil, Thailand, and Indonesia, highlighting the global nature of this research and its importance to countries with significant agricultural sectors. The results also reveal a notable presence of research from Oceania, with Australia and New Zealand among the top 15 countries. This suggests a strong focus on fruit quality and agricultural technology in these export-oriented economies. The distribution of research across continents indicates a global recognition of the importance of near-infrared spectroscopy in fruit studies, potentially for improving crop yields, quality assessment, and postharvest management. However, the stark differences in research output between countries also point to potential disparities in research funding, technological capacity, and prioritization of this specific field. This imbalance could affect global knowledge sharing and technological advancement, potentially widening the gap between leading and lagging countries in agricultural innovation and food security measures.

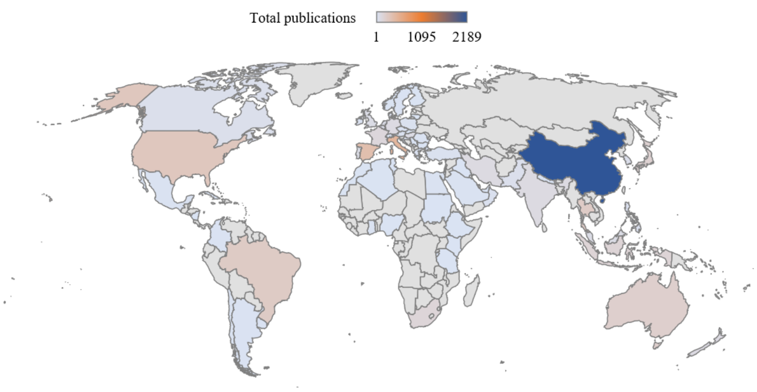


Figure 2.4: Country Scientific Production

Figure 2.5 shows a research collaboration map between countries and regions. This map visualization depicts international collaborations in near-infrared spectroscopy and fruit studies, offering valuable insights into global research partnerships. It can be seen that the map uses colour coding to highlight the intensity of collaboration, with China standing out in dark blue, indicating its central role as a major collaborator. The network of lines connecting various countries represents collaborative links, with thicker lines suggesting stronger or more frequent collaborations. The visualization reveals a complex web of international cooperation, with connections across continents. China's prominent position highlights its significant contribution to and engagement in global research efforts. This aligns with China's growing emphasis on agricultural technology and food science research, likely driven by its large population and the need for advanced food production and quality control methods (Sandrey & Edinger, 2009; Xu et al., 2016; Zhao & Huang, 2011).

The map shows strong collaborative ties between China and several Western countries, particularly the United States, as evidenced by the thick line connecting these nations. This suggests a robust exchange of knowledge and resources between these two scientific countries, potentially leading to accelerated advancements in the field. European countries, represented in lighter blue, also actively collaborate within Europe and with global partners. This reflects Europe's continued investment in agricultural research and its openness to international scientific cooperation. It can be observed that the visualization highlights connections with countries in the Southern Hemisphere, such as Australia, New Zealand, and parts of South America.

This global spread of collaborations indicates the universal relevance of near-infrared spectroscopy in fruit studies, spanning diverse climates and agricultural systems worldwide. It also suggests a healthy flow of knowledge and expertise across hemispheres, potentially leading to more comprehensive and globally applicable research outcomes. The varying thicknesses of the connecting lines imply different levels of collaborative intensity between countries. Although some connections appear strong and well-established, others are thinner, possibly indicating emerging or less frequent collaborations. This pattern may signify differences in research priorities, funding availability, or historical scientific ties between nations.

The implications of these collaborative patterns are significant. First, it suggests that research in near-infrared spectroscopy and fruit studies benefits from a global perspective, incorporating diverse agricultural contexts and challenges. This international approach will likely lead to more robust and widely applicable findings. Second, the strong collaborative networks centered around countries like China and the United States may accelerate innovation in this field as resources and knowledge are shared more efficiently. However, it also raises questions about equitable access to research findings and technologies for countries less prominently featured in these collaborations.

Furthermore, this global collaboration network can potentially address shared challenges in fruit production, quality assessment, and postharvest management worldwide. It may facilitate the transfer of technology and methodologies from more advanced research centres to regions where such capabilities are still developing. Finally, this collaborative approach in near-infrared spectroscopy and fruit studies could contribute significantly to global food security, sustainable agriculture practices, and the overall advancement of agricultural science.

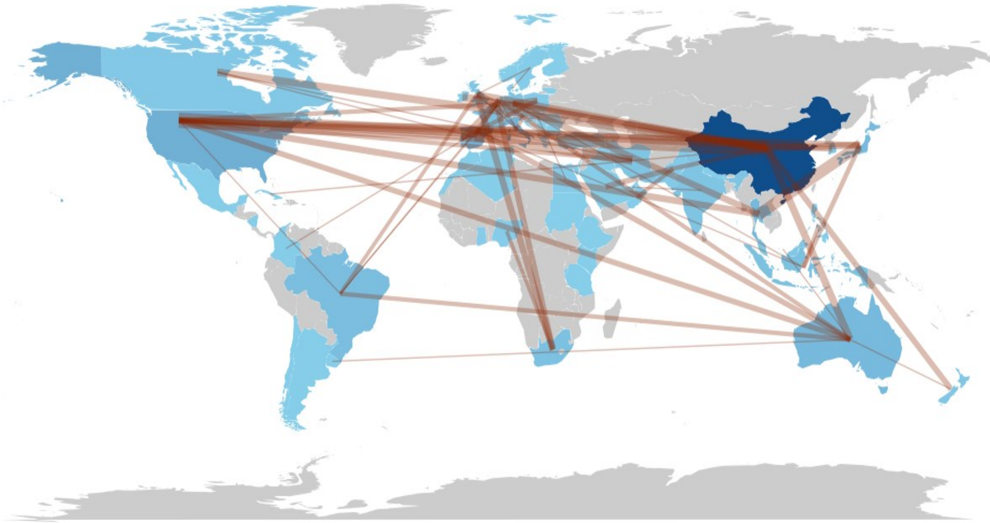


Figure 2.5: Country Collaboration Map

Figure 2.6 displays a detailed breakdown of single country publications (SCP) and multiple country publications (MCP) for corresponding authors in the field of various countries. It can be observed that China stands out as the dominant contributor, with 340 SCPs and 53 MCPs, far surpassing other nations. This highlights China's significant investment in and focus on agricultural technology research, likely driven by its large population and the critical need for advanced food production and quality control methods. The substantial gap between China's output and that of other countries suggests a potential concentration of expertise and resources in this field within China.

European countries, particularly Spain, Italy, and France, show a strong presence in SCPs and MCPs, suggesting Europe's continued commitment to agricultural research and its openness to international collaboration. The United States, while ranking fourth in SCPs with 45 publications, demonstrates a relatively high proportion of MCPs, indicating a strong inclination towards international collaboration. This trend is even more pronounced in countries like the Netherlands, which has more MCPs than SCPs, suggesting a highly collaborative research environment.

Emerging economies such as Brazil, Thailand, and Indonesia also feature prominently, highlighting the global relevance of this research area and its importance to countries with significant agricultural sectors. The presence of these nations in both SCP and MCP categories indicates their growing domestic research capabilities and integration into global research networks.

However, some countries show a higher propensity for international collaboration than their total output. For instance, South Africa and Thailand have a notable number of MCPs compared to their SCPs, suggesting strong international research ties. This pattern could indicate strategic efforts to leverage global expertise and resources to advance their research capabilities.

These results have several implications. Firstly, the dominance of China in both SCPs and MCPs suggests that it may be setting the pace for global research in near-infrared spectroscopy and fruit studies. This could lead to China becoming a key source of innovation and methodologies in this field. Secondly, the varied ratios of SCPs to MCPs across countries highlight different national strategies in research. Countries with higher MCPs may benefit from knowledge transfer and shared resources, potentially accelerating their research progress.

However, the disparity in publication numbers also raises questions about global equity in research capabilities and access to advanced technologies in this field. Countries with fewer publications may face challenges in developing and applying cutting-edge techniques in their agricultural sectors. This could have implications for global food security and the ability of different nations to optimize their fruit production and quality control processes.

Furthermore, the results suggest a healthy level of international collaboration in this specialized field, which is crucial for addressing global challenges in agriculture. Such collaborations can lead to more comprehensive studies that account for diverse agricultural conditions and practices worldwide. However, the varying levels of involvement in MCPs also indicate that there might be room for increased international cooperation, particularly for countries currently showing lower levels of collaborative output.

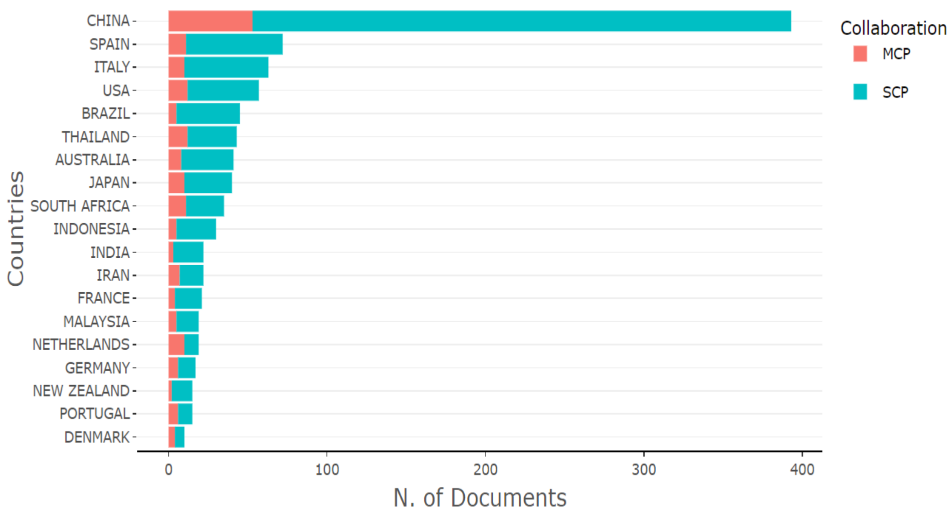


Figure 2.6: Corresponding Author's Countries

2.5. Analysis of Word Cloud and Trend Topics

The word cloud in Figure 2.7 comprehensively represents the most relevant keywords in near-infrared spectroscopy and fruit studies. This visualization provides key insights into the current focus areas, methodologies, and applications within this specialized agricultural and food science research domain. It can be observed that the most prominent terms in the word cloud are "nir spectroscopy," "chemometrics," and "fruit quality," indicating that these are central themes in the research field. Near-infrared (NIR) spectroscopy emerges as the primary analytical technique, likely due to its non-destructive nature and ability to rapidly assess various quality parameters of fruits (Amuah et al., 2019; Ncama, Magwaza, Mditshwa, & Tesfay, 2018). The prominence of "chemometrics" suggests that advanced statistical and mathematical methods are pivotal in analyzing the complex spectral data obtained from NIR measurements (Hidalgo et al., 2016; Mishra et al., 2020; Molognoni et al., 2020). "Fruit quality" is a key phrase emphasizing this technology's primary application in assessing and ensuring the quality of fruit products.

Other significant terms like "soluble solids content," "firmness," and "maturity" point to specific quality attributes that researchers are focusing on. These parameters are critical in determining fruit ripeness, taste, and overall quality, which are essential for consumers and the fruit industry. The presence of "non-destructive" in the cloud highlights one of the main advantages of NIR spectroscopy in fruit analysis – its ability to assess fruit quality without damaging the sample (Amuah et al., 2019; Shah et al., 2020).

Specific fruit types like "apple," "tomato," and "mango" are visible, indicating that these fruits are common subjects of study, possibly due to their economic importance or suitability for spectroscopic analysis. Including "classification" and "partial least squares regression" suggests that machine learning and statistical modeling techniques are frequently used to interpret spectral data and predict fruit quality parameters. Terms such as "hyperspectral imaging" and "visible/near-infrared spectroscopy" indicate that researchers are exploring various spectroscopic techniques beyond traditional

NIR methods. This implies a trend towards more sophisticated and comprehensive analytical approaches in fruit quality assessment. The presence of "principal component analysis" and "artificial neural network" further emphasizes the importance of advanced data analysis techniques in extracting meaningful information from spectral data. These methods are likely employed to handle spectroscopic datasets' high dimensionality and complexity.

The implications of these results are significant for the field of fruit quality assessment and the broader agricultural industry. The focus on non-destructive techniques suggests a move towards more efficient and cost-effective quality control methods that can be applied throughout the supply chain, from harvest to retail. The emphasis on chemometrics and advanced statistical techniques implies a growing need for interdisciplinary expertise, combining spectroscopy with data science and machine learning. Furthermore, the various quality parameters being studied (e.g., soluble solids, firmness, maturity) indicate a holistic approach to fruit quality assessment, which could lead to more comprehensive quality standards and improved consumer satisfaction. Including specific fruit types suggests that research might be tailored to address industry-specific challenges or consumer preferences for particular fruits.

Figure 2.7: Word Cloud (50 words)

Figure 2.8 shows the trend topics from 2003 to 2024. The figure shows a clear progression from more basic concepts and techniques towards increasingly sophisticated and specialized approaches. In the earlier years (2003-2010), the focus appears to be on fundamental aspects such as "sugar content," "modeling," "reflectance," and "non-destructive technique." These topics indicate that researchers were primarily concerned with establishing the basic methodologies and applications of near-infrared spectroscopy in fruit analysis. Moving into the 2010s, we can observe a shift towards more specific quality parameters and analytical techniques. Terms like "soluble solids content," "firmness," and "quality parameters" become prominent, suggesting a growing emphasis on comprehensive fruit quality assessment.

The mid-2010s marked a notable transition in research focus. There is an increased emphasis on advanced analytical methods and data processing techniques. Terms such as "chemometrics," "multivariate analysis," and "partial least squares" appear more frequently, indicating a growing sophistication in data analysis approaches. This shift likely reflects the need to handle the complex spectral data generated by near-infrared spectroscopy and extract meaningful insights from it.

Perhaps the most striking trend is the emergence of machine learning and artificial intelligence-related topics in recent years (2018-2022). Terms like "deep learning," "machine learning," and "artificial neural network" have become prominent, suggesting a significant move towards leveraging advanced computational techniques for spectral data analysis and interpretation. This trend aligns with the broader adoption of AI and machine learning across various scientific disciplines and indicates a new frontier in fruit quality assessment and prediction. The persistence of terms like "near-infrared spectroscopy," "NIR," and "fruit quality" throughout the entire period accentuates their fundamental importance to the field. However, the evolution of associated terms reflects the field's dynamic nature and responsiveness to technological advancements.

The implications of these trends are multiple. Firstly, they suggest that near-infrared spectroscopy in fruit studies is rapidly advancing, incorporating cutting-edge technologies and analytical methods. This evolution is likely driving more accurate, efficient, and comprehensive fruit quality assessments. The shift towards machine learning and AI techniques implies a future where fruit quality prediction could become increasingly automated and precise, potentially revolutionizing quality control processes in the fruit industry. Moreover, the trend toward more sophisticated data analysis methods suggests a growing need for interdisciplinary expertise.

Researchers and practitioners may need to combine spectroscopy, agriculture, and data science knowledge to stay at the forefront of these developments. This could lead to new educational and training requirements in the agricultural and food science sectors. The increasing complexity of analytical approaches also hints at the potential for a more nuanced understanding of fruit quality parameters. This could lead to developing more refined quality standards and potentially even tailored fruit production techniques to meet specific quality criteria. Lastly, the persistent focus on non-destructive techniques throughout the years underscores the importance of developing methods to assess fruit quality without damaging the produce. This continual emphasis suggests that future innovations in this field will likely continue to prioritize non-invasive assessment methods, which have significant practical and economic implications for the fruit industry.

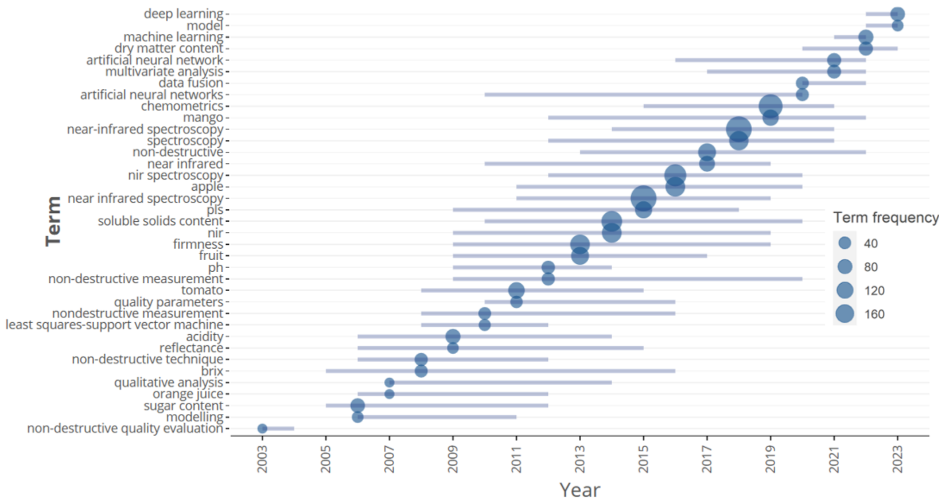


Figure 2.8: Trend Topics from 2003 - 2023

2.6.Co-Occurrence Network, Thematic Map, and Thematic Evolution

The co-occurrence network of keywords is shown in Figure 2.9. This provides a comprehensive overview of the interconnected themes and concepts within this research field. As seen, the centre of the network has a dominant term, "near-infrared spectroscopy," which serves as the core technique around which the research revolves. Key terms such as "chemometrics" and "NIR spectroscopy" are closely linked to this central node, indicating their fundamental importance in analyzing and interpreting spectral data.

The network shows several distinct clusters, each representing a different aspect or application of near-infrared spectroscopy in fruit studies. One prominent cluster, coloured in red, centres around "soluble solids content," highlighting the significance of this parameter in fruit quality assessment. This cluster's connection to terms like "partial least squares" suggests the common use of this statistical method in analyzing soluble solids data.

Another significant cluster, shown in blue, comprises various quality-related terms such as "fruit quality," "quality," and "apple." This grouping indicates a strong focus on applying near-infrared spectroscopy to assess overall fruit quality, with apples being a frequently studied fruit. The presence of terms like "non-destructive" in this cluster highlights the non-invasive nature of these spectroscopic techniques, a crucial advantage in fruit quality assessment.

The green cluster appears to be related to specific fruit characteristics and analytical methods, including terms like "firmness," "acidity," and "sugar content." This suggests a detailed focus on individual quality parameters contributing to overall fruit quality and consumer acceptance. The network also reveals connections to advanced data analysis techniques, as evidenced by terms like "artificial neural network," "machine learning," and "deep learning." These links indicate the growing importance of sophisticated computational methods in interpreting the complex spectral data obtained from near-infrared analysis. The presence of terms like "hyperspectral imaging" and "visible/near-infrared spectroscopy" suggests that researchers are exploring various spectroscopic techniques beyond traditional near-infrared methods, potentially to gain more comprehensive insights into fruit quality.

The implications of this co-occurrence network are significant for the field of fruit quality assessment and the broader agricultural industry. Firstly, it highlights the multidisciplinary nature of this research area, combining spectroscopy, chemometrics, machine learning, and agricultural science. Integrating diverse fields suggests that future advancements may require collaborative efforts across disciplines. The network also implies that researchers focus on developing comprehensive, non-destructive methods for simultaneously assessing multiple quality parameters. This could lead to more efficient and accurate quality control processes in the fruit industry, potentially reducing waste and improving consumer satisfaction. The prominence of specific fruits like apples in the network indicates that certain fruits may serve as model systems for developing and refining spectroscopic techniques.

However, the diversity of terms suggests that these methods are being applied to a wide range of fruits, pointing to the versatility and broad applicability of near-infrared spectroscopy in fruit studies. The inclusion of advanced data analysis techniques in the network suggests a trend toward a more sophisticated interpretation of spectral data. This could lead to more accurate predictions of fruit quality and potentially enable the detection of subtle quality variations that were previously difficult to measure.

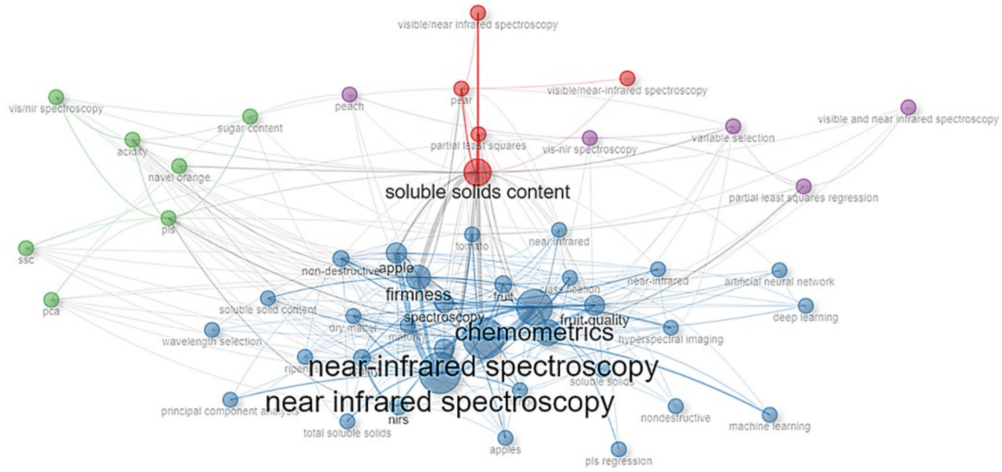


Figure 2.9: Keywords Co-occurrence Network

The thematic map of keywords, which explains the relationships between various research themes based on their development (density) and relevance (centrality), is displayed in Figure 2.10. It can be gleaned that the thematic map is divided into four quadrants. The basic themes quadrant comprises fundamental concepts and techniques that form the core of this research field. Terms like "chemometrics," "classification," "soluble solids content," and "partial least squares" are positioned here, indicating their widespread use and central importance. The presence of specific fruits like "apples," "pears," and "citrus" suggests that these are commonly studied subjects. The inclusion of "deep learning" in this quadrant is particularly interesting, as it implies that this advanced computational technique has become a standard tool in the field, potentially revolutionizing data analysis approaches.

The motor themes quadrant represents highly developed and central topics likely driving the field forward. "Principal component analysis," "ripeness," and "neural network" appear here, suggesting that these analytical methods are at the forefront of current research efforts. The presence of "internal quality" and "modeling" indicates a strong focus on developing comprehensive fruit quality assessment techniques. "Non-invasive" methods are also highlighted, underscoring the importance of non-destructive analysis in fruit studies.

The niche themes quadrant contains highly developed but less central topics, representing specialized research areas. "HPLC" (High-Performance Liquid Chromatography) appears here, suggesting its use as a complementary technique to near-infrared spectroscopy. "Dry matter content," "*prunus persica*" (peach), and "index of absorbance difference" are also in this quadrant, indicating focused research on specific fruit characteristics or species. Though not central to the field, these niche areas may represent critical specialized applications or emerging research directions.

The emerging or declining Themes quadrant is particularly intriguing, as it can indicate new trends and fading topics. "Near-infrared spectroscopy (NIRS)" and "near-infrared (NIR) spectroscopy" appear here, which might seem counterintuitive given the field's focus. However, this could suggest that these terms are evolving or being refined into more specific applications. "Feature extraction," "PLS-DA" (Partial Least Squares Discriminant Analysis), and "SIMCA" (Soft Independent Modeling of Class Analogy) are also in this quadrant, potentially indicating emerging analytical techniques gaining traction in the field.

The implications of this thematic map are significant for researchers, practitioners, and stakeholders in the fruit industry. It suggests that while traditional spectroscopic and chemometric methods remain fundamental, there is a clear trend toward integrating advanced computational techniques like deep learning and neural networks. This evolution could lead to more accurate and sophisticated fruit quality assessment tools. The prominence of non-invasive techniques across different quadrants underscores the importance of developing methods to assess fruit quality without damaging the produce, a crucial factor for practical applications in the industry. The map also highlights the multidisciplinary nature of this field, combining spectroscopy, data analysis, and fruit science. This suggests that future advancements may require collaborative efforts across these disciplines. Specific fruits in different quadrants indicate that while some fruits (like apples) are extensively studied, there are opportunities for specialized research on other fruit types. Furthermore, positioning terms related to internal quality and modeling in the motor themes quadrant suggest a trend toward developing more comprehensive and predictive quality assessment methods. This could significantly improve fruit quality standards, reduce waste, and enhance consumer satisfaction.

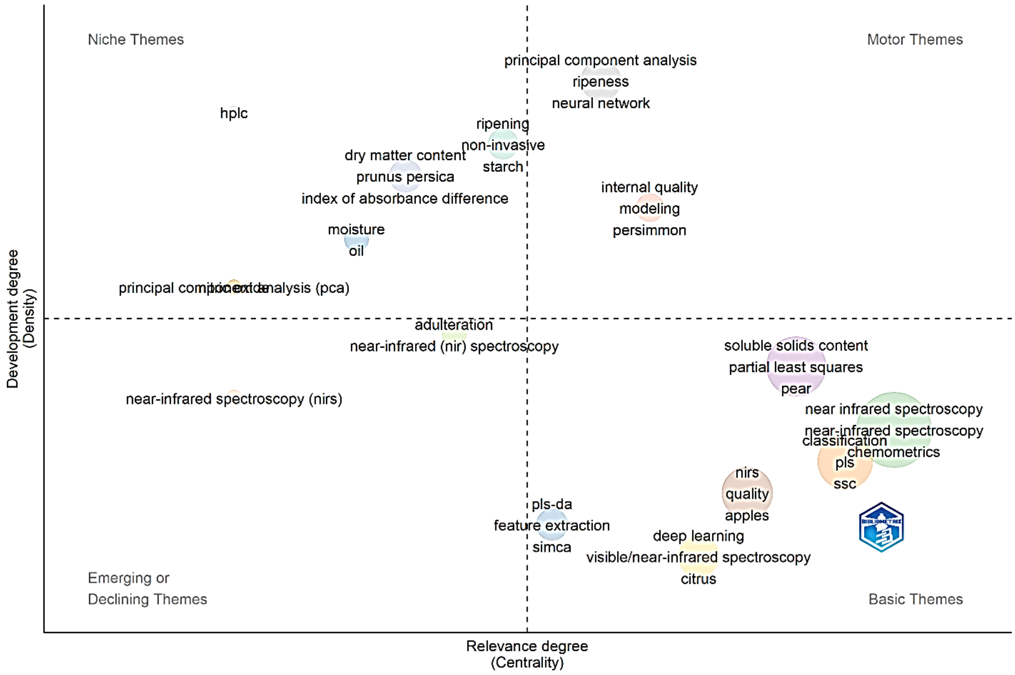


Figure 2.10: Thematic Map

Figure 2.11 presents the thematic evolution of keywords. This visualization provides critical insights into how the field has developed and where it may be heading. In the initial period (2003-2014), we see a focus on fundamental concepts and techniques. Keywords like "non-destructive," "dry matter," and "chemometrics" are prominent, indicating a primary concern with establishing basic methodologies for non-invasive fruit analysis. The presence of "partial least squares (pls)" suggests an early adoption of this statistical method for spectral data analysis. Interestingly, specific fruit types like "chestnut" appear, hinting at early applications to particular crops.

The middle period (2015-2020) shows a significant expansion in the breadth and depth of research topics. "Near-infrared spectroscopy" remains central, but we see the emergence of more sophisticated analytical techniques such as "principal component analysis," "multivariate analysis," and "feature selection." This suggests a growing emphasis on advanced data processing methods to extract more meaningful information from spectral data. The appearance of "vis-nir spectroscopy" indicates an expansion of the spectral range under study. Terms like "classification" and "quality" become more prominent, reflecting an increased focus on practical applications in fruit quality assessment.

In the most recent period (2021-2023), we observe a further evolution towards more specialized and advanced concepts. "Near-infrared spectroscopy" continues to be central. However, it is now accompanied by terms like "antioxidant," "absorption spectroscopy," and "fruits and vegetables," suggesting a broadening of application areas and a more nuanced understanding of fruit composition. The appearance of "visible/near-infrared spectroscopy" indicates a trend toward integrating multiple spectral ranges for more comprehensive analysis. Notably, "dry matter content" persists across all periods, underlining its enduring importance in fruit quality assessment.

The implications of this thematic evolution are significant for the field of fruit quality analysis and the broader agricultural industry. Firstly, it demonstrates a clear trend towards more sophisticated and comprehensive analytical approaches. The persistent presence of non-destructive methods throughout all periods underscores the ongoing importance of developing techniques to assess fruit quality without damaging the produce. This has crucial implications for reducing waste and improving efficiency in the fruit industry.

The evolution also reflects the increasing integration of advanced data analysis techniques, suggesting a growing need for interdisciplinary expertise combining spectroscopy, chemometrics, and data science. This trend is likely to continue, potentially leading to more accurate and predictive models for fruit quality assessment.

Expanding keywords related to specific fruit components (e.g., antioxidants) and broader categories (fruits and vegetables) implies a move towards more detailed and diverse applications of near-infrared spectroscopy. This could lead to more tailored quality assessment methods for different types of produce, potentially improving overall food quality and consumer satisfaction.

Furthermore, the persistent and evolving nature of specific keywords (like chemometrics and principal component analysis) across all periods indicates their fundamental importance to the field. However, the emergence of new terms in recent years suggests that the field is still dynamic and open to innovation.

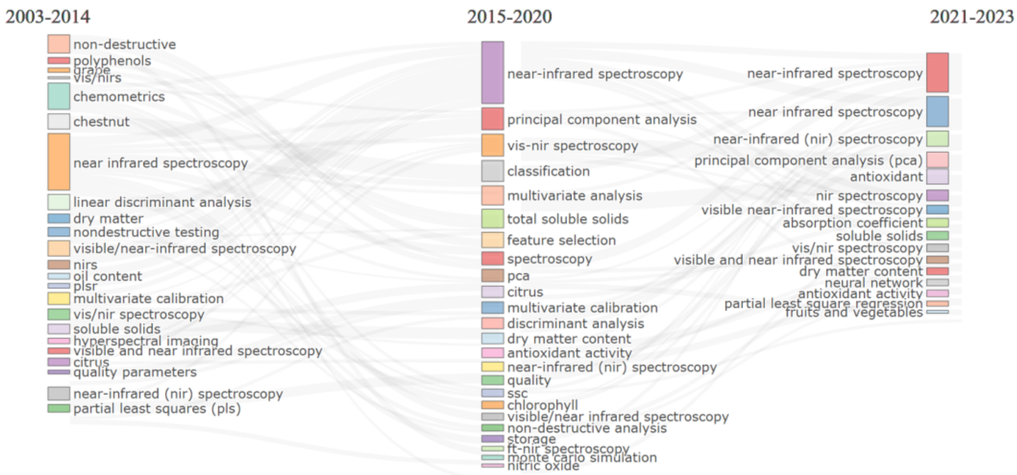


Figure 2.11: Thematic evolution

2.7.Journals, Institutions, and Citations

Figure 2.12 presents the top 20 journals based on their total publications in the field. Acta Horticulturae has the highest number of publications, with 65 articles. This highlights the strong connection between this technology and horticultural sciences. The prominence of journals like Postharvest Biology and Technology (56 articles) and Journal of Food Engineering (41 articles) underlines the significant application of NIRS in post-harvest management and food processing technologies. The Journal of Near Infrared Spectroscopy's high ranking (53 articles) indicates the technical depth and specialization of research in this area. The presence of broader scope journals such as Food Chemistry (44 articles) and Journal of the Science of Food and Agriculture (35 articles) suggests the wide-ranging implications of this technology across food science and agricultural research. Including journals focused on optics and spectroscopy, like Proceedings of SPIE and Infrared Physics and Technology, highlight the interdisciplinary nature of the field, bridging physics, engineering, and agricultural sciences.

Notably, the figure features journals dedicated to computational and electronic applications in agriculture (Computers and Electronics in Agriculture) and analytical methods (Food Analytical Methods), indicating the growing importance of data analysis and automation in this field. The diversity of journals, ranging from those focused on sensors and measurement techniques to those covering broader aspects of food science and horticulture, reflects the multiple applications of NIRS in fruit analysis. This spread across various disciplines implies that the technology is advancing in terms of spectroscopic methods and its practical applications throughout the fruit production and supply chain. The presence of highly specialized journals and those with a broader scope suggests that research in this field is simultaneously deepening our understanding of the technology while expanding its practical applications in the food industry.

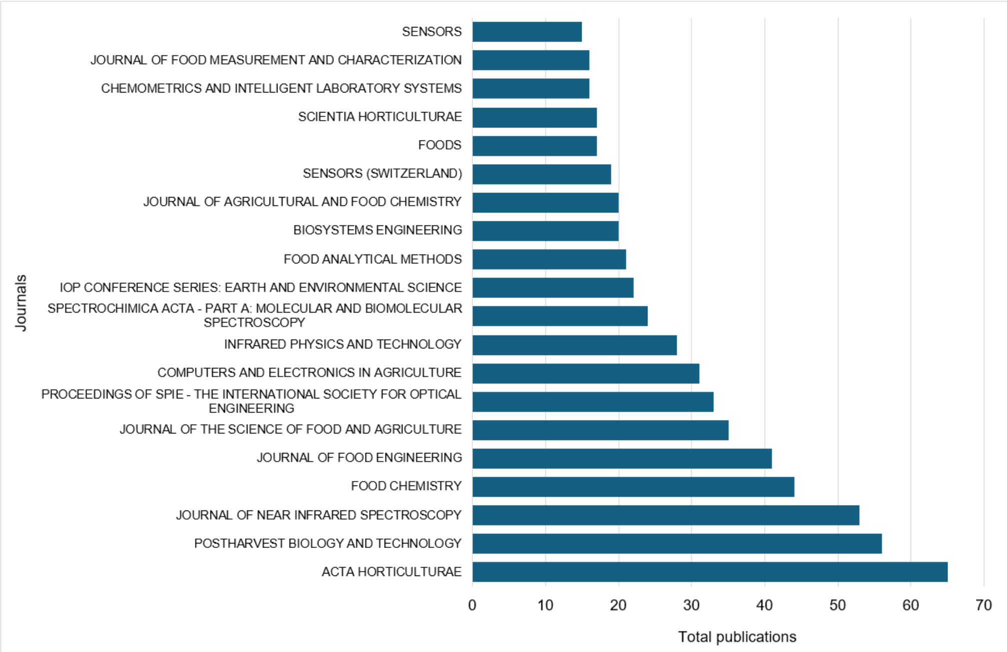


Figure 2.12: Top 20 Most Relevant Journals Based on Total Publications

Figure 2.13 displays the top 20 institutions actively involved in NIRS applications in fruit analysis research. It can be seen that Zhejiang University emerges as the clear leader in this field, with a substantial 280 publications, more than double the output of the next highest contributor. This dominance suggests that Zhejiang University has established itself as a major hub for research in this area, likely possessing advanced facilities and a dedicated team of experts. Chinese institutions, including East China Jiaotong University, Northwest A&F University, and Jiangsu University, feature prominently in the top ranks, indicating China's strong focus on this technology in agriculture. The international nature of this research is evident, with institutions from South Africa (Stellenbosch University), Spain (University of Cordoba), Italy (Università degli Studi di Milano), Thailand (Kasetsart University), Australia (Central Queensland University), Hungary (Szent István University), and the United States (Colorado State University) all appearing in the list.

This global distribution recognises NIRS's importance in fruit analysis across various agricultural contexts and climates. The presence of multiple agricultural universities, such as China Agricultural University, Nanjing Agricultural University, and Huazhong Agricultural University, underlines the direct application and relevance of this technology to agricultural practices and food science. The number of publications from these institutions suggests a concentrated effort to advance the field, potentially driven by the need for improved fruit quality assessment, sorting, and monitoring techniques in the global fruit industry. This research focus could have significant implications for enhancing fruit production efficiency, quality control, and consumer satisfaction in the international fruit market.

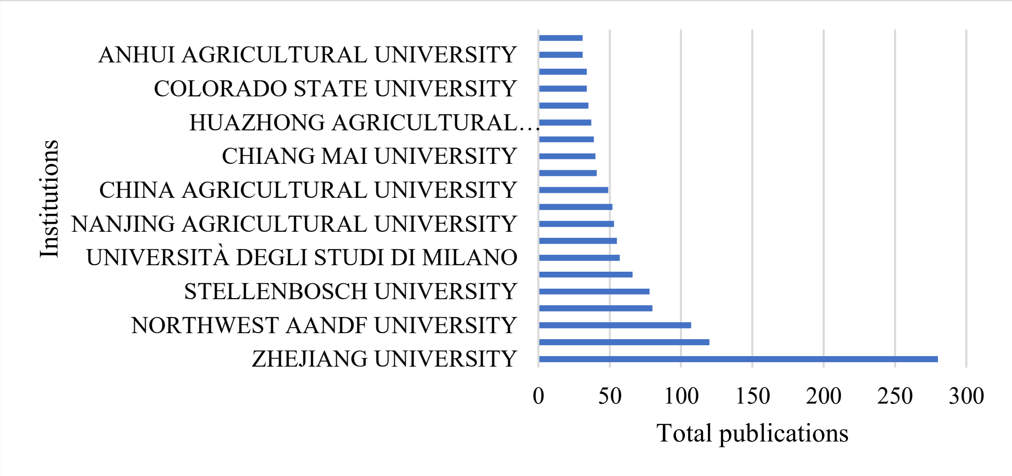


Figure 2.13: Top 20 Most Relevant Institutions Based on Total Publications

Figure 2.14 shows a comprehensive view of the impact of global research in the field as measured by the total citations per country. China has 9,368 citations, more than three times the citations of the next highest country. This dominance suggests that China is at the forefront of research and innovation in this area, likely driven by its large agricultural sector and emphasis on technological advancements in food production and quality control. Following China, a group of countries including Spain (2,616 citations), the USA (2,159), South Africa (1,965), Italy (1,753), and Australia (1,609) form a second tier of significant contributors, indicating a strong research presence in both European and non-European nations. The diversity of countries, spanning from Brazil to Japan and Iran to Canada, emphasizes the global relevance of this technology in fruit analysis. Interestingly, smaller countries like Romania (1,542 citations) and Ireland (1,031 citations) have a notably high impact relative to their size, suggesting focused research efforts or particularly influential studies in these nations.

The presence of both developed and developing countries highlights the widespread application and importance of NIRS in various agricultural and economic contexts. Countries with traditionally strong agricultural sectors, such as Brazil, Thailand, and Turkey, feature prominently, indicating the practical applications of this research in enhancing fruit production and quality assessment. The distribution of citations across multiple continents also suggests a global exchange of knowledge and collaboration in this field. However, the significant disparity in citation counts between the top-ranked countries and those at the bottom points to potential gaps in research capacity or focus among different nations. This map implies that while NIRS in fruit analysis is a globally relevant topic, there are clear leaders in the field. China is at the helm, driving innovation and setting research agendas that have far-reaching implications for global fruit production, quality control, and trade.

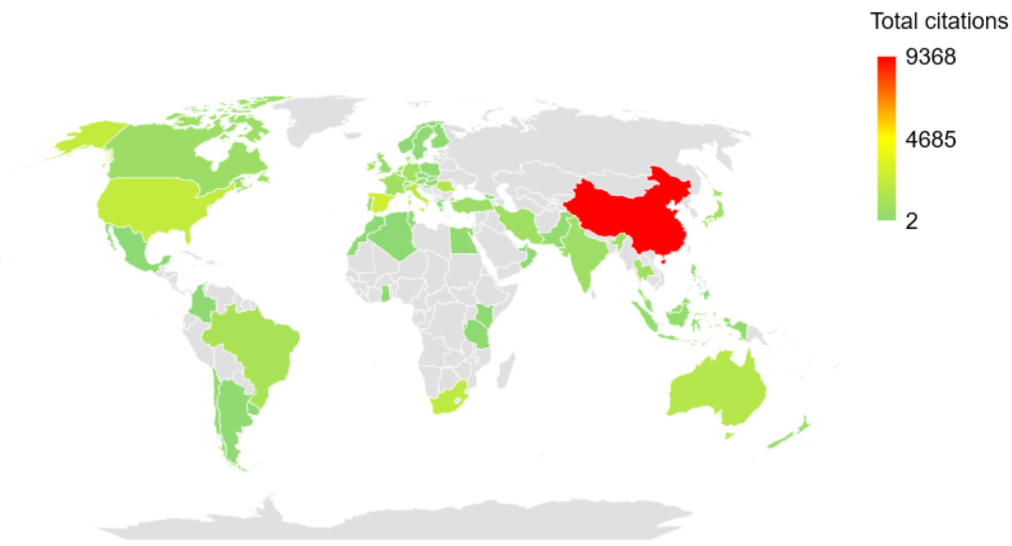


Figure 2.14: Total Citations per Country

2.8. Review of the Top 50 Most Relevant Papers on NIRS in Fruit Analysis

This section discusses the findings and recommendations from the top 50 most relevant papers included in the 1048 papers used for the bibliometric analysis. The selected studies span various fruit types, research methodologies, and practical applications, providing a comprehensive overview of NIRS applications in fruit analysis techniques. Some of these studies include Huang, Yu, Xu, and Ying (2008), who reviewed the use of NIR spectroscopy for on/in-line monitoring in the food and beverage industry over 30 years. The authors found that NIR has been successfully applied to various products like meat, fruits, grains, dairy, and beverages. The study highlighted the importance of chemometric treatment in interpreting NIR data. Dos Santos, Lopo, Páscoa, and Lopes (2013) assessed the applications of portable NIR instruments in the agro-food sector.

The study revealed that handheld NIR devices have been successfully used for multiple on-site applications, from assessing fruit quality to soil analysis. NIRQuest spectrometer and SCIO NIR spectrometer are displayed in Figure 2.15. Cortés, Blasco, Aleixos, Cubero, and Talens (2019) investigated using VIS-NIR spectroscopy for in-line monitoring of postharvest agro-food products. The authors found that while VIS-NIR shows great potential for real-time quality control, current research lacks real in-line applications.

Wang, Sun, Pu, and Cheng (2017) examined NIR applications in liquid food analysis. It was found that NIR is effective for analyzing various liquid foods and detecting adulteration. Qu et al. (2015) reviewed NIRS applications in food safety evaluation. The study found NIRS to be a promising technique for food safety inspection due to its speed and ease of use. Nicolaï et al. (2014) reviewed non-destructive measurement of fruit and vegetable quality, discussing various techniques for measuring internal and external quality attributes. Walsh, McGlone, and Han (2020) reviewed NIRS applications in postharvest decision support.

Several studies have focused on using NIR spectroscopy to assess various quality parameters in fruits, including soluble solids, acidity, firmness, and sugar content. For instance, Gómez, He, and Pereira (2006) used Vis/NIRS to predict soluble solids content (SSC), acidity (pH), and firmness in Satsuma mandarins. The authors achieved good prediction models, with the best model for SSC having an r^2 of 0.94 and RMSEP of 0.33 Brix. Saranwong, Sornsrivichai, and Kawano (2004) used NIR to predict the eating quality of ripe mangoes. Dry matter and starch content at harvest, measured by NIR, were good predictors of soluble solids content in ripe mangoes. Li, Huang, Zhao, and Zhang (2013) used Vis/NIR spectroscopy to predict pears' SSC, pH, and firmness. The findings indicate that LS-SVM models outperformed PLS models and that using effective wavelengths improved prediction accuracy. Wu, He, Nie, Cao, and Bao (2010) used Vis-NIR spectroscopy to predict pH and SSC in grape juice. They found that UVE-SPA-MLR models outperformed other models in predicting these parameters.

Golic and Walsh (2006) used NIRS to assess the internal quality of stone fruits. The study found good calibration models for total soluble solids across different varieties. Xu, Qi, Sun, Fu, and Ying (2012) investigated the online determination of soluble solid content in pears using Vis-NIR spectroscopy. The results revealed that GA-SPA-MLR on selected wavelengths provided good prediction accuracy. Fan, Zha, Du, and Gao (2009) used NIR spectroscopy to measure SSC and firmness in apples. The findings revealed that fruit orientation and light source influenced prediction results, with multi-lamp setups improving accuracy. Walsh, Golic, and Greensill (2004) evaluated NIRS to assess various fruits' SSC and dry matter content. The authors found the technology well-suited for some fruits (e.g., apples) and less accurate for others (e.g., papaya). Moghimi, Aghkhani, Sazgarnia, and Sarmad (2010) used Vis/NIR spectroscopy to predict SSC and pH in kiwifruit. They found good prediction models using SNV with median filter and first derivative preprocessing.

Bureau et al. (2009) evaluated NIRS for predicting apricot quality traits. The results show a good prediction performance for soluble solids and titratable acidity but less accurate for other traits. Larraín, Guesalaga, and Agosín (2008) developed an NIR instrument to measure ripeness parameters in wine grapes. The study found good Brix and pH measurement results, with promising results for anthocyanin concentration. McGlone, Fraser, Jordan, and Künnemeyer (2003) compared different NIR measurement modes for predicting SSC and titratable acidity in Satsuma Mandarins. The authors found the direct transmission mode most accurate for SSC prediction. Sun, Lin, Xu, and Ying (2009) evaluated the online measurement of pear SSC using Vis/NIR spectroscopy. It was observed that fruit moving speed had little effect on model performance at speeds of 0.3-0.7 m/s. Rungpichayapichet, Mahayothee, Nagle, Khuwijitjaru, and Müller (2016) evaluated the effect of harvest season on NIRS calibrations for mango quality prediction. The authors highlighted that combining data from multiple years greatly enhanced prediction accuracy.

Other studies have applied NIRS to detect and quantify specific compounds in fruits and vegetables, such as carotenoids, sugars, and polyphenols. For example, Baranska, Schütze, and Schulz (2006) used FT-Raman, ATR-IR, and NIR spectroscopy to quantify lycopene and β-carotene in tomatoes. The authors found that IR spectroscopy gave the best prediction quality. Xie, Ye, Liu, and Ying (2009) used NIR spectroscopy to determine glucose, fructose, and sucrose in bayberry juice. The study found PLS models good for predicting these sugar concentrations. Zhang et al. (2020) applied NIR hyperspectral imaging with deep learning to determine chemical compositions in black goji berries. It was observed that deep learning approaches performed well for both modeling and feature extraction. Baranski, Baranska, and Schulz (2005) applied NIR-FT-Raman spectroscopy for in situ analysis of carotenoids in living plant samples.

The study successfully detected carotenoid changes in various plant tissues. Pissard et al. (2013) employed NIRS to determine apples' vitamin C and total polyphenol content. The study achieved good prediction performance, especially for polyphenol and sugar content. Clément, Dorais, and Vernon (2008) measured lycopene content and other quality parameters in intact tomatoes via NIRS and found accurate predictions for lycopene content and colour variables. Janik, Cozzolino, Dambergs, Cynkar, and Gishen (2007) compared PLS regression and artificial neural networks for predicting anthocyanin concentration in red-grape homogenates. The results revealed that the ANN outperformed PLS, especially for new vintage samples.

Several studies have employed advanced techniques or methodologies in conjunction with NIR spectroscopy, such as hyperspectral imaging, data fusion, and novel modeling approaches. In view of this, Magwaza et al. (2012) focused on NIR spectroscopy for non-destructive quality assessment of citrus fruits. The study found that NIR can measure both internal and external quality attributes of citrus fruit, reducing waste and allowing repeated measurements on the same fruit over time. The authors also highlighted the potential of emerging technologies like multispectral and hyperspectral imaging. Opara and Pathare (2014) reviewed technologies for bruise measurement in fresh produce. It was found that advanced non-invasive technologies like NIR spectroscopy, hyperspectral imaging, thermal imaging, and MRI show promise for bruise detection. Lee et al. (2014) used NIR hyperspectral imaging to detect bruises on pears. They achieved 92% accuracy in bruise detection using waveband ratio analysis.

Mendoza, Lu, and Cen (2012) evaluated multi-sensor data fusion for predicting apple firmness and SSC. The authors revealed that sensor fusion improved predictions compared to individual sensors. Chen, Qiao, Xu, Feng, and Cai (2019) developed a fuzzy optimisation strategy for RBF LSSVR models in Vis-NIR analysis of pomelo fruit maturity. The results revealed that the strategy effectively reduces computational complexity and improves predictions. Nogales-Bueno, Hernández-Hierro, Rodríguez-Pulido, and Heredia (2014) used hyperspectral imaging to determine grapes' phenolic and technological maturity parameters. The authors achieved good prediction results for both red and white grapes.

Malegori et al. (2017) compared a miniaturized NIR device with a benchtop FT-NIR spectrometer for measuring quality parameters in acerola fruit. The study found that the portable device performed well, especially when using SVM algorithms. Camps and Christen (2009) used portable NIR spectroscopy to determine apricot quality parameters. The authors achieved good results for SSC prediction, with variable results for firmness and titratable acidity. Marques, de Freitas, Pimentel, and Pasquini (2016) evaluated a handheld NIR spectrometer for quality control of 'Tommy Atkins' mangoes and found good prediction models for soluble solids and dry matter content. Chen and Opara (2013) investigated food texture measurement techniques. The results revealed that while sensory evaluations are useful, instrumental methods like NIR spectroscopy are increasingly used for rapid and cost-effective texture measurement. Magwaza and Opara (2015) reviewed methods for determining sugar content and sweetness in horticultural produce. It was seen that while chromatographic techniques are accurate, NIR spectroscopy provides a rapid, simple, and cost-effective alternative for routine sugar analysis.

Arendse, Fawole, Magwaza, and Opara (2018) reviewed non-destructive measurement techniques for fruits with thick rinds. It was found that thick rinds can interfere with NIR measurements of internal quality, but other techniques like X-ray micro-CT and NMR show promise. Sankaran, Mishra, Maja, and Ehsani (2011) employed VIS-NIR spectroscopy to detect Huanglongbing disease in citrus trees. The findings achieved high classification accuracies (up to 98%) using quadratic discriminant analysis. Xie, Ying, Ying, Yu, and Fu (2007) applied VIS-NIR spectroscopy to detect transgenic tomatoes. The authors achieved 100% correct classification using PLSDA after derivative spectral pre-treatment.

Fan, Zhang, Li, Huang, and Wang (2016) assessed the influence of spectrum measurement position on the NIR analysis of apple SSC. They found that a global position model with effective wavelengths gave the best prediction results. de Oliveira, Bureau, Renard, Pereira-Netto, and de Castilhos (2014) investigated the impact of fruit structure on NIR predictions of SSC and titratable acidity. The study found good apricot results but less accurate predictions for fruits with thick skins or heterogeneous structures. Aday and Caner (2014) used NIR spectroscopy to evaluate the effects of various treatments on strawberry storage life. Fu, Ying, Lu, and Xu (2007) compared transmission and reflectance modes of VIS/NIR spectroscopy for detecting brown hearts in pears. The results revealed a transmission mode more effective, achieving 91.2% classification accuracy. Magwaza and Tesfay (2015) reviewed methods for determining avocado fruit maturity. The authors highlighted emerging optical and imaging techniques promising for non-destructive avocado maturity and quality analysis. Guo, Ni, and Kokot (2016) used NIRS and chemometrics to analyse and classify jujube fruit from different origins. They found that LS-SVM models produced the best quantitative prediction results. Ncama, Opara, Tesfay, Fawole, and Magwaza (2017) used Vis-NIR spectroscopy to quantify flavor-related parameters in oranges and grapefruits. They achieved good prediction models for BrimA, TSS:TA ratio, and TSS, demonstrating the potential of Vis/NIRS for non-destructive flavor assessment.

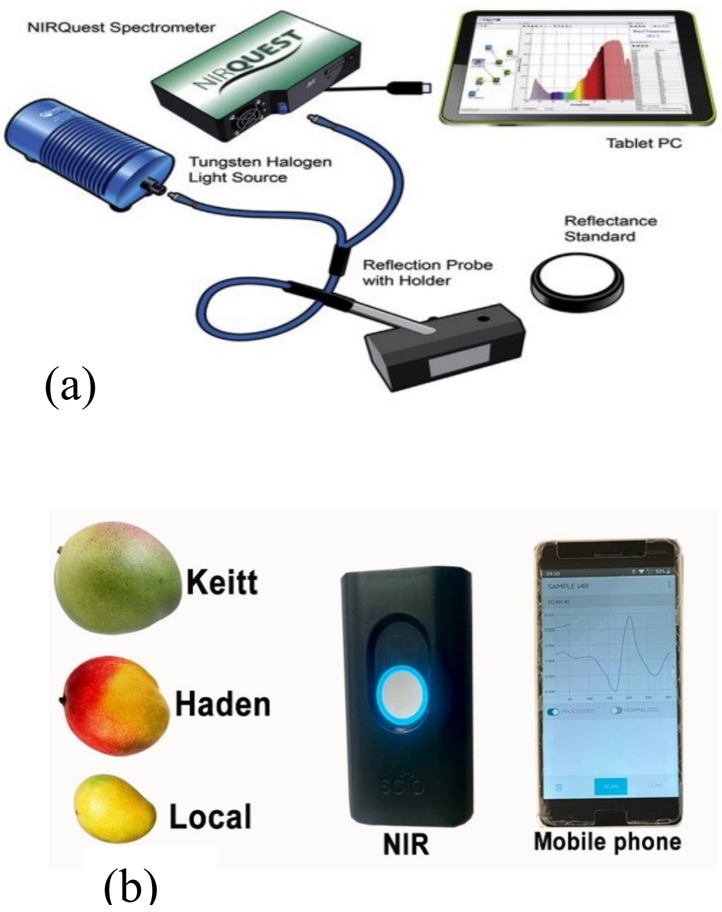


Figure 2.15: (a) NIRQuest spectrometer (Raghavendra et al., 2021) (b) SCIO NIR spectrometer with the three varieties of mango fruits (Lamptey, Teye, Abano, & Amuah, 2023) (Published under open access)

2.9 Advances in NIRS in Mango Fruit Analysis

This section comprehensively discusses the advancement in NIRS applications in mango fruit analysis. It categorized the papers included in the systematic review into the following sections: (1) mango quality assessment and prediction using NIRS, (2) non-destructive testing and monitoring, (3) spectral data analysis and model optimisation, (4) technological innovations and instrumentation and (5) Applications of machine learning and artificial intelligence.

2.9.1 Mango Quality Assessment and Prediction using NIRS

NIRS has proven to be a valuable tool for assessing various quality attributes in mangoes, providing a non-destructive and rapid approach for predicting key parameters like total acidity (TA), soluble solids content (SSC), fruit maturity, and internal physiological disorders. Across the reviewed studies, different regression models, pre-processing techniques, and prediction models have been employed, each contributing valuable insights to the field. For example, Munawar, Meilina, and Pawelzik (2022) demonstrated the potential of NIRS for predicting TA in intact mangoes, with ANN outperforming other models like support vector machine regression (SVMR) and partial least squares regression (PLSR).

The study highlighted that using the first four principal components (PCs) as input to the ANN model provided optimal prediction accuracy. This high degree of correlation ($r^2cal = 0.97$, $r^2pred = 0.89$) indicates the robustness of ANN models, suggesting that they can better capture the non-linear relationships within the spectral data compared to traditional methods like PLSR. This is consistent with findings from Shah et al. (2020), which compared PLSR and non-linear models (e.g., ANN, LSSVM) for fruit maturity estimation across various fruits. Shah et al. (2020) found that while PLSR is commonly used, non-linear models often outperform it, particularly for complex fruits like mangoes. Both studies underline the need for more advanced models to maximize the accuracy of NIRS predictions.

On the other hand, Nordey, Davrieux, and Léchaudel (2019) highlighted a different challenge in predicting postharvest mango quality. The study found weak correlations between quality traits measured at harvest and those measured after ripening, except for dry matter content (DM), which showed a moderate correlation ($r^2 = 0.61$). Interestingly, pulp colour was identified as a strong indicator of fruit shelf life ($r^2 = 0.7$). These results suggest that although NIRS can predict some traits with moderate accuracy, post-ripening quality indicators are more sensitive to environmental and growing conditions, complicating predictions. This finding emphasizes distinguishing traits that can be reliably predicted at harvest and those that require postharvest assessment. Rungpichayapichet et al. (2017) took a novel approach by combining hyperspectral imaging (HSI) with NIRS to predict firmness, TA, and SSC in mangoes. Their study showed significant correlations between HSI data and mango firmness ($r^2 = 0.81$) and TA ($r^2 = 0.81$), with a moderate correlation for SSC ($r^2 = 0.5$).

The development of prediction maps enabled visualization of the spatial distribution of quality attributes within the fruit. This is particularly relevant for the industry, where non-destructive techniques like HSI could enhance grading processes by detecting quality variations across the fruit. Compared to the purely spectral approach of Polinar, Yaptenco, Peralta, and

Agravante (2019), which focused on using PLSR to predict DM and SSC in 'Carabao' mango, the HSI method offers additional spatial information that could improve industrial sorting and grading capabilities. Further emphasizing industrial applications, Guru, Raghavendra, and Rao (2021) explored postharvest handling using AI-based solutions. Their work builds upon the foundational NIRS-based research by incorporating machine learning and deep learning approaches to automate the sorting and grading process. The comparative study between different AI methods demonstrated the promise of AI in improving postharvest handling efficiency, showing alignment with Polinar et al. (2019) in the potential for automation in the fruit supply chain.

The importance of pre-processing methods in improving the accuracy of NIRS models was highlighted by both Purwanto et al. (2015) and Phuangsombut, Phuangsombut, and Terdwongworakul (2020). Purwanto et al. (2015) worked on developing calibration models for predicting both SSC and pH in 'Gedong Gincu' mangoes. The study found that proper pre-processing, such as smoothing and first derivative Savitzky-Golay filtering, significantly enhanced the accuracy of the models. For SSC, the best results were achieved using the smoothing method ($r = 0.82$), while for pH, first derivative filtering improved accuracy ($r = 0.74$). On the other hand, Phuangsombut et al. (2020) focused solely on predicting SSC in peeled and unpeeled mangoes.

The study demonstrated that the accuracy of SSC predictions was higher for peeled fruit ($r = 0.88$) than for unpeeled fruit ($r = 0.84$). Applying an empirical approach to account for differences between the peel and flesh spectra further improved SSC predictions ($r = 0.87$), highlighting the importance of pre-processing in refining NIRS models for SSC prediction.

These findings align with Munawar, Hörsten, Mörlein, Pawelzik, and Wegener (2013), who found that spectral pre-processing (SNV for SSC and MSC for TA) improved the calibration model performance for predicting mango sweetness and acidity. These results demonstrate that pre-processing is critical in enhancing NIRS model accuracy beyond the choice of prediction models (PLSR, ANN, etc.). Mogollón et al. (2020) contributed valuable insights into detecting internal physiological disorders like jelly seed and black flesh using Vis-NIR spectroscopy. Although the study showed moderate accuracy (Logistic model accuracy of 71% after storage), it highlighted the challenge of differentiating between various internal disorders, suggesting that future research should focus on refining spectral models to better distinguish between these conditions.

2.9.2. Non-Destructive Testing and Monitoring

This section discusses studies on non-destructive techniques using NIR and Visible-NIR (VNIR) spectroscopy to assess mangoes' internal quality without damaging the fruit. These methods provide rapid and reliable predictions of key quality parameters, allowing producers and suppliers to make informed decisions, reduce postharvest losses, and improve product quality. The findings across several studies show how these technologies are applied for various purposes, from internal browning detection to quality determination. Some studies include Gabriëls, Mishra, Mensink, Spoelstra, and Woltering (2020), who demonstrated the effectiveness of VNIR spectroscopy, combined with Artificial Neural Networks (ANN), in detecting internal browning in mangoes, achieving a classification accuracy of over 80%.

The ability to non-destructively differentiate between healthy and browned mangoes represents a significant advancement for postharvest quality control, where internal disorders are often difficult to detect through external inspection. A robust classification system like this can minimize postharvest losses by enabling early detection and removal of affected fruit from the supply chain. Similarly, Zakaria et al. (2021) used NIR and ANN models to detect Insidious Fruit Rot (IFR) in Harumanis mangoes, achieving an impressive prediction accuracy of 98.05%. This high level of accuracy demonstrates the potential of NIR combined with machine learning techniques in identifying internal fruit disorders, further supporting the findings of Gabriëls et al. (2020). Such early detection technologies are vital for ensuring the quality of exported fruits, particularly those destined for long-distance markets where internal disorders could go unnoticed.

Marques et al. (2016) explored the feasibility of using handheld NIR spectrometers for assessing quality parameters such as soluble solids (SS), dry matter (DM), titratable acidity (TA), and pulp firmness (PF) in 'Tommy Atkins' mangoes. The study achieved high coefficients of determination for most parameters, with the SS model showing an r^2 value of 0.92, indicating that portable NIR devices can effectively monitor physico-chemical changes in fruit. This aligns with the findings of Taira et al. (2017), who also demonstrated that portable NIR devices help evaluate mango quality non-destructively.

Developing portable devices is particularly important for field applications, allowing real-time decision-making in quality control and harvest timing. Subedi and Walsh (2011) evaluated the accuracy of visible-short wavelength NIR (VIS-SWNIR) for predicting DM and total soluble solids (TSS) in mangoes. While the study found good performance for DM prediction across all ripening stages ($r^2 > 0.75$), TSS was accurately predicted only in ripened fruit. This outcome suggests that VIS-SWNIR may have limitations in distinguishing between starch and soluble sugars during the early stages of ripening. This finding contrasts Saranwong et al. (2004), who demonstrated that NIR spectroscopy could accurately predict ripe-stage eating quality (SSC) based on DM and starch content measured at harvest. This discrepancy may be attributed to differences in the spectral range or model calibration approaches used in the two studies.

Integrating NIR spectroscopy with deep learning models for quality prediction has also shown promise. Nuanmeesri and Poomhiran (2022) combined image-based deep learning with NIR spectral data to classify the sweetness of ripe mangoes. Using enhanced spectral data, their model achieved a training accuracy of 99.66% and a validation accuracy of 94.20%. The study illustrates the potential of combining spectral data with advanced machine learning techniques to improve prediction accuracy. This approach aligns with the findings of Jha et al. (2012), who showed that NIR could predict TSS and pH in mangoes with reasonable accuracy, although their models did not achieve the same level of precision. Watanawan, Wasusri, Srilaong, Wongs-Aree, and Kanlayanarat (2014) examined the relationship between harvest maturity and NIRS values in export mangoes, focusing on how harvest quality correlates with NIRS readings and ripe-stage quality attributes. The study found that NIRS values were strongly correlated with dry matter content ($r^2 = 0.96$) and negatively correlated with fruit firmness ($r^2 =$

0.99). Moreover, the NIRS model accurately predicted TSS in ripe fruit with 99% accuracy.

This study demonstrates the utility of NIRS in determining optimal harvest maturity, particularly for export mangoes, ensuring better consistency in fruit quality. Rungpichayapichet et al. (2023) compared the use of NIR and hyperspectral imaging (HSI) to evaluate the internal quality of mangoes during ripening. While NIR provided better prediction accuracy for attributes like firmness and soluble solids, HSI was valuable for assessing spatial variation in fruit quality. The combination of NIR and HSI could potentially provide a more holistic approach to mango quality assessment, improving sorting and grading efficiency in industrial applications.

2.9.3. Spectral Data Analysis and Model Optimisation

This section discusses studies that focused on spectral data and the optimisation of models for predicting fruit quality, particularly using NIRS. Various studies have demonstrated that combining spectral pre-processing, advanced algorithms, and model tuning can significantly enhance the prediction accuracy of key quality traits in mangoes and other fruits. For instance, Mishra, Rutledge, Roger, Wali, and Khan (2021) examined the effect of chemometric pre-processing on NIR spectra. They found that pre-processing can sometimes reduce models' predictive performance contrary to traditional assumptions. Partly least squares (PLS) and deep learning (DL) models performed better with raw absorbance than pre-processed data.

The best results were achieved with DL models, which had a 13% lower root mean squared error of prediction (RMSEP) compared to PLS models. These findings suggest that raw spectral data contain useful scattering information that could be lost through excessive pre-processing, highlighting the need to carefully consider pre-processing techniques in spectral data analysis. This contrasts with more conventional approaches where pre-processing has been considered essential for enhancing prediction accuracy. In contrast, Khumaidi and Raafi'udin (2022) found that pre-processing, particularly spectral transformation methods such as smoothing and scatter correction, significantly improved classification accuracy for mango cultivars. Their study further showed that balancing the dataset using oversampling techniques like SMOTE (synthetic minority oversampling technique) led to higher classification accuracy in machine learning models, with Support Vector Machine (SVM) models achieving 100% accuracy.

Passos and Mishra (2023) explored the application of convolutional neural networks (CNNs) for dry matter (DM) prediction in fresh fruit. The study findings demonstrated that CNNs outperformed traditional PLS and locally weighted PLS (LW-PLS) models, particularly when many training samples were used (around 500). The study highlights that CNNs, when properly optimized, offer a significant advantage in handling complex spectral data and generalizing across diverse fruit types. This result aligns with Mishra and Passos (2021), who also found that DL models could outperform PLS models when updated to account for seasonal variability, especially when large datasets are available. These studies collectively point to the growing importance of DL models in NIR spectral analysis, as they can provide robust and generalizable predictions with less reliance on traditional chemometric techniques.

The introduction of a model updating approach by Yang et al. (2022) further highlights the potential of DL models in handling variations across multiple seasons and sample sizes. Their model fine-tuning approach, which involved using limited samples to update the calibration, improved prediction accuracy across multiple fruit types, including mangoes. This dynamic updating process reduced RMSE by at least 9.2% compared to traditional global models or recalibration methods, demonstrating the effectiveness of deep learning in managing variability in spectral data collected across different growing conditions and seasons.

The integration of machine learning techniques into spectral data analysis has also been explored by Chen et al. (2023), who employed Gaussian Process Regression (GPR) to improve the prediction of fruit traits. GPR outperformed PLSR, reducing RMSE by 14% in in-distribution tests. Additionally, incorporating variance into the model allowed for eliminating poorly predicted samples, further improving model accuracy. This approach presents an alternative to standard one-class classification methods, showing the potential of GPR for more refined spectral data analysis. Wohlers, McGlone, Frank, and Holmes (2023) demonstrated the value of data augmentation in enhancing the robustness of CNN models for NIR spectral analysis. By augmenting training data to mimic spectra collected from multiple devices, their study reduced overfitting and improved the generalization of the models.

In addition to the focus on model improvement, Funsueb, Thanavanich, Theanjumpol, and Kittiwachana (2023) introduced a flexible approach to calculating Fruit Quality Indices (FQIs) by aggregating several quality parameters (e.g., total soluble solids, titratable acidity, firmness, dry matter, etc.). The use of NIR spectroscopy combined with PLS regression allowed for rapid, non-destructive prediction of these indices, showing the broad applicability of NIR in fruit quality monitoring. Hayati, Munawar, and Fachruddin (2020) and Munawar, Hayati, and Fachruddin (2021) demonstrated the importance of enhancing NIR datasets and employing optimized spectral datasets to improve model performance in predicting quality parameters such as vitamin C and total acidity in mangoes. These studies highlight that even within traditional chemometrics, careful optimization of spectral datasets can significantly improve prediction accuracy and robustness, showing that data refinement remains a critical component of spectral analysis.

2.9.4. Technological Innovations and Instrumentation

The studies included in this section emphasized the transition toward more portable, efficient, and accurate tools for non-destructive fruit quality evaluation. The development of neural networks for spectroscopy has led to significant advancements in model accuracy and reliability. Dirks and Poole (2022) demonstrated that ensembling techniques combined with hyperparameter optimisation enhance the performance of neural networks in processing VNIR spectroscopy data. Using the latest data samples for validation further improved model robustness, making it a noteworthy contribution to predictive modeling in spectroscopy. This innovation opens possibilities for more automated configurations of machine learning models, ensuring higher accuracy in time-based predictions.

The automated configuration not only minimizes the need for manual intervention but also makes real-time quality control more feasible in agricultural processes. Cheng et al. (2019) contributed significantly by integrating VIS/NIR spectroscopy and ASCA analysis to assess mango drying processes. The study findings highlight the potential of spectral techniques to evaluate colour and chemical changes in mangoes under different drying conditions. The use of longwave NIR spectra to measure chemical alterations and shortwave VIS/NIR spectra to assess fruit maturity and batch effects highlights the versatility of these techniques. The study's findings demonstrate that the tunnel dryer provided a more consistent drying process than traditional dryers, emphasizing the importance of controlled drying environments in preserving mangoes' nutritional and chemical quality. This supports the broader use of NIR in optimizing postharvest processing methods.

The exploration of miniaturized NIR spectrometers by Praiphui and Kielar (2023) demonstrates their potential for use in small-scale operations and portable devices. Although SCiO and Linksquare instruments showed promise for parameters like dry matter (DM), total soluble solids (TSS), and pH, other devices such as DLP NIRscan Nano underperformed. This suggests that while miniaturized spectrometers hold potential, further optimisation is needed for broader parameter evaluation. The development of such devices aligns with the industry's need for cost-effective, portable tools to deliver rapid and reliable quality assessments in real-time. Developing a microcontroller-based portable NIR device by Izneid and Al-kharazi (2013) represents a step forward in creating affordable, user-friendly tools for smallholder farmers or industry operators. The voltage readings correlated with mango ripening stages offer a simple yet effective method of assessing fruit quality, which is vital for postharvest monitoring.

The ability of NIR spectroscopy to detect internal mango disorders, as demonstrated by Seehanam et al. (2022), provides a new avenue for improving postharvest management by identifying defects early. While the linear discriminant analysis (LDA) model achieved moderate accuracy, the artificial neural network (ANN) showed improved classification rates, especially in distinguishing black-streaked vascular tissue (BSV) from internal breakdown (IBD). This study emphasizes the relevance of non-linear analysis techniques, such as ANN, in enhancing the classification accuracy of internal disorders. Integrating fuzzy logic with NIR data to classify mango maturity presents an innovative method for handling overlapping maturity classes. Khumaidi, Purwanto, Sukoco, and Wijaya (2022) demonstrated that this approach significantly enhances the accuracy of maturity classification, particularly in complex datasets where traditional classification methods may fail. Combining fuzzy logic with machine learning techniques, such as support vector machines (SVM) and partial least squares (PLS), offers a robust framework for non-destructive quality assessment in mangoes.

Tan and Chia (2023) provided valuable insights into the effects of pre-processing methods on machine learning model performance in spectroscopy. Their work shows that artificial neural networks (ANNs) and principal component analysis (PCs-ANN) offer a robust, non-destructive internal quality prediction solution. The comparative performance of ANN and PCs-ANN in different validation and prediction sets highlights the importance of pre-processing in optimizing model accuracy. The findings suggest that ANN, despite its tendency to overfit, can be effectively applied to spectroscopy when combined with the right data pre-processing techniques.

The ability to non-destructively determine vitamin C content using NIR reflectance spectroscopy, as shown by Munawar, Hayati, and Wahyuni (2019), further supports the growing application of NIRS in rapid fruit quality assessment. The high correlation between NIR readings and vitamin C levels demonstrates that spectroscopy can offer qualitative and quantitative evaluation of key nutritional components in fruits. This finding is particularly relevant for enhancing the nutritional value monitoring of mangoes and other perishable fruits in the supply chain.

Shah et al. (2021) successfully developed a hand-held NIR device to classify the maturity state of mangoes with notable accuracy using K-nearest neighbors (KNN) algorithm. This portable device has the potential to revolutionize in-field maturity assessments, allowing farmers and suppliers to determine optimal harvest times non-destructively. With the accuracy achieved, such tools can reduce waste by minimizing premature or delayed harvesting and optimizing the supply chain from farm to market. Fourier transform NIR spectroscopy, as applied by (Munawar, von Hörsten, Wegener, Pawelzik, & Mörlein, 2016), demonstrated excellent potential for predicting soluble solids content (SSC), titratable acidity (TA) and ascorbic acid (AA) in mangoes. Using scatter-corrected spectra with multivariate calibration models (PLS and PCR) showed promising results, further validating the utility of NIR in non-destructive chemical analysis. The study adds to the growing evidence that NIRS, combined with chemometric techniques, is an effective method for rapid quality assessment in the food industry.

2.9.5. Applications of Machine Learning and Artificial Intelligence

This section explores the applications of machine learning (ML) and artificial intelligence (AI) in conjunction with near-infrared spectroscopy (NIRS) for fruit classification, quality prediction, and ripening detection. The integration of ML techniques, including deep learning models, has shown significant potential in enhancing the accuracy and efficiency of these processes. Several studies have demonstrated the effectiveness of AI models in fruit recognition and quality assessment. For instance, Khanh Ninh, Doan, Khanh Ninh, Xuan Nguyen-Thi, and Le Thanh (2021) utilized a combination of NIRS and deep neural networks (DNNs) to classify different fruit types, achieving a high accuracy of approximately 99% for recognizing five fruit types, including apple, avocado, and mango. This study highlighted the superiority of DNN models, especially ResNet-based architectures, over traditional classifiers like k-nearest neighbors and support vector machines. Including spectral derivatives further improved the model's recognition accuracy by over 8%, indicating the importance of advanced feature extraction techniques in boosting model performance.

Similarly, Chia and Suarin (2022) investigated the use of neural networks and XGBoost in quality prediction using a large dataset of NIR spectral data. Their findings revealed that while XGBoost achieved satisfactory accuracy, the Bayesian regularized neural network significantly outperformed it, with higher r^2 values and lower error margins. This suggests that neural networks, particularly those with regularization, are more effective in handling large and complex datasets in NIR spectroscopy applications, making them a preferred choice for predicting fruit quality parameters.

In addition to these models, Solihin, Zekui, Ang, Heltha, and Rizon (2021) introduced a visual programming approach using Orange Data Mining software to calibrate NIR data, demonstrating its accessibility for non-programmers. This study emphasized the practical use of ML for NIRS data calibration, making sophisticated chemometric techniques more accessible to a broader audience, particularly in agricultural sectors where programming expertise may be limited. Moreover, Zeb, Qureshi, Ghafoor, and O'Sullivan (2022) focused on using shortwave NIR spectroscopy and the QDA classifier for fruit classification. The study achieved a test data accuracy of 93.02%, showing that fruit classification can be efficiently performed using key spectral features related to O-H and C-H overtones. This method has practical implications for developing LED-based devices for automated fruit classification, highlighting the potential of NIR-ML integration in commercial applications.

Mishra and Passos (2021) explored transfer learning (TL) for updating deep learning models in spectral data processing in the context of model scalability. Their results indicated that TL approaches successfully adapted models to new scenarios, outperforming recent benchmarks. This suggests that TL can be crucial in making DL models more flexible and widely applicable, especially in dynamic environments such as agriculture, where data may change across seasons and geographical locations. Furthermore, Ulya, Chamidah, and Saifudin (2021) employed a multi-predictor local polynomial regression model to predict Avomango sweetness, achieving a mean absolute percentage error (MAPE) of 8.554%, classified as highly accurate.

Regarding food safety, Lakade et al. (2019) developed a novel NIR method for detecting artificially ripened mangoes using calcium carbide. The method distinguished between naturally and artificially ripened mangoes, with ICP-MS confirming higher arsenic levels in artificially ripened fruits. This rapid, non-invasive approach demonstrates how AI and NIRS can ensure food safety by detecting harmful substances in fruit ripening processes. Watanawan, Wasusri, Wongs-Aree, Srilaong, and Kanlayanarat (2012) evaluated the optimal maturity stage for harvesting 'Nam Dok Mai' mangoes using handheld NIR devices. Their results correlated NIR values with key quality parameters like dry matter content and total soluble solids, providing a non-destructive method for determining the optimal harvesting time for export-quality mangoes. This method underscores the utility of AI-NIR integration in improving postharvest decision-making and reducing fruit losses during export.

2.10. Summary and Directions for Future Research

The bibliometric analysis highlights the field's multidisciplinary nature, combining spectroscopy, chemometrics, agricultural science, and, increasingly, data science and artificial intelligence. Key research themes include non-destructive quality assessment, prediction of various fruit parameters (e.g., soluble solids content, firmness, acidity), and the application of advanced data analysis techniques. The evolution of research topics shows a clear shift from basic concepts and techniques in the early years (2003-2014) to more sophisticated approaches in recent years (2021-2023), including integrating machine learning and artificial intelligence for spectral data analysis. Prominent journals in the field span horticulture, food science, and spectroscopy, indicating the diverse applications of NIRS in fruit analysis.

Acta Horticulturae leads in publications, followed by journals focused on postharvest biology, food engineering, and near-infrared spectroscopy. This diverse range of journals indicates the broad impact of NIRS across various aspects of fruit production, quality assessment, and food science (Cen & He, 2007; Ncama et al., 2018). Institutions from multiple countries are actively involved in research, with Zhejiang University leading in publications and other Chinese institutions and universities from countries like South Africa, Spain, Italy, and Australia. The review of top papers reveals successful applications of NIRS in predicting various quality parameters across different fruits.

These studies have demonstrated the effectiveness of NIRS in measuring soluble solids content, acidity, firmness, and other quality attributes in fruits such as apples, pears, mangoes, grapes, and citrus fruits (Fan et al., 2009; Gómez et al., 2006; Li et al., 2013; Magwaza et al., 2012). The research also shows the potential of portable NIRS devices for on-site analysis, which could revolutionize quality control processes in the fruit industry (Pérez-Marín, Paz, Guerrero, Garrido-Varo, & Sánchez, 2010). Moreover, studies have explored the integration of NIRS with other technologies like hyperspectral imaging and data fusion techniques to enhance prediction accuracy and expand the range of measurable parameters (Lee et al., 2014; Nogales-Bueno et al., 2014).

Advanced analytical techniques, including chemometrics, principal component analysis, and partial least squares regression, have been widely used to interpret the complex spectral data obtained from NIRS measurements (Amuah et al., 2019; Mahanti & Chakraborty, 2020). More recently, there has been a growing trend toward applying machine learning and deep learning algorithms to improve prediction accuracy and handle the high-dimensional nature of spectral data (Ninh, Phan, Ninh, & Le Thanh, 2022; Rong, Wang, Ying, Zhang, & Zhang, 2020). This shift towards more sophisticated data analysis methods suggests a growing need for interdisciplinary expertise combining spectroscopy, agriculture, and data science. The research also highlights the importance of non-destructive measurement techniques in fruit analysis.

NIRS has been shown to be effective in assessing both internal and external quality attributes of fruits without damaging the sample, which is crucial for reducing waste and allowing repeated measurements of the same fruit over time (Arendse et al., 2018; Shah et al., 2021). This non-destructive nature of NIRS makes it particularly valuable for applications throughout the fruit supply chain, from harvest to retail. Studies have also focused on specific compounds in fruits, such as carotenoids, sugars, and polyphenols (Lamptey et al., 2023; Martínez-Valdivieso et al., 2014; Toledo-Martín et al., 2018). NIRS has been successfully applied to quantify these compounds, providing valuable information about fruit nutritional content and maturity (Lamptey et al., 2023). This application of NIRS extends beyond basic quality control to more detailed compositional analysis, which could have significant implications for nutritional studies and targeted fruit breeding programs.

The analysis also reveals a growing interest in applying NIRS for detecting and classifying fruit diseases, defects, and origins (Eisenstecken et al., 2019; Ghooshkhaneh, Golzarian, & Mollazade, 2023; Raghavendra et al., 2021). Several studies have demonstrated the potential of NIRS, often in combination with other spectroscopic techniques or imaging methods, to identify diseased fruits, detect internal defects, or classify fruits based on their geographical origin (Eisenstecken et al., 2019; Ghooshkhaneh et al., 2023; Raghavendra et al., 2021). This expanding application area could significantly impact fruit production and trade, enhancing disease management strategies and supporting geographical indication certifications.

Based on these findings, several potential research directions can be deduced: first, there is a need for further research on integrating NIRS into fruit processing lines for continuous, real-time quality assessment (Grassi & Alamprese, 2018). This could involve developing robust calibration models to handle the variability introduced by moving fruits and changing environmental conditions. Also, future research could focus on combining NIRS with other non-destructive techniques like hyperspectral imaging, thermal imaging, or nuclear magnetic resonance to provide a more comprehensive fruit analysis (de Carvalho, Pereira, de Morais, de Lima, & de Almeida Teixeira, 2019; Varith, Hyde, Baritelle, Fellman, & Sattabongkot, 2003; Zhang et al., 2020).

This multi-sensor approach could potentially overcome some limitations of individual techniques and provide a more complete picture of fruit quality. Furthermore, the trend towards machine learning and artificial intelligence in spectral data analysis will likely continue. Future research could explore more sophisticated algorithms, including deep learning architectures designed explicitly for spectral data, to improve prediction accuracy and handle increasingly complex datasets (Ninh et al., 2022; Rong et al., 2020). While much research has focused on common fruits like apples and citrus, there is potential to expand NIRS applications to a broader range of fruits, including tropical and exotic varieties.

Additionally, research could explore new quality parameters or compounds that can be measured using NIRS. Developing standardized protocols for NIRS measurements and data analysis could enhance comparability across studies and facilitate the transfer of calibration models between different instruments or locations. This could be a crucial step towards wider industrial adoption of NIRS technology.

Further research into using NIRS for early, non-destructive detection of fruit diseases or internal defects could significantly impact fruit production and storage practices (Eisenstecken et al., 2019; Ghooshkhaneh et al., 2023; Raghavendra et al., 2021). This could involve developing more sensitive instruments or refining data analysis techniques to detect subtle spectral changes associated with early disease or defect development stages. Furthermore, exploring the use of NIRS in field conditions for assessing fruit quality on the tree could support precision agriculture practices (Saranwong, Sornsrivichai, & Kawano, 2003). This might involve developing robust, portable NIRS devices capable of withstanding field conditions and providing real-time data to inform harvesting decisions.

Further research into miniaturizing NIRS devices and reducing their cost could promote widespread adoption, particularly among smaller producers or developing countries (Izneid & Al-kharazi, 2013). This could involve exploring new materials or designs for NIRS components. Research could also focus on how NIRS can be used throughout the fruit supply chain to reduce food waste. This might involve developing prediction models for shelf life or optimal ripeness, allowing for better management of fruit inventories (Rungpichayapichet et al., 2016; Shah et al., 2021).

Furthermore, as climate change affects fruit production (Bhattacharjee, Warang, Das, & Das, 2022), research could explore how NIRS can be used to monitor and predict changes in fruit quality due to varying environmental conditions. This could provide valuable data for adapting agricultural practices to changing climates. Research into developing consumer-friendly NIRS devices for assessing fruit quality at the point of purchase could empower consumers and drive improvements in fruit quality throughout the supply chain (Lee et al., 2017).

Further research is needed to develop calibration models that account for biological variability in fruits, including differences due to cultivar, growing conditions, and season (Anderson, Walsh, Subedi, & Hayes, 2020). This could involve exploring new ways to incorporate this variability into predictive models. Finally, expanding research into NIRS for detecting food fraud or authenticating premium fruit products could significantly affect food safety and international trade (Torres, Sánchez, Vega-Castellote, & Pérez-Marín, 2021).

2.11. Conclusion

Near-infrared spectroscopy (NIRS) has demonstrated significant potential in the non-destructive assessment of fruit quality, where it has shown high accuracy in predicting essential quality parameters such as total acidity, soluble solids content, and internal physiological disorders. Integrating machine learning models, particularly artificial neural networks and deep learning, has further enhanced the accuracy of NIRS in predicting fruit quality traits, enabling more efficient and precise assessments.

The development of portable NIRS devices and hyperspectral imaging have made NIRS applications more accessible for in-field use, offering real-time quality assessments that can significantly reduce postharvest losses. This study has revealed the extensive research conducted in NIRS and fruit quality, highlighting the global interest and collaborative efforts that drive advancements in this area. The future of NIRS in fruit quality assessment lies in improving real-time capabilities, enhancing device portability, and expanding its application to a broader range of fruit varieties. As the technology continues to evolve, NIRS has the potential to significantly impact global fruit production, quality control, and sustainability efforts.

Conflict of Interest Statement

The authors have declared no conflicts of interest for this article.

Data Availability Statement

Data for this article are available at Open Science Framework at https://osf.io/hwc8m/files/osfstorage

Acknowledgments

We express our profound gratitude for the assistance rendered by the CSIR-Food Research Institute, and World Food Programme.

Credit Authorship Contribution Statement:

Conceptualization: FPL, ET, FO, GSO, EEA, CLYA, SAK; Formal analysis: FPL, FO, CLYA, ET; Investigation: FPL, ET, FO, GSO, EEA, CLYA, SAK; Funding acquisition: FPL, ET, FO; Methodology: FPL, ET, FO, GSO, EEA, CLYA, SAK; Writing – original draft: FPL, ET, FO, GSO, EEA, CLYA, SAK; Data curation: FPL, FO, CLYA, EEA, ET; Funding acquisition: FPL, ET; Resources: FPL, ET, FO; Supervision: CLYA, EEA, ET; Writing – review & editing: FPL, ET, FO, GSO, EEA, CLYA, SAK.

CHAPTER THREE

APPLICATION OF HANDHELD NIR SPECTROMETER FOR SIMULTANEOUS IDENTIFICATION AND QUANTIFICATION OF QUALITY PARAMETERS IN INTACT MANGO FRUITS.

*Francis Padi Lamptey[1,2,4], Ernest Teye[1,4], Ernest Ekow Abano[1,] Charles L. Y. Amuah[3,4],

[1]Department of Agricultural Engineering, School of Agriculture, College of Agriculture and Natural Sciences, University of Cape Coast, Cape Coast, Ghana

[2]Department of Food Science and Postharvest Technology, Cape Coast Technical University, P. O. Box DL 50, Cape Coast, Ghana.

[3]Department of Physics, Laser and Fibre Optics Centre, School of Physical Sciences, College of Agriculture and Natural Sciences, University of Cape Coast, Cape Coast, Ghana

[4]University of Cape Coast, Africa Centre of Excellence for Food Fraud and Safety, AfriFoodinTegrity Centre, Cape Coast, Ghana

Correspondence: francis.lamptey@cctu.edu.gh

Published, Smart Agricultural Technology (Scopus / Sci, IF = 6.3)

Statement of contribution of Joint Authorship

Francis Padi Lamptey: (Candidate)

Conceptualised the topic, established methodology, data collection and analysis, preparation of tables and figures, writing and compilation of the original manuscript.

Teye, E.: (Principal Supervisor)

Conceptualised the topic, established methodology, supervised and edited the manuscript and co-author of manuscript.

Ernest Ekow Abano: (Co-Supervisor)

Conceptualised the topic, supervised and edited the manuscript and co-author of manuscript.

Charles L. Y Amuah: (Research Team member)

Validation, visualization and data curation and co-author of manuscript

Abstract

There are several varieties of mango fruits, and the most important quality indicators for determining mango maturity are pH and total soluble solids (TSS). The study examined the possibility of using a handheld NIR spectrometer (NIRS) with a wavelength range of 740 nm to 1070 nm and multivariate algorithms in combination with a smartphone to determine the varieties and maturity of mangoes. A total of 198 intact mango fruits were scanned with the NIR spectrometer, while a digital refractometer and pH meter were used to measure TSS and pH from the extracted mango juice. After using several preprocessing methods, multivariate classification models were created using support vector machine (SVM), linear discriminant analysis (LDA), random forest (RF), neural network (NN), and a new classifier (LDA-SVM) to identify the varieties. Partial least square regression methods, such as interval partial least square (IPLS), synergy partial least square (Si-PLS), and back interval partial least square (Bi-PLS), were used to build quantitative models for determining TSS and pH of the fruits. Among the identification techniques, the RAW, MC, SNV, FD, and SD plus LDA-SVM could identify mango fruit varieties 100% accurately in the training set and 97.44% in the prediction set. For quantification, the best model for TSS and pH measurements in mango is Si-PLS, with an r^2 value of 0.63, an RMSEP value of 1.83, an r^2 value of 0.81, and an RMSEP value of 0.49, respectively. The study demonstrated that rapid and non-destructive assessment of TSS and pH could be achieved using handheld NIR coupled with suitable chemometric tools.

Keywords: Mango, NIR spectrometer, Total soluble solids, pH, Varieties

3.1 Introduction

Mango (*Mangifera sp*.) is one of the fruits that contributes the most to global economies. It belongs to the Anacardiaceae family's Mangifera genus. The mango fruit is prized for its delicious flavor, distinctive fragrance, and health advantages (Lawson, Lycett, Mayes, Ho, & Chin, 2020). A wide range of minerals and phytochemicals are present in the fruit. The fruit contains prebiotic dietary fibre carotenoids, vitamin C, and polyphenols. The chemical composition of mango pulp varies with the location of cultivation, variety, and stage of ripeness (Tharanathan et al., 2006). The most crucial element affecting fruit quality and storage life is maturity at harvest. When ripe, immature fruits have a lower-quality flavor and are more prone to mechanical damage and shriveling. Soon after harvest, overripe fruits will likely taste bland and become mealy and mushy. Physiological issues after harvest are more likely to affect crops harvested too early or late in their season (Kader, 1997).

According to Jha et al. (2007), fruits harvested before they are fully ripe do not ripen consistently and may show severe shrinkage and poor sweetness levels. Even ethylene or acetylene treatment cannot fully develop immature fruits with the right fragrance, flavor, and taste. Besides, harvested overripe fruits have a shorter shelf life and are more prone to infection. Some chemical indicators have also determined mango ripeness, including starch, TSS, TA, phenolic compounds, carotenoids, and DM content (Jha, Kingsly, & Chopra, 2006). TSS and pH are significant inner quality indicators of fruit ripeness and postharvest quality among mango fruit internal quality characteristics. For instance, when estimating fruit maturity and harvest time and assessing and grading fruits, TSS is one of the most valuable internal quality parameters (Peng & Lu, 2008).

Currently, destructive methods are used in the standard procedure for determining the internal quality criteria of fruits like mangoes. Typically, this approach is wasteful and labor-intensive. It results in expensive analysis costs and prevents the examination of the entire crop of fruits because it necessitates specialized equipment and sophisticated methods and requires trained workers to operate them (Cortés et al., 2016). In addition, a representative sample is frequently used to forecast all the crops, which typically results in estimation errors. Therefore, a quick, non-destructive prediction of TSS in mango fruits would be beneficial in figuring out when to harvest them for the most excellent consumption quality. This would be the perfect time to meet the growing customer demand for trustworthy, high-quality fruits. Innovative technologies, notably non-destructive evaluation techniques, are quickly used to track fruit quality modifications throughout the postharvest life. These quick and non-destructive techniques can provide conclusive criteria to attain better-quality mango products and encourage eating fruits with improved health benefits. The first factor affecting some of a commodity's quality characteristics is its genetic makeup (Faniadis, Drogoudi, & Vasilakakis, 2010). Currently, mango cultivars are identified phenotypically. However, due to the impact of the environment and the dearth of distinguishing features, cultivar identification solely based on phenotypic traits is erroneous (Jha et al., 2013).

Different molecular techniques, including simple sequence repeats (Razak et al., 2020; Ajayi, Olawuyi, Ayodele, & Faneye, 2019), RAPD, and

ISSR, have been used to attempt some molecular identification of mango cultivars (Tu, 2019). However, these methods are destructive, and the tested fruit cannot be sold. Analyzing mango quality factors can alternatively be done using near-infrared spectroscopy. This technique offers quick and non-destructive food safety and quality detection. It has been adopted for quantitative and qualitative testing in the food industry. Near-infrared spectroscopy is a helpful tool that is simple, fast, non-destructive, and requires little to no sample preparation. Some researchers have employed NIRS to measure some quality parameters in pineapples, cocoa, and strawberries (Amuah et al., 2019; Anyidoho, Teye, & Agbemafle, 2021; Anyidoho, Teye, Agbemafle, Amuah, & Boadu, 2021; Mancini et al., 2020). Others employed NIRS to determine the quality of mango fruits like total soluble solids (Schmilovitch, Mizrach, Hoffman, Egozi, & Fuchs, 2000), vitamin C and TA (Munawar et al., 2019), and β-carotene (Rungpichayapichet et al., 2015). At the time of this research and to the best of our knowledge, no research has addressed the feasibility of using a handheld NIRS combined with chemometric methods to discriminate between different mango cultivars and predict TSS and pH without causing damage.

3.2. Materials and Methods

3.2.1. Mango Fruit Samples

One hundred and ninety-eight (198) mango fruits were obtained from Ministry of Food and Agriculture registered farmers, Somanya, Eastern Region, Ghana. These fruits were harvested at different ripening stages and conveyed to the School of Agriculture Teaching and Research Laboratory, University of Cape Coast (UCC). These fruits include 34 Keitt mango fruits, 66 Haden fruits, and 98 indigenous fruits (local). Before measurements were performed, these fruits were stored for two days at 26°C (1°C).

3.2.2. NIR Spectroscopy Measurement

Each mango's spectrum was collected in the reflectance mode utilizing a portable NIRS (SCIO^{TM}), having a spectral range of 740 nm to 1070 nm and a resolution of 1 nm. The centre of all the fruits was scanned thrice after being rotated 120 degrees. Fruits were scanned at a temperature of 26 °C and a relative humidity of 60%. Figure 3.1 shows the setup of the scanning processing using a SCIO NIR spectrometer with the three varieties of mango fruits.

3.2.3. Reference Method (TSS/°Brix and pH)

A digital refractometer (model: PAL-1, °Brix range of 0-35%; Atago, Tokyo, Japan) was used to calculate the total soluble solids (TSS) contents following the protocols described by others (Abarra, Serrano, Sabularse, Mendoza, & Rosario, 2018). The results of three replicate readings were represented as degrees Brix. The juice from the fruits was dropped directly into the refractometer for the TSS measurement, and the refractometer was calibrated and cleaned after each measurement using distilled water.

pH of the scanned mangoes was measured destructively with a digital pH meter, and the average values were noted in triplicate.

3.3. Statistical Analysis

3.3.1. Data Partition

The raw dataset (from the 198 samples) was divided into two subsets after the proper preprocessing using Kennard stone algorithm: the calibration set (data from 139 samples) for model creation and the prediction set (data from 59 samples) to evaluate the predictive capacity of the model built. To prevent bias, the calibration set was selected from 70% of the data, and the prediction set was selected from the remaining data of all types.

3.4 Software Device

Utilizing a research license obtained from the SCIO lab, recordings of spectral data were downloaded, along with the associated reference value recorded at the time of scanning, and imported into MATLAB version 9.5.0 (Mathworks Inc., USA) using Windows 10 Basic for data.

3.5 Spectral Data Preprocessing

Five preprocessing strategies, comprising mean centering (MC), multiplicative scatter correlation (MSC), standard normal variant (SNV), first derivative (FD), and second derivative (SD), were utilized in comparison to achieving the best performance out of the model. To create precise, dependable, and constant calibration models, noise from the raw spectra and other background data were removed during preprocessing. As shown in Figure 3.2(a), the raw near-infrared spectra of the mango fruits provide useful and unwanted information. This can be due to interferences from light scattering from the materials, inconsistent spectra, temperature changes, or background sounds (Jha & Garg, 2010). As a result, it was decided to utilize chemometric pretreatment of the dataset to maintain the similarities and differences between the preliminary observations while only acquiring the useful features of the samples.

To achieve this, five spectrum preprocessing techniques—MC (mean centering), MSC (multiplicative scatter correction), SNV (standard normal variant), FD (first derivative), and SD (second derivative)—were utilized in

MATLAB version 9.6.0, as illustrated in Figures 3.2(b)-3.2(f). Mean centering is a method of spectral preprocessing that involves calculating the dataset's mean spectrum and subtracting the mean from each spectrum. The first derivative preprocessing method reduces baseline effects by comparing the spectra of two subsequent measurement locations. The second derivative transformation algorithm separates overlapped peaks and improves resolution, removing the additive and multiplicative baseline within the spectra. The Savitzky-Golay technique was utilized to smooth the NIR spectra before applying the SD preprocessing procedure. Generally speaking, the linearity and corrected offset in NIR data were most enhanced using the Savitzky-Golay smoothing SD technique.

3.6. Principal Component Analysis (PCA)

After the initial preprocessing treatment of the spectra data, this work employed principal component analysis (PCA) as an unsupervised pattern recognition approach to display data trends in a dimensional space as score plots. A common technique for lowering the size of a data matrix is PCA, which divides the data into basic components with understandable variables. The most valuable data in PCA is found in the top three principal components (PCs), which frequently highlight pertinent information while minimizing or eliminating redundant data. Typically, in descending sequence, PC1, PC2, PC3, PC4, PC5, etc., explain and provide pertinent information (Anyidoho et al., 2021).

3.7. Identification Model

Different multivariate characterization and quantification procedures were systematically explored after carefully choosing the optimum spectrum pretreatment method, and choosing the appropriate type is crucial in subsequent analyses. Support vector machine (SVM), linear discriminant analysis (LDA), random forest (RF), neural network (NN), and a new classifier, LDA-SVM, were the multivariate systems used in this experiment (Alolfe, Mohamed, Youssef, Mohamed, & Kadah, 2009; Xiong & Cherkassky, 2005), for the identification of the problem. The outcomes of the LDA-SVM classifier were contrasted with those of the SVM, LDA, RF, and NN. With customized kernel functions, SVM learning methods can imitate complicated non-linear boundaries while exhibiting good performance when it comes to generalization. SVM has recently been instituted in chemometrics and has proven effective in classifying near-infrared spectra. SVM maximizes the inter-class geometric margin while minimizing the empirical classification error (Devos, Ruckebusch, Durand, Duponchel, & Huvenne, 2009). The typical approach for reducing the spectral dimension is linear discriminant analysis. It uses the samples' prior knowledge and experience during dimensionality reduction. Projecting spectral data from more excellent measurements to lesser dimensions while maximizing space within classes and minimizing space within classes is the main goal of LDA (Qi, Wu, Yang, Wu, & Fu, 2022).

Random Forest (RF) has drawn more interest in vis-NIR spectral studies in several fields. It has several benefits, including resilience to disturbances, the capacity to be employed regardless of whether the predictor variables exceed the data, facing minor overfitting, and evaluating variable relevance. When introducing variability to the general spectrum library, random forest can manage non-linear and hierarchical behaviours to anticipate local samples (Nawar & Mouazen, 2019). The NN is one of the machine learning (ML) methods. The neural network is the most commonly utilized model because of the vast dataset that is now available, robust computing capabilities, and advanced algorithm architecture. NN is a type of supervised ML, and it has been established that a network with only a hidden layer but sufficient neurons can express an arbitrary function. Neural networks have strong evaluation abilities for denoting complex, non-linear relations between input and output features. The input layer, hidden layer, and output layer are the three layers that make up the architecture of a NN (Qi, Chen, Li, Cheng, & Li, 2019).

3.7.1 Quantification Model

In the process of building the models for TSS and pH of mangoes, full spectrum partial least square regression method and different types of partial least square regressions, interval partial least regression (IPLS), synergy partial square regression (Si-PLS), and back interval partial least square regression (Bi-PLS) were utilized to create subsequent models amongst the spectra fingerprinted data and wet chemistry data. The performance of the multivariate models was estimated employing the coefficient of determination (r^2, Eq.(1)), the root mean square error of calibration (RMSEC, Eq.(2)), the root error of prediction (RMSEP, Eq.(3)), bias (bs, Eq.(4)), and the residual predictive deviation (RPD, Eq. (5)). As a good rule of judgment; a suitable robust model should have a high RPD, a high coefficient of determination value, a low root mean square error, and bias values (Liu, He, Wang, & Sun, 2011). MATLAB was used to compute all the statistical analyses. These parameters were calculated using the equation other researchers used

(Arendse, Fawole, Magwaza, Nieuwoudt, & Opara, 2017; Liu, Sun, & Ouyang, 2010; Teye et al., 2023; Williams, 2014).

$$r^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (1)$$

$$RMSEC = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \quad (2)$$

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \quad (3)$$

$$\text{Bias} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \quad (4)$$

$$\text{RPD} = 1/\,(1 - r^2)^{1/2} \quad (5)$$

where n = the number of samples

y_i = the reference measurement results for sample i

$\hat{y}_i$ = the estimated results of the model for the sample i

and $\bar{y}$ = the mean of the reference measurement results for all samples in the data set.

3. 8 Results

3.8.1 Reference Measurement and Data Partitioning

The data set of 198 sample spectra was split into two subsets: the calibration set (139, used for building the model) and the prediction set (59, used for testing the model's applicability). All 198 samples were randomly picked for calibration and prediction sets to prevent selection bias. The calibration set was made up of three samples out of every five, while the prediction set was made up of the remaining samples. Table 3.1 shows the mean and standard deviation of the mango varieties' pH and TSS parameters.

3.8.2. Spectra Profile

Attractive traits can be seen in the mango fruits' scanned spectrum profiles. As seen in Figure 3.2, the raw spectra (Figure 3.2a) indicated some specific information with little observed peaks and some noise; as a result, the raw spectra were preprocessed using several techniques. Figures 3.2b, 3.2c, and 3d indicate several peaks in the MC, MSC, and SNV around 975. While SD exhibited several peaks at 820, 840, 920, 940, 980, 1010, 1020, and 1040, FD showed peaks around 825 and 950. This is because spectral derivation is a great mathematical technique for removing baseline drafts and improving and sharpening spectral features (Hong et al., 2019). To eliminate the vertical offsets and linearly sloping baselines, the first and second derivatives are frequently computed and smoothed using SG smoothing (Jiao, Li, Chen, & Fei, 2020).

The wavelength range corresponds to the H_2O, ROH, ArOH (OH bond on the aromatic group), and NH_2 functional group (N-H 3rd overtone, O-H 2nd overtone, and N-H 2nd overtone) (Cen & He, 2007; Stuart, 2004). The main components of mangoes, including their water, glucose, sucrose, and cellulose, are associated with these groups. Total soluble solid is an organic molecule containing C-H, O-H, C-O, and C-C bonds that might be non-destructively measured using NIR spectroscopy (Amuah et al., 2019). The mean spectrum of the scanned samples was determined to make sure the samples differed from one another, as seen in Figure 3.3.

3.8.3. Principal Component Analysis

The spectra cluster trends were found using PCA. According to the results, PCA performed on raw, MSC, and SNV preprocessed spectra data did not provide any clear trends or separations, as shown in Figures 3.4 (a), (c), and (d). However, as seen in Figure 3.4(f), SD-PCA provided a separation with a distinct cluster trend. From the 198 samples, the top three PCs were PC1 (66.56%), PC2 (26.44%), and PC3 (1.40%). It demonstrates that the top three PCs can account for 94.40% of the variance data from the spectra dataset, which includes the pertinent biochemical data in the samples. Mango fruits with the same traits are grouped closely by employing the PCA approach, which extracts relevant information and eliminates unnecessary ones. As a result, the graphical output may be utilized to identify differences between the various types of mango fruits that were used. Three varieties of mango fruits were employed in the study, as shown in Figure 3.4(f).

The groupings comprise a broader array of mango fruits. The graphical plot provides pertinent details that might be utilized to distinguish between Keitt, Haden, and local varieties of mangoes. Although PCA is not a classification method, it could visualize dimension space and reveal data trends (Anyidoho et al., 2021). Figure 3.5 illustrates the PC loadings with three principal components, which explain the cluster trend observed in the PCA score plot. The major peaks revealed were around 742nm, 745nm, 788nm, 812nm, 835nm, 853nm, 878nm, 899nm, 933nm, 959nm, 982nm, 1005nm, 1044nm, and 1068nm, for PC1, 744nm, 747nm, 787nm, 813nm, 824nm, 843nm, 859nm, 875nm, 926nm, 945nm, 974nm, 1004nm, 1033nm, and 1062nm for PC 2, and 745nm, 748nm, 752nm, 767nm, 781nm, 796nm, 815nm, 834nm, 848nm, 864nm, 884nm, 898nm, 924nm, 957nm, 980nm, 999nm, 1034nm, and 1068nm for PC3. These wavelengths are associated with the third overtones of C-H, N-H, O-H and the 3rd overtones of O–H, N–H, C–H (Ozaki, Christy, & McClure, 2006).

3.8.4. Performance of the Identification Models.

The PCA was unable to effectively categorize mango fruits according to their varieties. Thus, a qualitative analysis was carried out using SVM, LDA, NN, RF, and LDA-SVM. The outcomes of several classification models for differentiating between mango varieties are shown in Table 3.2. There are advantages and disadvantages to every multivariate classification algorithm. Table 3.2 demonstrates how, except for the LDA model, the SD processing significantly improved the performance of all multivariate classification methods in the calibration set and prediction set. Other researchers obtained similar results and found that raw preprocessing performed better than MSC and SNV (Li, Peng, Li, Yang, & Chao, 2020).

In the calibration set, the SVM model's best classification rate was 91.25%, and in the prediction set, it was 92.50% at the optimum number of 3 PCs. The optimal classification rate for NN was 87.50% for the calibration set and 77.50% for the prediction set. The best classification rate for RF was 90.63% for the calibration set and 92.50% for the prediction set. However, for LDA, the best classification rate for the calibration set was 81.25% and 85.00% for the prediction set. From the study results, the LDA-SVM model was superior to the other models. In the training set, the LDA-SVM model attained a classification accuracy of 100%, and the training set attained a classification accuracy of 97.44% for samples preprocessed with MC, SNV, FD, and SD.

3.8.5. Optical Variable Selection

Effective variable selection is utilized to boost a model's performance to the highest possible level. The ideal range for pH and total soluble solids (TSS) using SiPLS is shown in Table 3.3. The intervals chosen for pH were [2,7,10,17], and for TSS, were [7,13,16] with RMSECV = 1.85355 and 0.52015, respectively. As shown in Figure 3.6, the wavelengths chosen for TSS were 830–844, 910–993, and 959–972 nm in the complete spectrum employed (740–1050 nm). These wavelengths are associated with the third overtones of C-H and N-H and the second overtones of N-H and O-H. To determine the pH, the wavelengths selected 758–775 nm, 848–865 nm, 901–917 nm, and 1020–1036 nm correlate to the 3rd overtones of O–H, N–H, and C–H, and the 2nd overtones of N–H and O–H, which are related to acidity in the mango sample (Ozaki et al., 2006).

3.8.6. Comparative Performance of TSS and pH Models

The effective variable selection performance comparison models for pH and TSS are displayed in Table 3.4. From the table, it was seen that various PLS algorithms have distinctive strengths. The performance of full PLS and IPLs denoted poor results comparatively. These were below 0.66 and 0.79 for TSS and pH, respectively. This could be explained by the fact that the whole spectrum comprises some spectral variables that are unrelated and collinear. This could influence its ability to predict and stabilize the derived model. Even though IPLS chose the broadest spectrum that matches TSS and pH, fixing the full PLS's flaw, it only utilized intervals, ignoring other relevant spectral information.

To overcome this, other variable selection techniques were used; Si-PLS and Bi-PLS algorithms were tried to give the optimum model for determining TSS and pH. Si-PLS for TSS and pH outperformed the other models with r^2= 0.63 and 0.81 under the first derivative preprocessing technique for TSS and the second derivative preprocessing technique for pH, respectively. For TSS, the RPD ranges from 1.25 - 1.78, whereas for pH, the RPD ranges from 1.54 - 3.35. An RPD value of 2-2.5 suggests that the model can make coarse quantitative predictions, while values of 2.5–3 or higher indicate that the model can make good and excellent predictive accuracies. An RPD of 1.5–2 indicates that the model can distinguish between low and high response variables (Amodio, Ceglie, Chaudhry, Piazzolla, & Colelli, 2017). The bias and residual prediction deviation (RPD) for the best model for each predicted parameter are given in Table 3.4.

3.9 Discussion

This study shows significant potential for the straightforward usage of portable NIR technology by mango farmers to predict the TSS and pH of mangoes in domestic and international markets. Figures 3.7a and b present the correlation between the reference recommended method values and the NIR spectroscopic estimated values for TSS and pH in mango fruits, respectively. It must be noted that the final TSS of mangoes ranges from 14 to 20% (Subedi & Walsh, 2011), at which the fruit is ripe, sweet, and has a good flavor. The TSS level in mangoes is also essential for determining the best time for harvesting. Harvesting at the right time ensures the fruit has the right sugar content, flavor, and texture. Mangoes harvested too early may have low total soluble solids, while those harvested too late may have high total soluble solids and may be overripe (Halepotara, Kanzaria, Rajatiya, Solanki, & Dodiya, 2019). Mangoes are considered acidic fruits, with a pH range of 3.8 to 4.5 (Bekele, Satheesh, & Sadik, 2020).

The pH level of mangoes can be affected by factors such as variety, maturity stage, and environmental conditions. A low pH level in mangoes is desirable as it helps to enhance their flavor and can also act as a natural preservative, preventing the growth of spoilage microorganisms, mainly bacteria (Saranraj & Geetha, 2012). The results demonstrated that NIR spectroscopy has the potential to forecast mango ripeness and quality parameters without causing damage to the fruit. The study's findings would enable the growers to maximize their earnings by only harvesting mangoes at ideal maturity. Above all, this might be used as the foundation for commodity pricing, quality assurance, and automatic sorting systems.

The second derivative treatment outperformed the other preprocessing methods in the comparative analysis of the PCA cluster by clearly displaying a cluster trend, as seen in Figure 3.4f. The biochemical makeup of each mango fruit, which varies depending on the variety, can be used to explain the clustering. The contributions of the three topmost PCs were 94.40% of the total variation in the initial data. However, PCA is not a classification method; hence, it cannot provide precise identification. However, lowering dimensionality may conserve a lot of variation in a high-dimensional space.

Five pattern recognition algorithms that can solve identifying issues were used in this study. Pattern recognition techniques such as SVM, LDA, RF, NN, and LDA-SVM were used to create a classification model. The LDA-SVM model attained a classification accuracy of 100% in the training set. The training set attained a classification accuracy of 97.44% for samples preprocessed with MC, SNV, FD, and SD. In contrast, the classification accuracies for the calibration set and prediction set were 91.25% and 92.50% for SNV (SD), 90.63% and 92.50% for RF (SD), and 87.50% and 77.50% for NN (SD). These results demonstrated that LDA-SVM gives the best results compared to SVM, LDA, RF, and NN since it combines attractive properties of both SVM and LDA approaches (Alolfe et al., 2009; Xiong & Cherkassky, 2005).

When a handheld NIRS was connected to a mobile phone to select wavelength variables step-by-step effectively, Si-PLS gave the preeminent procedure for determining total soluble solids (TSS) and pH in mango fruit non-destructively. The association between the TSS and pH values predicted by NIR spectroscopy and the reference method's recommended values are displayed respectively in Figures—3.7a and 3.7b. The reference measurement results accounted for over 70% of the near-infrared sensor data variance, indicating that correlation values above 0.71 might be suitable for screening and other approximate calibrations (Polinar et al., 2019; Williams, 2001). These results compared well with those of other researchers, where the r^2 for TSS and pH were found to range from 0.66 to 0.88 (Agulheiro-Santos, Ricardo-Rodrigues, Laranjo, Melgão, & Velázquez, 2022; Arendse et al., 2017; Theanjumpol et al., 2012).

However, compared to the range used (740-1070 nm), the range employed by these authors was wider (1064-1640 nm). This further supports the idea that handheld NIR might be used to assess the TSS and pH of mango fruit. However, other researchers also discovered a better coefficient of determination (r^2) between 0.93 and 0.98 for TSS and pH (Jiang et al., 2012; Ncama et al., 2017; Teye et al., 2015). Despite these results being relatively superior, the author's use of an extra costly and greater NIRS restricts their use for on-site discovery at the farm gate for instantaneous analysis. More specifically, the model's effectiveness in this research might be credited with the strength of Si-PLS. Si-PLS left out additional redundant data that can affect the model's functioning and chose several relevant wavelengths associated with the quality parameters of interest (TSS & pH). Following optimisation by the RMSEC in the SiPLS model, all feasible groupings of 2,3 or 4 intervals were considered, and the number of intervals that yielded the groupings of intervals exhibiting the least RMSEC was selected (Guo et al., 2011; Teye, 2022). Additionally, SiPLS provides the ability to look for the ideal interval groupings to generate the optimal model for reliable calibration. The synergy partial least square regression method divided the whole spectrum into a certain number of intervals (variable-wise). It calculated every possible PLS model pairing of more intervals (2, 3, and 4) to produce the best performance for this research (Mantanus et al., 2009; Teye, 2022). In this investigation, the ideal TSS and pH subintervals were 23 and 19, respectively.

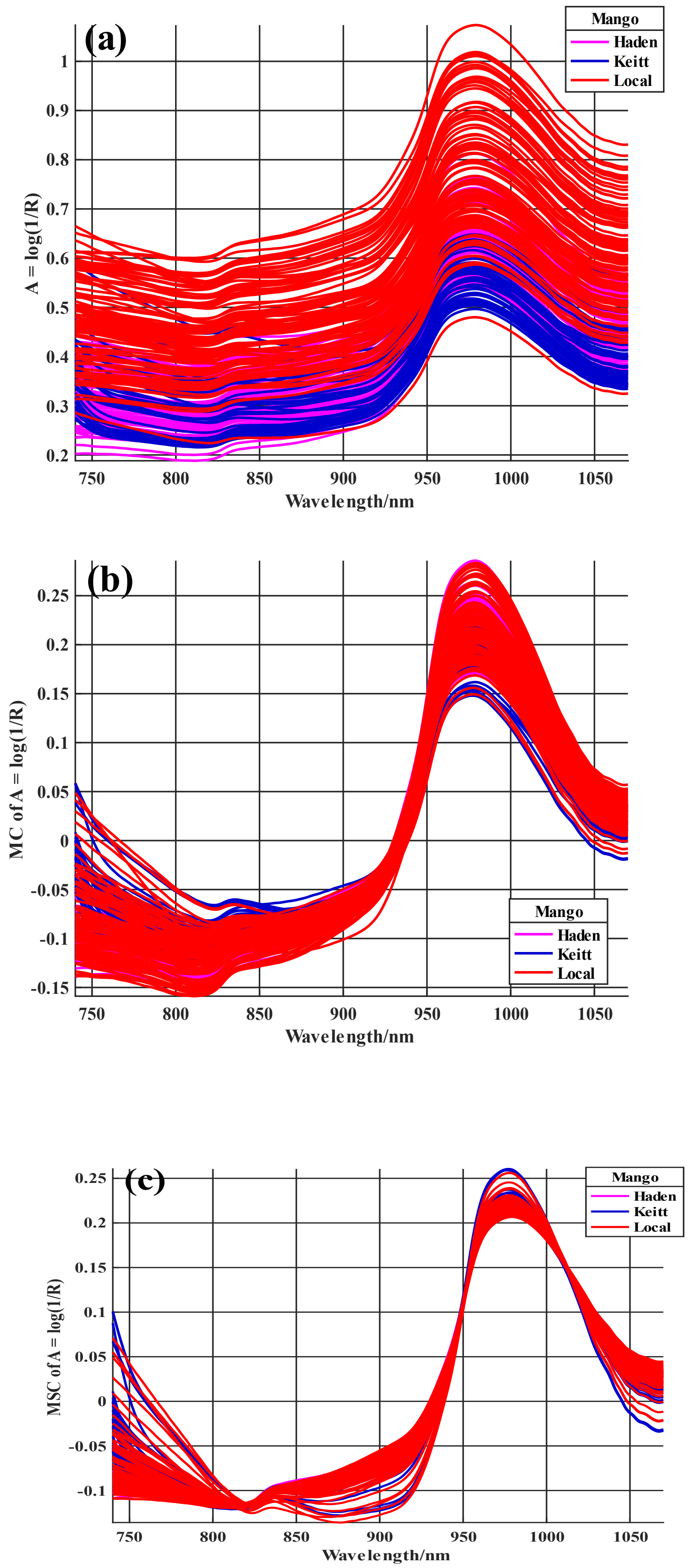
Additionally, Si-PLS was chosen above back interval partial least square and other variable selection algorithms. However, the back interval partial least square divides the data into a predetermined number of intervals, and PLS models are generated with each interval left out (Xiaobo, Jiewen, Povey, Holmes, & Hanpin, 2010). The interval that has been left out may impact the model's efficiency in this work. This study further establishes that synergy interval selection might also be incorporated into different models, as done by Ouyang, Chen, Zhao, and Lin (2013).

Table 3.1: Reference Measurements of TSS and pH

Subset	Variety	Number of samples	Min	Max	Mean	Std
TSS	Keitt	34	4	14	9.13	1.90
	Haden	66	7	13.5	11.15	1.25
	Local	98	10	18	14.48	1.57
pH	Keitt	34	3.98	5.86	4.82	0.55
	Haden	66	3.72	4.93	4.26	0.29
	Local	98	3.72	6.81	5.65	0.76

Note: TSS (Total soluble solids), Max (Maximum), Min (Minimum), and Std (Standard deviation)

Figure 3.1. General set up for Scanning and the Three Mango Varieties.


(a)
Mango
Haden
Keitt
Local
A = log(1/R)
Wavelength/nm
(b)
MC of A = log(1/R)
Wavelength/nm
(c)
MSC of A = log(1/R)
Wavelength/nm

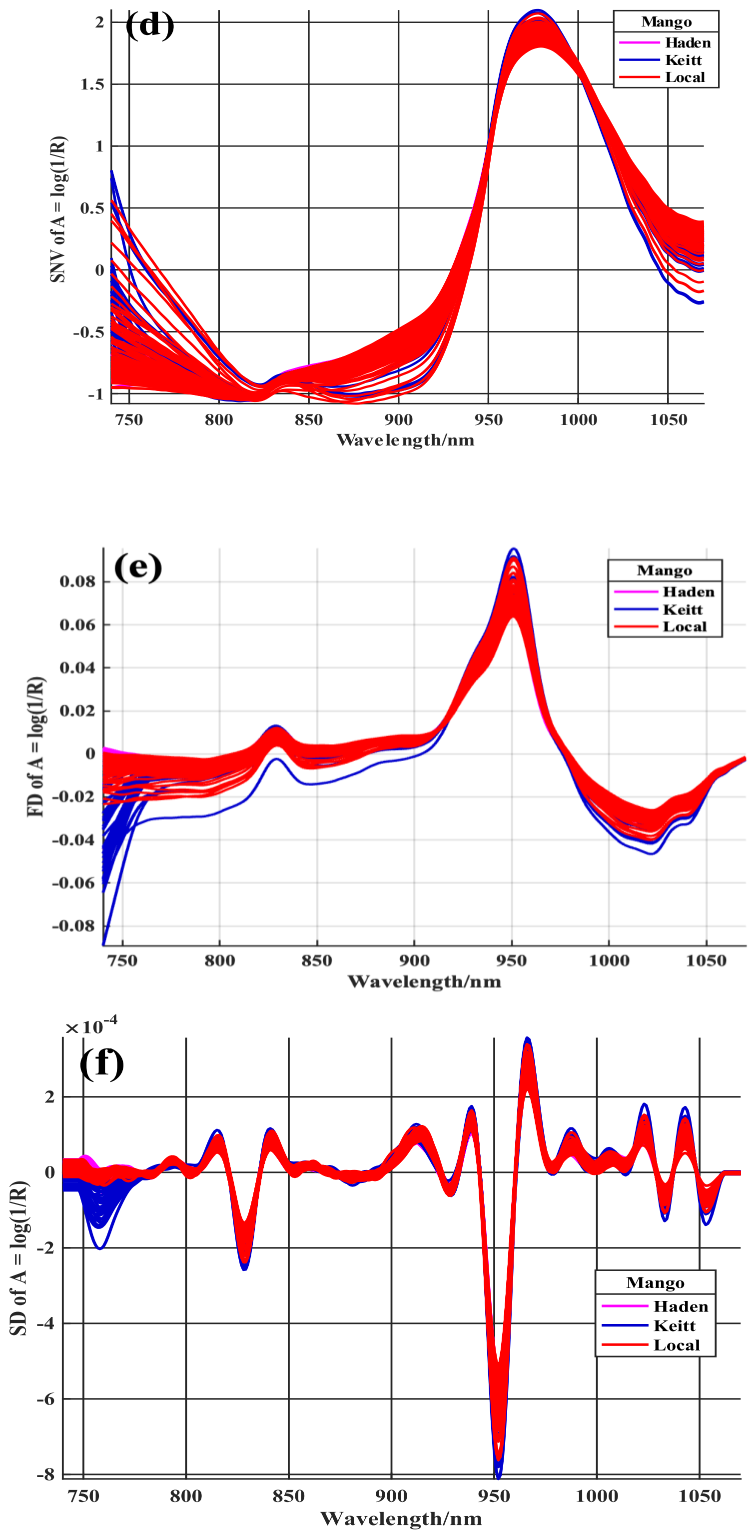


Figure 3.2. Spectra of RAW (a), MC (b), MSC (c), SNV (d), FD (e) and SD (f)

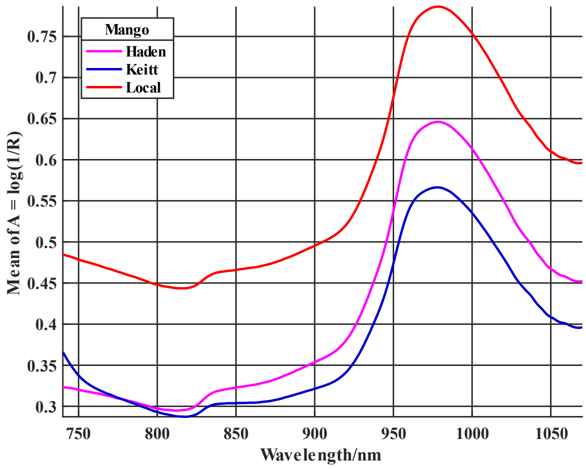
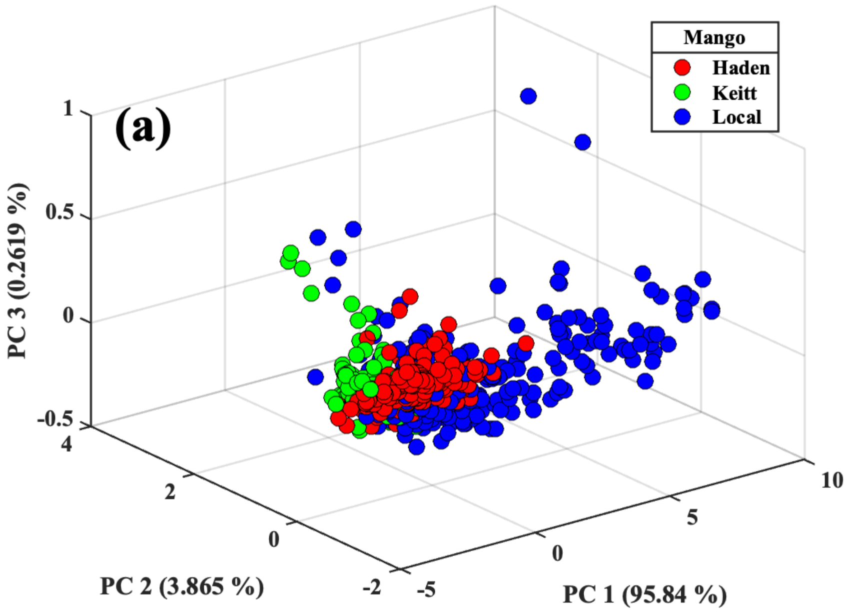
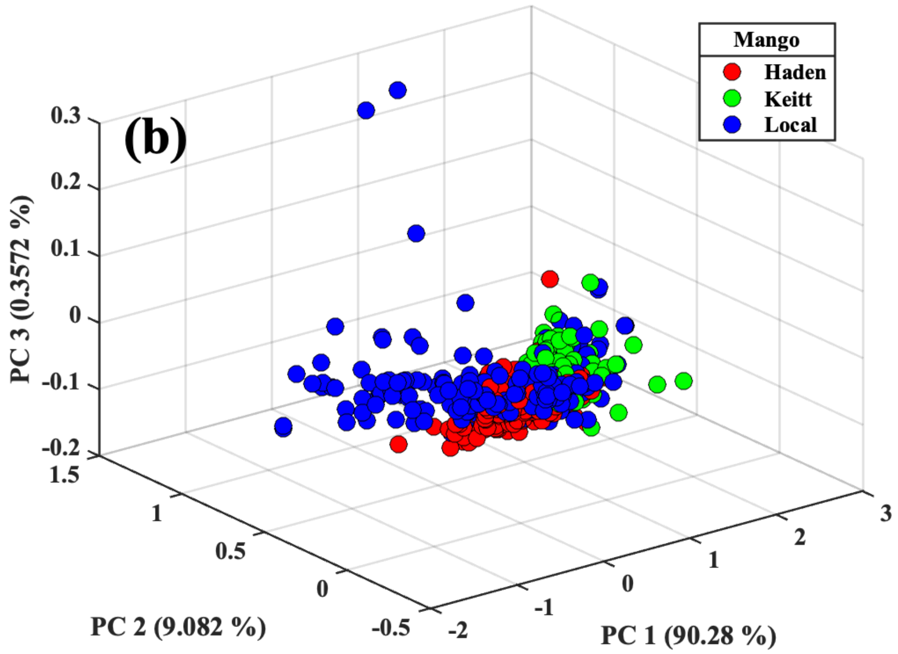
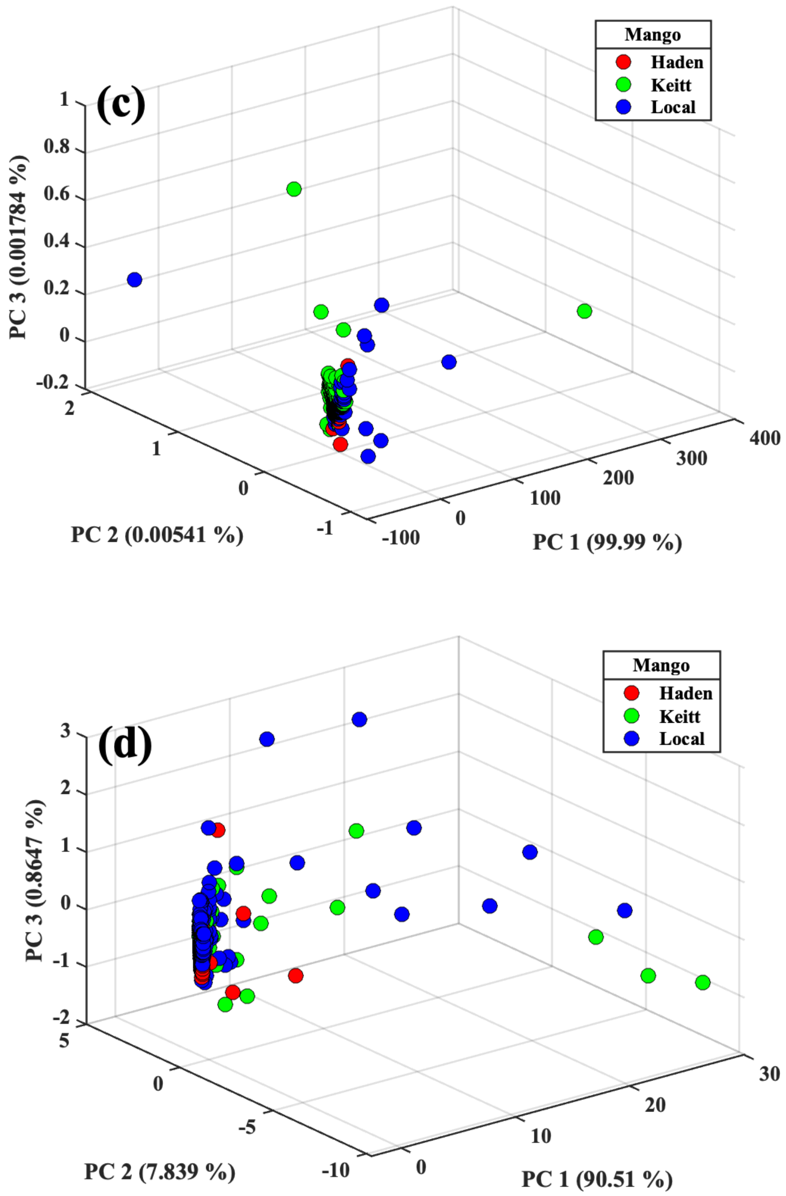


Figure 3.3. Mean spectra profile of Haden, Keitt and Local fruits.


(a)
Mango
Haden
Keitt
Local
PC 3 (0.2619 %)
PC 2 (3.865 %)
PC 1 (95.84 %)


(b)
Mango
Haden
Keitt
Local
PC 3 (0.3572 %)
PC 2 (9.082 %)
PC 1 (90.28 %)


(c)
Mango
Haden
Keitt
Local
PC 3 (0.001784 %)
PC 2 (0.00541 %)
PC 1 (99.99 %)
(d)
Mango
Haden
Keitt
Local
PC 3 (0.8647 %)
PC 2 (7.839 %)
PC 1 (90.51 %)

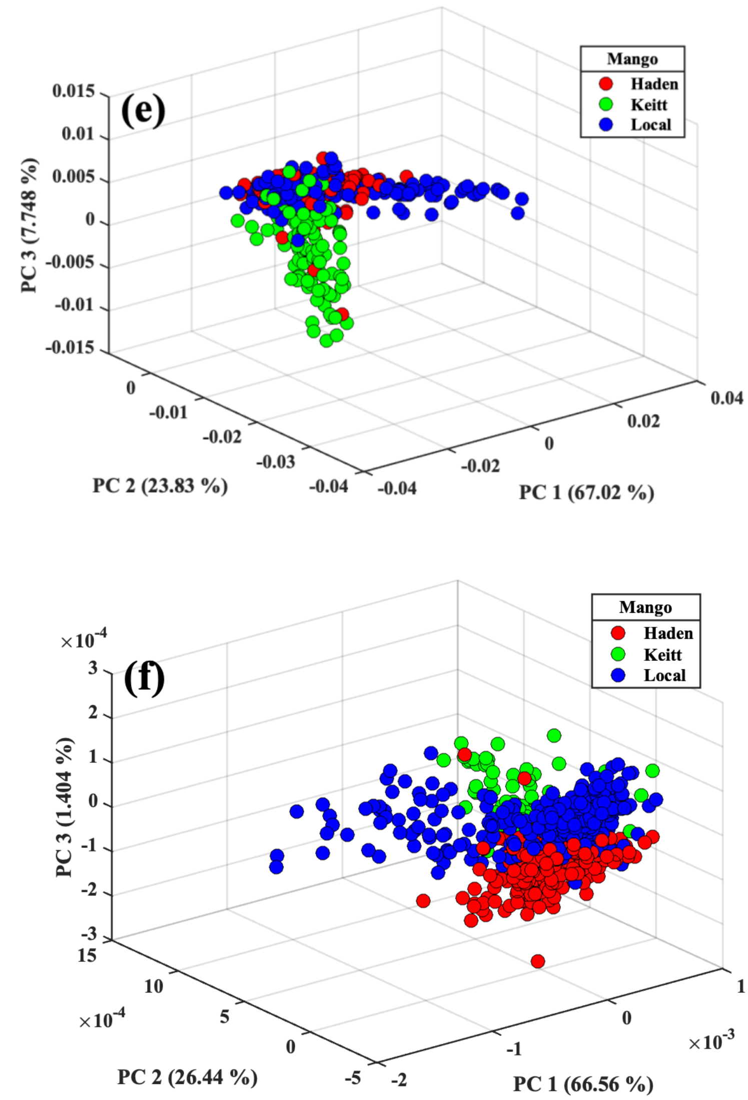


Figure 3.4. PCA Score Plot of Spectra Data. RAW (a), Preprocessed with MC (b), Preprocessed with MSC (c), Preprocessed with SNV (d), Preprocessed with FD (e), and Preprocessed with SD (f), SD with PCA Technique Gave a Separation with Clear Cluster Trend.

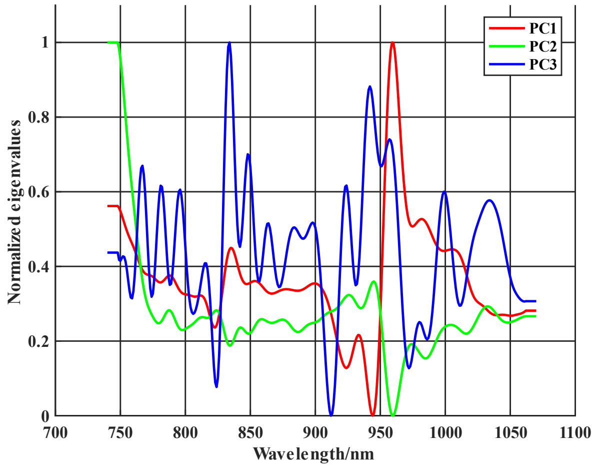


Figure 3.5. PC Loadings with Three Principal Components

Table 3.2: Performance of Identification Model for Mango Varieties at 3PCs

Classification Model	Preprocessing treatment at 3PCs											
	RAW		MC		MSC		SNV		FD		SD	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
SVM	84.38	87.50	70.00	75.00	79.38	67.50	78.13	67.50	73.75	67.50	91.25	92.50
LDA	81.25	85.00	70.00	65.00	75.00	82.50	75.63	80.00	65.63	65.00	65.63	65.00
NN	79.38	72.50	66.25	67.50	68.13	75.00	71.25	77.50	67.50	77.50	87.50	77.50
RF	80.00	82.50	70.00	50.00	77.50	65.00	76.25	72.50	71.88	77.50	90.63	92.50
LDA-SVM	100.00	97.44	100.00	97.44	65.41	53.85	100.00	97.44	100.00	97.44	100.00	97.44

Note: SVM (support vector machine), LDA (linear discriminant analysis, NN (neural network), RF (random forest), MC (mean centering), MSC (multiplicative scatter correction), SNV (standard normal variate), FD (first derivative), and SD (second derivative).

Table 3.3: Performance of SiPLS Model with some Selected Optimal Spectra Regions

Items	No. of subintervals	PLS factor	Selected subinterval	RMSEC
TSS	10	8	[1,24]	1.94421
	11	7	[4,7,8]	1.92433
	12	7	[4,6,8]	1.92644
	13	8	[1,2,5]	1.91618
	14	11	[5,9,10]	1.91304
	15	12	[6,9,10]	1.88604
	16	6	[5,6,9]	1.91459
	17	6	[5,6,8]	1.88996
	18	6	[6,7,10]	1.89257
	19	9	[6,11,13]	1.87829
	20	12	[7,9,13]	1.85864
	21	14	[1,2,3]	1.8871
	22	9	[7,13,15]	1.882
	23	10	[7,13,16]	1.85355
	24	14	[1,2,8]	1.89207
	25	12	[8,15,16]	1.86698
pH	10	3	[1,4,6,8]	0.53529
	11	4	[1,5,6,10]	0.53572
	12	3	[2,5,9,12]	0.54233
	13	2	[5,7,9,13]	0.53746
	14	5	[2,6,10,14]	0.54301
	15	4	[2,6,11,15]	0.53856
	16	5	[2,9,11,16]	0.52945
	17	4	[2,7,9,15]	0.52199
	18	4	[2,7,10,16]	0.53489
	19	3	[2,7,10,17]	0.52015
	20	5	[2,8,10,18]	0.53516
	21	4	[2,8,11,19]	0.52475
	22	4	[2,9,12,20]	0.52506
	23	4	[2,9,12,21]	0.52849
	24	3	[2,3,9,13]	0.5359
	25	10	[2,8,17,21]	0.53257

Note: PLS (Partial least squares), RMSEC (root mean square error of calibration), and TSS (Total soluble solids).

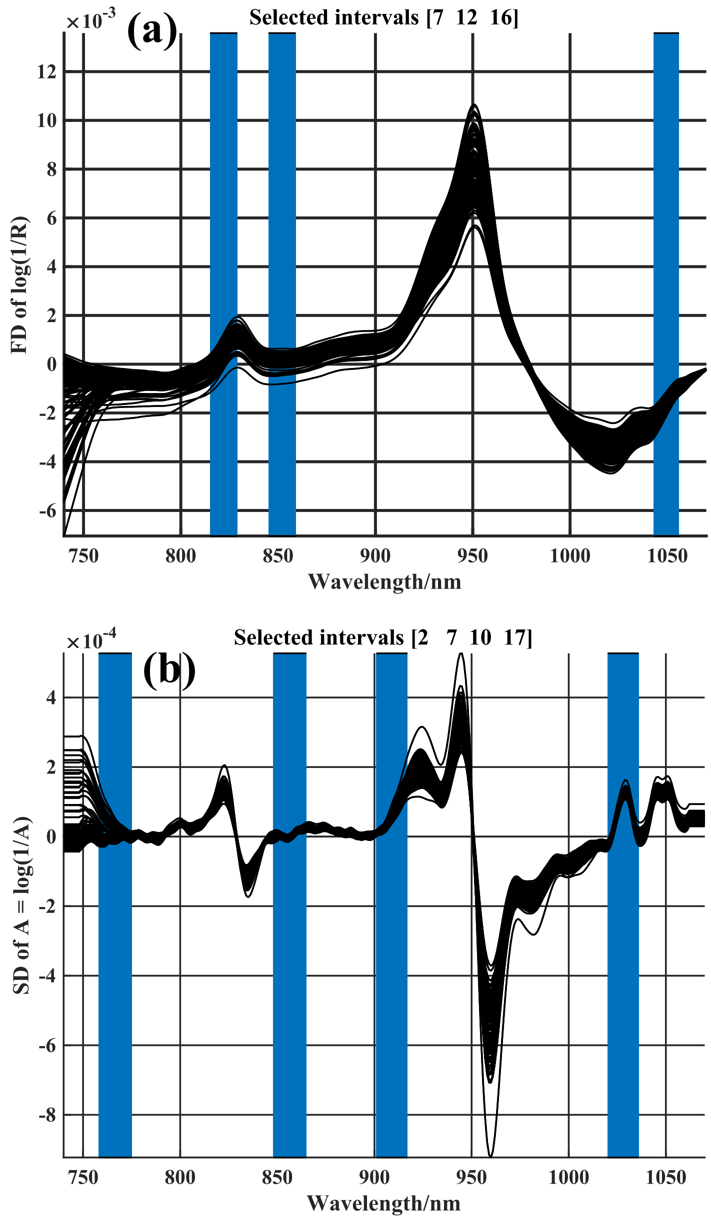


Figure 3.6. Optimal Selection and Combination of Spectra Region by SiPLS.

Table 3.4: Efficient Variable Selection Performance Comparative Models for TSS and pH

	Model	Preprocessing	r^2	RMSEC	Bias	r^2	RMSEP	Bias	RPD
TSS	PLS	RAW	0.5495	2.1828	0.1851	0.586	1.9915	0.2593	1.5542
		MC	0.5615	2.1452	0.182	0.5557	2.0507	0.267	1.5002
		MSC	0.5217	2.1829	0.1851	0.5363	2.1758	0.2833	1.4685
		SNV	0.5083	2.271	0.1926	0.4917	2.0441	0.2661	1.4026
		FD	0.6029	2.1218	0.18	0.6589	1.7296	0.2252	1.7122
		SD	0.6569	1.9165	0.1626	0.6843	1.8932	0.2465	1.7798
	IPLS	RAW	0.5918	2.1122	-0.0466	0.5664	2.0415	0.0588	1.5186
		MC	0.3908	2.3976	0.0152	0.3603	2.3526	-0.4715	1.2503
		MSC	0.4416	2.2986	-0.0325	0.5005	2.3296	-0.0315	1.4149
		SNV	0.573	2.1677	-0.0089	0.5727	1.9537	-0.3363	1.5298
		FD	0.5919	2.1443	-0.0418	0.6003	1.94	-0.6165	1.5817
		SD	0.5735	2.0857	-0.0587	0.5673	2.1567	0.2413	1.5202
	Si-PLS	RAW	0.6414	2.0082	-0.0102	0.6569	1.864	0.0468	1.7072
		MC	0.7014	1.8535	-0.0313	0.4782	2.3383	-0.7163	1.3844
		MSC	0.6856	1.8755	-0.0661	0.4342	2.3469	-0.1625	1.3294
		SNV	0.6687	1.9624	0.0219	0.5842	1.9811	-0.5417	1.5508
		FD	0.6389	2.049	0.0243	0.6682	1.8315	-0.6523	1.7360
		SD	0.6652	1.8981	0.0066	0.6006	2.0851	0.166	1.5823
	Bi-PLS	RAW	0.6176	2.0598	0.0078	0.6703	1.8454	0.148	1.7416
		MC	0.652	1.9762	0.0159	0.4893	2.2862	-0.6483	1.3993
		MSC	0.6376	1.9755	0.0001	0.5384	2.1873	-0.224	1.4719
		SNV	0.5748	2.1715	-0.0182	0.5553	2.026	-0.5287	1.4996
		FD	0.615	2.0983	-0.0069	0.6298	1.8719	-0.5599	1.6435
		SD	0.6128	2.0178	0.002	0.5981	2.102	0.2448	1.5774
pH	PLS	RAW	0.7813	0.5361	0.0455	0.7138	0.5859	0.0763	1.8692

	MC	0.7422	0.5826	0.0494	0.8729	0.4094	0.0533	2.8050
	MSC	0.7185	0.6048	0.0513	0.8352	0.463	0.0603	2.4633
	SNV	0.6597	0.6401	0.0543	0.8891	0.4024	0.0524	3.0029
	FD	0.7455	0.5799	0.0492	0.911	0.3462	0.0451	3.3520
	SD	0.7877	0.5461	0.0463	0.5787	0.6245	0.0813	1.5407
IPLS	RAW	0.7484	0.5714	-0.0013	0.7314	0.5788	0.0275	1.9295
	MC	0.7329	0.5923	-0.014	0.7931	0.5144	-0.0025	2.1985
	MSC	0.6486	0.6624	-0.0017	0.7231	0.5934	-0.0588	1.9004
	SNV	0.6326	0.6601	0.006	0.775	0.607	0.0517	2.1082
	FD	0.6811	0.6376	0.0049	0.7157	0.5873	-0.0092	1.8755
	SD	0.745	0.5916	-0.0027	0.7448	0.5128	0.0363	1.9795
Si-PLS	RAW	0.8074	0.5083	-0.0172	0.6853	0.6155	-0.0026	1.7826
	MC	0.7642	0.5616	-0.0031	0.8267	0.0246	-0.0246	2.4022
	MSC	0.745	0.581	-0.005	0.865	0.4494	-0.0456	2.7217
	SNV	0.7323	0.5814	0.0097	0.8868	0.4648	0.0472	2.9722
	FD	0.7652	0.5605	-0.0024	0.8121	0.5001	-0.0274	2.3069
	SD	0.8097	0.5201	-0.0023	0.7691	0.4944	-0.0569	2.0811
Bi-PLS	RAW	0.7716	0.5473	-0.0158	0.6936	0.605	0.0068	1.8066
	MC	0.7391	0.5866	-0.01	0.8209	0.5048	0.0812	2.3629
	MSC	0.6926	0.6289	0.0002	0.8377	0.4899	-0.0417	2.4822
	SNV	0.6439	0.6548	0.001	0.8547	0.5526	0.1164	2.6234
	FD	0.7066	0.6171	0.0026	0.8775	0.4385	0.0073	2.8571
	SD	0.7704	0.5661	-0.0071	0.769	0.4969	-0.0508	2.0806

Note: PLS (Partial least square), IPLS (interval partial least square), SiPLS (synergy partial least square), and BiPLS (back interval partial least square), MC (mean centering), MSC (multiplicative scatter correction), SNV (standard normal variate), FD (first derivative), and SD (second derivative).

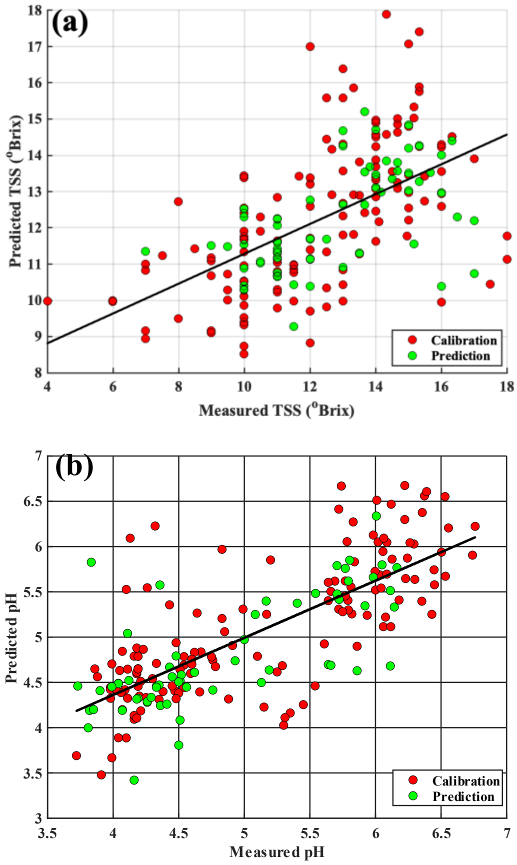


Figure 3.7. Scatter Plot of Reference Measured Versus Handheld NIR Predicted (a) TSS and (b) pH by SiPLS Model.

3.10 Conclusion

The study has demonstrated the possibility for quick, non-destructive evaluations of mango quality using a portable near-infrared spectroscopic technique. SD provided clear separation in the first three PCs and the best PCA cluster trend. Among the identification techniques, the raw, MC, SNV, FD, and SD plus LDA-SVM could be used to identify mango fruit varieties with 100% accuracy in the training set and 97.44% in the prediction set. However, the best model for TSS and pH measurements in mango is synergy partial least square, with an r^2 value of 0.63, an RMSEP value of 1.83, an r^2 value of 0.81, and an RMSEP value of 0.49, respectively. The findings indicated that Si-PLS could be employed for quick, on-site, non-destructive mango TSS and pH testing in developing nations.

Acknowledgments

We acknowledge the support provided by the World Food Program and CSIR-Food Research Institute, Ghana. The assistance provided by Gideon Kwesi Nakotey, Joel Welbeck, and Abigail Sapong is highly acknowledged.

Ethics statement

Not applicable: This manuscript does not include human or animal research.

Conflict of Interest

The authors have declared no conflicts of interest for this article.

Author Contributions

Francis Padi Lamptey: Formal analysis; Investigation; Methodology; Writing-original draft. Ernest Teye: Conceptualization; Data curation; Funding acquisition; Methodology; Resources; Supervision.

Ernest Ekow Abano: Supervision; Writing-review & editing.

Charles Lloyd Yeboah Amuah: Formal analysis; Software; Writing-review & editing.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

CHAPTER FOUR

RAPID AND NON-INVASIVE PREDICTION OF ETHEPHON LEVELS IN MANGO FRUITS USING NEAR-INFRARED SPECTROSCOPY

Francis Padi Lamptey[a,b,e], Charles L. Y. Amuah[c], Vida Gyimah Boadu[a,d,e], Ernest Ekow Abano[a], *Ernest Teye[a,e]

[a] *University of Cape Coast, College of Agriculture and Natural Sciences, School of Agriculture, Department of Agricultural Engineering, Cape Coast, Ghana*

[b] *Cape Coast Technical University, School of Applied Sciences and Technology, Department of Food Science and Postharvest Technology, Cape Coast, Ghana*

[c] *University of Cape Coast, College of Agriculture and Natural Sciences, School of Physical Sciences, Department of Physics, Laser and Fibre Optics Centre, Cape Coast, Ghana*

[d] *Akenten Appiah-Menka University of Skills Training and Entrepreneurial Development,*

Department of Hospitality and Tourism Education, Kumasi, Ghana

[e] *Africa Centre of Excellence for Food Fraud and Safety, AfriFoodinTegrity Centre, University of Cape Coast, Cape Coast, Ghana*

Tel: 233-243170302/+233-243514504

**Email:ernest.teye@ucc.edu.gh / francis.lamptey@cctu.edu.gh*

Submitted: Food Additives & Contaminants: Part A

Statement of Contribution of Joint Authorship

Lamptey F. P.: (Candidate)

Conceptualised the topic, established methodology, data collection and analysis, preparation of tables and figures, writing and compilation of the original manuscript.

Teye, E.: (Principal Supervisor)

Conceptualised the topic, established methodology, supervised and edited the manuscript and co-author of manuscript.

Abano E. E.: (Co-Supervisor)

Conceptualised the topic, supervised and edited the manuscript and co-author of manuscript.

Amuah C. L. Y.: (Research Team member)

Validation, visualization and data curation and co-author of manuscript

Boadu V. G.: (Research Team member)

Validation, visualization and data curation and co-author of manuscript

Abstract

This study investigates using near-infrared (NIR) spectroscopy combined with chemometrics to rapidly and non-invasively predict the concentration of ethephon, a ripening agent, in mangoes. Traditional methods for detecting ethephon are time-consuming and destructive, limiting their applicability in large-scale operations. By analyzing the spectral data of mangoes ripened naturally and those treated with different ethephon concentrations (250, 750, and 1000 ppm), this study aimed to develop predictive models using NIR spectroscopy. The mangoes were scanned with a portable NIR spectrophotometer, and the spectral data were processed using multiple preprocessing techniques, including multiplicative scatter correction (MSC) and standard normal variate (SNV). Principal component analysis (PCA) was used to group samples based on spectral differences, while classification models such as neural networks (NN) and random forests (RF) were applied to categorize the ethephon concentrations. The best-performing model was the neural network (NN) combined with MSC preprocessing, achieving 100% training and test set accuracy. Additionally, partial least squares (PLS) regression demonstrated excellent predictive accuracy for ethephon concentration when preprocessed with SNV, yielding an r^2 of 0.904 and RMSEC of 0.313 for the training set, and an r^2 of 0.996, RMSEP of 0.068, and an RPD value of 15.253 for the test set. These findings suggest that NIR spectroscopy offers a rapid, non-destructive, and effective method for monitoring the artificial ripening process in mangoes.

Keywords: Near-infrared; chemometrics; mango; ethephon; ripening

4.1 Introduction

Fruits are a vital component of human nutrition, providing essential nutrients such as vitamins, minerals, dietary fibre, and antioxidants that promote growth and overall health (Asif, 2011; Meena, Yadav, & Meena, 2020). As part of a balanced diet, fruits should be consumed regularly. Climacteric fruits like mangoes (Mangifera indica) are particularly notable because they continue to ripen after being harvested, provided they reach the end of their growth phase (Meena et al., 2020; Paul, Pandey, & Srivastava, 2012). Ripening in climacteric fruits is accompanied by physiological and biochemical changes, including the development of flavor, aroma, softening of tissues, and changes in pigmentation (Anwar, Mattoo, & Handa, 2018; Kou et al., 2021). This ripening process is typically marked by increased respiration and ethylene production (Kou & Wu, 2018; Meena et al., 2020).

While natural ripening of climacteric fruits at room temperature is possible, it is often slow and can result in uneven ripening and reduced fruit quality (Payasi & Sanwal, 2010; Siddiqui & Dhua, 2010). Artificial ripening techniques are widely used in commercial fruit production to address this issue, with ethylene gas being the standard agent for inducing uniform ripening (Payasi & Sanwal, 2010). Calcium carbide and ethephon are the commonly used artificially ripened agents. Fruits ripened with ethephon display more consistent colour and longer shelf life compared to those ripened naturally or with calcium carbide (CaC_2), a hazardous chemical agent banned in many countries (Adeyemi, Bawa, & Muktar, 2018; Kesse et al., 2019; Siddiqui & Dhua, 2010).

The use of CaC_2 is particularly concerning due to its contamination with arsenic and phosphorus, which pose significant health risks (Siddiqui & Dhua, 2010). Despite this, CaC_2 remains used in countries such as India, Pakistan, Bangladesh, Nepal, and Ghana (Kesse et al., 2019; Siddiqui & Dhua, 2010). Health hazards associated with CaC_2 include neurological effects, mood disturbances, and more severe conditions such as seizures (Adeyemi et al., 2018). Ethephon, another chemical ripening agent, is considered less harmful when used in permissible concentrations, though it can still cause skin and eye irritation and long-term health issues (Ruwali et al., 2022). Although regulatory bodies aim to safeguard food safety, many African countries still face challenges in establishing and enforcing specific regulations to control artificial fruit ripening (Islam et al., 2016).

Traditional detection methods for ethephon involve complex and labor-intensive chemical analyses that require specialized laboratory equipment. While accurate, these methods are time-consuming and destructive, making them impractical for large-scale or field-based applications. As a result, there is a growing need for rapid, non-destructive techniques that can reliably predict the concentration of ripening agents in fruits. Near-infrared (NIR) spectroscopy has gained widespread use as a non-destructive, rapid, and online alternative to traditional destructive methods for measuring the quality components of food products. The recent advancements in NIR spectroscopy have led to the development of battery-powered, handheld devices. Combined with chemometrics, NIR spectroscopy offers several advantages over conventional laboratory analyses, including faster analysis, non-invasive measurements, minimal or no sample preparation, zero chemical use, and portability for on-site quality control (Mahanti & Chakraborty, 2020). NIR spectroscopy has been successfully utilized in mango quality assessment and prediction (Nordey et al., 2019; Polinar et al., 2019), food safety, and detection of adulteration (Lakade et al., 2019; Mahanti & Chakraborty, 2020; Teye, Elliott, Sam-Amoah, & Mingle, 2019).

This study explores the feasibility of using NIR spectroscopy to identify and predict the concentration of ethephon used in the artificial ripening of mango fruits. By analyzing spectral data obtained from mango samples treated with different ethephon concentrations, we seek to develop robust predictive models that can be employed for quality control and regulatory compliance. The successful application of NIR spectroscopy would provide a valuable tool for ensuring the safety and integrity of artificially ripened mangoes in the marketplace.

4.2 Methodology

Unripe mangoes were purchased from a farmer in the Eastern region of Ghana. The mangoes were manually cleaned to remove all the adhering dust and dirt. A total of 80 mangoes were used in this study; 20 were ripened naturally, and the remaining were ripened artificially at different concentrations. The selected mango fruits Cv. Keitt was kept in the ripening chamber at the University of Cape Coast, School of Agriculture laboratory, Cape Coast, Ghana. For the preparation of 250, 750, and 1000 ppm of ethephon solutions, 0.64, 1.92, and 2.56 mL of ethrel (2-Chloroethylphosphonic acid 39% aqueous solution under the brand name chemophon 480 SL), respectively, were dissolved in 1litre of distilled water (Lavanya, Rao, Edukondalu, Lakshmypathy, & Rao, 2019). Uniform sized fruits were dipped in ethephon solution for 5 min and air-dried to remove excess moisture. Fruits were treated with different concentrations, such as 250, 750, and 1000 ppm. All the sets were allowed to ripen artificially for 48–72 h at room temperature, and then the signature spectra of each mango sample were obtained and analysed for interpretation.

4.2.1 NIR Spectra Acquisition

The spectral acquisition was performed using a portable NIR spectrophotometer (Innospectra Co., Hsinchu, Taiwan), having a wavelength specific to 900– 1700 nm, as shown in Figure 4.1. The spectrometer included a measurement unit connected to a computer. Each fruit was scanned in such a way that the maximum surface of the fruit could be covered. Each fruit was rotated clockwise, and six scans were taken from each fruit. Then, the obtained NIR spectra were used for further statistical analysis.

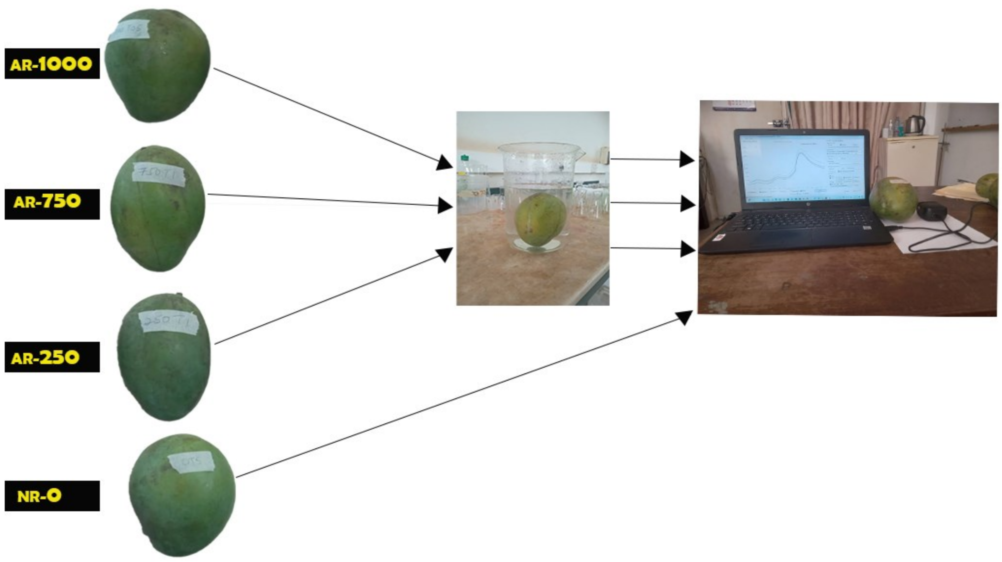


Figure 4.1. Experimental set-up

4.2.2 Statistical Analysis, Data Processing and Analysis

Spectral data recordings stored on the computer are imported into chemometric software packages: Matlab version 9.5.0 (Mathworks Inc., USA) with Windows 10 Basic for all data processing. In this study, the total data set was divided into two groups, one for building the model and the other for testing the model's actual predictability. This study employed repeated cross-validation to partition the spectral data into training (calibration) and test subsets. A tenfold repeated cross-validation technique was employed. The data was partitioned into ten groups, with nine groups serving as calibration sets and the remaining group as a test set. The test set was subsequently modified until all groups had been tested (Ramírez-Morales, Rivero, Fernández-Blanco, & Pazos, 2016).

4.2.3 Chemometric Analysis

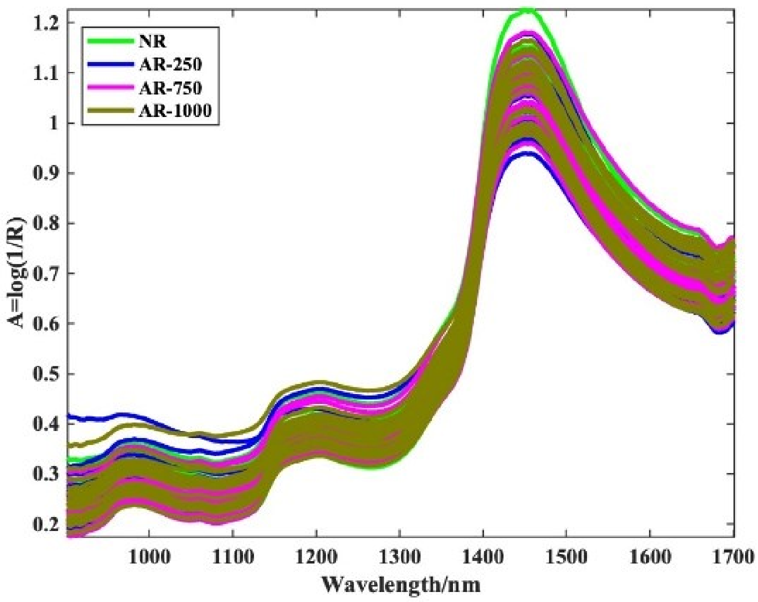
Five pre-processing methods, mean centering (MC), multiplicative scatter correction (MSC), standard normal variate (SNV), first derivative (FD), and second derivative (SD), were used comparatively to control external influences so that the outcome of the model would be based on the chemical fingerprint from the spectral information acquired (Teye, Elliott, et al., 2019). For preliminary classification, by an unsupervised technique, principal component analysis (PCA) was performed on the selected spectra, and score plots between different principal components (PCs) were plotted to see the formation of different groups of samples based on their varieties. After ascertaining the group, the final classification was carried out using supervised methods, partial least squares (PLS), to discriminate the analysis. After that, the PLS regression method was employed for the classification of the samples, extracting the contributed variables, and then validating and predicting the results obtained (Lakade et al., 2019). The performance of the multivariate models was evaluated using several metrics, including the coefficient of determination (r^2), root mean square error of prediction (RMSEP), bias (bs), Range Error Ratio (RER), and residual predictive deviation (RPD). For a model to be considered robust, it should exhibit a high RPD, a high r^2 value, and low RMSEP and bias values (Lamptey et al., 2023; Liu, He, Wang, & Sun, 2011). All statistical analyses were performed using MATLAB.

4.3 Results and Discussions

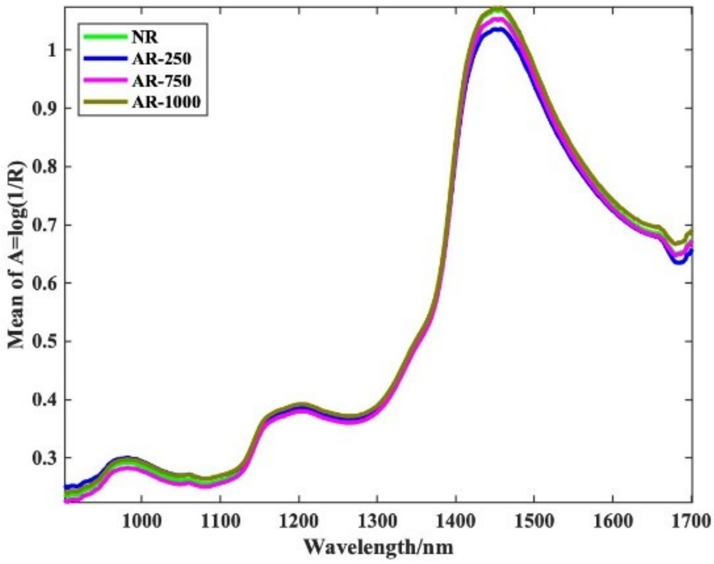
4.3.1 Spectral Analysis and Preprocessing

The spectral absorbance values of mangoes treated with 250, 750, and 1000 ppm ethephon and naturally ripened mangoes are shown in Figure 4.2a. While all the spectral curves follow a similar trend, the absorbance intensities vary, suggesting that the fruits contain the same internal substances but in different quantities (Mahanti & Chakraborty, 2020). Figure 4.2b displays the mean spectral absorbance of ethephon-treated mangoes at different concentrations and naturally ripened mangoes. The spectra are dominated by the second and third overtone regions of the C–H bonds associated with sugar solutions, showing peak absorption at 970 nm, 1200 nm, and 1450 nm (Guthrie, Walsh, Reid, & Liebenberg, 2005; Mahanti & Chakraborty, 2020). The peaks around 970 nm and 1200 nm are also linked to water absorption (Fan, Li, Huang, & Chen, 2017). Mango, a fleshy fruit, contains more than 80% water, and its sugar content increases as it ripens (Léchaudel & Joas, 2007; Padda, do Amarante, Garcia, Slaughter, & Mitcham, 2011). Ethephon (2-chloroethylphosphonic acid) releases ethylene, accelerating ripening

(Fleming, Martinez, Mallea, & Guerra, 2014). This artificial ripening process may cause differences in sugar accumulation, organic acid metabolism, and water content compared to naturally ripened fruits. The peak at 1450 nm is attributed to the first overtone of the O-H stretching band and a combination band (Omar, Atan, & MatJafri, 2012). Moreover, the peak at 1650 nm may be related to fruit attributes such as hardness and the first overtone of the C–H bonds in carotenoids (Abarra, Serrano, Sabularse, Mendoza, & Del Rosario, 2018; Toledo-Martín et al., 2018).



a. Raw



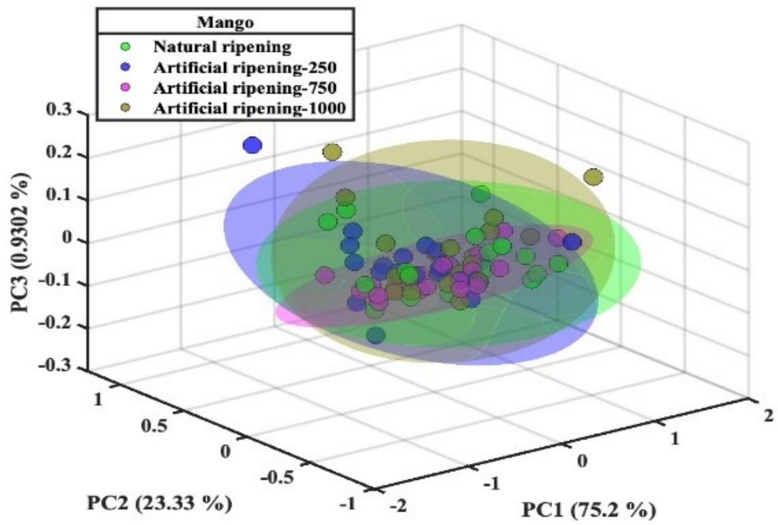
b. Mean

Figure 4.2. Raw and mean Spectral profile of naturally ripened mangoes and artificially treated mangoes with 250, 750, and 1000 ppm ethephon.

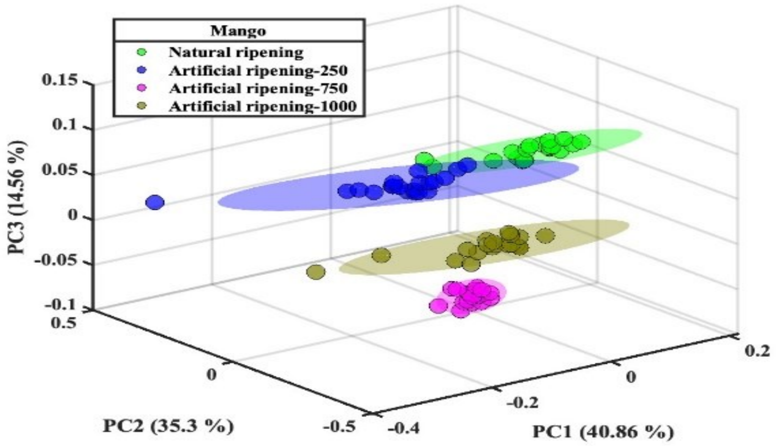
4.3.2 PCA (Principal Component Analysis)

Principal component analysis (PCA), as an unsupervised pattern recognition technique, brings out useful information so that similar samples are grouped closer to each other. This approach enables the visualization of data trends in a dimensional space (Teye, Elliott, et al., 2019). In this regard, the graphical output could be useful in ascertaining the differences between the categories of different mango samples used in our study. Figure 4.3 (a-b) shows the PCA score plots for raw and MSC (multiplicative scatter correction) preprocessed spectral data, respectively. From these figures, it can be seen that four major groups of mango samples were used in this experiment: mangoes treated with 250, 750, and 1000 ppm ethephon and naturally ripened mangoes. In the raw spectral data PCA plot (Figure 4.3a), there was some overlap between the groups, indicating that the raw spectral data alone may not provide perfect separation between the four categories.

The PCA score plot for the MSC preprocessed data (Figure 4.3b) shows a marked improvement in the separation between mangoes treated with 250, 750, and 1000 ppm ethephon and naturally ripened mango samples. The clusters for each category appear more compact and distinctly separated, demonstrating the effectiveness of MSC preprocessing in enhancing the spectral differences between the four groups. These graphical plots provide vital information that could be applied to differentiate within and between groups used in this study. The clear separation observed in the MSC preprocessed data suggests that NIR spectroscopy combined with PCA can effectively distinguish between mangoes treated with 250, 750, and 1000 ppm ethephon and naturally ripened mangoes. The neat separation of mangoes treated with 250, 750, and 1000 ppm ethephon and naturally ripened mango samples in the MSC preprocessed PCA plot suggests that NIR spectroscopy, combined with appropriate preprocessing and multivariate analysis techniques, has significant potential as a rapid and non-destructive method for detecting artificial ripening in mangoes.



a. Raw



b. MSC

Figure 4.3. PCA plot of the topmost three PCs for naturally ripened mangoes and artificially treated mangoes with 250, 750, and 1000 ppm ethephon

4.3.3 Classification Models

The classification results (in Table 4.1) reveal that neural networks (NN) and random forests (RF) generally outperformed other models, particularly when combined with MSC preprocessing. The high accuracy achieved (100% for both training and test sets with MSC preprocessing for

NN, and 100% for training and 98.75% for test sets for RF) indicates the robust capability of NIR spectroscopy in distinguishing between different levels of ethephon treatment. The superior performance of NN and RF models can be attributed to their ability to capture complex, non-linear relationships in the spectral data (Nawar & Mouazen, 2019; Qi et al., 2019). Even though NN and RF showed excellent performance, they also have the potential for overfitting, especially with limited sample sizes (Lamptey et al., 2023). The high accuracy in the training set (100% for RF across all preprocessing methods) compared to the generally lower test set performance suggests some overfitting may be present. The k-nearest neighbors (K-NN) algorithm showed notably high performance with MSC preprocessing (97.78% for training and 97.50% for test sets), suggesting that this model can also be effective when combined with appropriate preprocessing techniques.

4.3.4 Quantitative Prediction of Ethephon Concentration

The Partial Least Squares (PLS) regression results demonstrate the potential of NIR spectroscopy for quantitative prediction of ethephon concentration, as illustrated in Table 4.2. The best performance was achieved with SNV preprocessing, yielding an r^2 of 0.996, RMSEP of 0.068, and RPD of 15.253 for the test set. These metrics indicate excellent predictive capability, as 8.1 or higher is considered excellent for any application (Polinar et al., 2019). The high RPD and RER values across different preprocessing techniques (particularly for SNV, MSC, and second derivative) suggest that the developed models are robust and can predict ethephon concentration across a wide range of values. MSC and SNV can eliminate additive and multiplicative impacts within spectra (Padhi et al., 2024). This finding is consistent with other studies in identifying artificially ripened fruit using NIR spectroscopy, such as the work by Mahanti and Chakraborty (2020) on sapota, where SNV preprocessing also yielded superior results. The superior performance of the SVM model can be attributed to its strong self-learning and self-adjustment capabilities. Other researchers have also noted that SVM's advantage lies in incorporating the structural risk minimization principle, which reduces the upper boundary on expected risk compared to other techniques (Teye et al., 2019).

4.4 Conclusion

This study demonstrates the powerful potential of NIR spectroscopy combined with chemometric analysis as a rapid and non-destructive method for detecting and quantifying ethephon in artificially ripened mangoes. The Neural Network model, particularly when paired with MSC preprocessing, exhibited exceptional performance, achieving 100% accuracy in both the training and test sets. Furthermore, the Partial Least Squares (PLS) regression model, preprocessed with SNV, demonstrated outstanding predictive capability, with an r^2 of 0.996 and an RPD value of 15.253 for the test set, indicating its robustness for quantitative predictions. These findings contribute significantly to developing more efficient and reliable methods for ensuring food safety and authenticity in the fruit industry.

Table 4.1: Performance of Spectral Pre-Processing Techniques During Model Development and Cross Validation

Models	RAW		MC		MSC		SNV		FD		SD	
	Preprocessing											
	Training set	Test set	Training set	Test set	Training set	Test set	Training set	Test set	Training set	Test set	Training set	Test set
K-NN	52.64	28.75	55.69	21.25	97.78	97.50	77.08	43.75	65.97	38.75	65.97	38.75
NN	90.83	31.25	91.67	31.25	100.00	100.00	93.61	48.75	94.86	35.00	95.69	35.00
RF	100.00	32.50	100.00	33.75	100.00	98.75	100.00	56.25	100.00	38.75	100.00	42.50

Note: K-NN (K-nearest neighbor), NN (Neural network), RF (Random forest), MC (mean centering), MSC (multiplicative scatter correction), SNV (standard normal variate), FD (first derivative), and SD (second derivative).

Table 4.2: Prediction of Concentration of the Ethephon

PLS		Training set				Testing set					
	Factors	r^2	RMSEC	SEC	Bias	r^2	RMSEP	SEP	Bias	RPD	RER
raw	11	0.841	0.403	0.407	0.054	0.948	0.230	0.230	0.047	4.499	11.142
mc	11	0.873	0.360	0.363	0.048	0.980	0.144	0.144	0.029	7.192	17.812
msc	11	0.902	0.317	0.319	0.042	0.989	0.108	0.108	0.022	9.591	23.753
snv	11	0.904	0.313	0.316	0.042	0.996	0.068	0.068	0.014	15.253	37.775
fd	11	0.765	0.491	0.495	0.066	0.957	0.210	0.210	0.043	4.926	12.201
sd	11	0.973	0.166	0.167	0.022	0.993	0.087	0.087	0.018	11.901	29.474

Note: r^2 (Coefficient of determination), PLS (Partial least squares), RMSEC (root mean square error of calibration), RMSEP (root mean square error of prediction), TSS (Total soluble solids), MC (mean centering), MSC (multiplicative scatter correction), SNV (standard normal variate), FD (first derivative), and SD (second derivative).

Acknowledgments

We appreciate the help provided by Judith Awudi, Department of Agricultural Engineering, CSIR-Food Research Institute, World Food Programme, and School of Agriculture, UCC.

Conflict of interest statement

The authors have declared no conflicts of interest for this article.

Credit authorship contribution statement:

Conceptualization: FPL, CLYA, VGB; ET, Formal analysis: FPL, CLYA, ET; Investigation: FPL, VGB, ET; Funding acquisition: FPL, ET; Methodology: FPL, CLYA, EEA, VGB, ET; Writing – original draft: FPL, ET; Data curation: FPL, CLYA, EEA, VGB, ET; Funding acquisition: FPL, ET; Resources: FPL, CLYA, VGB, ET; Supervision: CLYA, EEA, ET; Writing – review & editing: FPL, CLYA, EEA, VGB, ET.

CHAPTER FIVE

FEASIBILITY STUDY ON FINGERPRINTING ORGANIC AND CONVENTIONAL MANGO FRUITS, CHIPS, AND JUICE USING PORTABLE NEAR-INFRARED SPECTROSCOPY

*Francis Padi Lamptey[1,2,5], *Ernest Teye[1,5], Sandra Ama Kaburi[2], Flavio Odoi-Yorke[3], Charles L. Y. Amuah[4,5], Ernest Ekow Abano[1], Gifty Serwaa Otoo[2]

[1] University of Cape Coast, College of Agriculture and Natural Sciences, School of Agriculture, Department of Agricultural Engineering, Cape Coast, Ghana

[2] Cape Coast Technical University, School of Applied Sciences and Technology, Department of Food Science and Postharvest Technology, Cape Coast, Ghana

[3] Cape Coast Technical University, School of Engineering, Department of Renewable Energy Technology, Cape Coast, Ghana

[4] University of Cape Coast, College of Agriculture and Natural Sciences, School of Physical Sciences, Department of Physics, Laser and Fibre Optics Centre, Cape Coast, Ghana

[5] Africa Centre of Excellence for Food Fraud and Safety, AfriFoodinTegrity Centre, University of Cape Coast, Cape Coast, Ghana

Tel: 233-243170302/+233-243514504

*Email: ernest.teye@ucc.edu.gh / francis.lamptey@cctu.edu.gh

Abstract

This research examined the distinction between organic and conventional mango fruits, chips, and juice using portable near-infrared (NIR) spectroscopy. A comprehensive analysis was conducted on a sample of 100 mangoes (comprising 50 organic and 50 conventional) utilising a portable NIR spectrometer that spans a wavelength range from 900 to 1700 nm. The mangoes were assessed in their entirety and their juice and chip forms. The spectral data underwent pre-processing through methodologies such as multiplicative scatter correction (MSC), standard normal variate (SNV), and derivatives to enhance the precision of the models. Principal component analysis (PCA) and various multivariate classification algorithms, including linear discriminant analysis (LDA), random forest (RF), k-nearest neighbors (KNN), and partial least squares discriminant analysis (PLSDA), were utilised to categorise the samples effectively. The findings indicated that the random forest method and specific pre-processing techniques achieved the highest classification accuracy for distinguishing organic and conventional mango products. For mango fruit and chips, it achieved 88.76% and 77.98% accuracy, respectively, when pre-processed using the second derivative, while for juice, it achieved 87.53% accuracy without pre-processing. This investigation demonstrates the efficacy of portable NIR spectroscopy as a dependable and non-invasive method for verifying organic mango products, thereby enhancing the integrity of food labelling and fostering consumer confidence.

Keywords: Classification, Mango products, Near infrared spectroscopy, Organic, Pre-processing.

5.1. Introduction

The global market for organic food has experienced substantial growth in recent years, primarily propelled by environmental sustainability considerations. Organic agriculture, which employs organic techniques and avoids synthetic pesticides, is pivotal in enhancing food security, product quality, and overall sustainability (Andika & Bidayati, 2024). Among the array of organic products, the mango (*Mangifera indica L.*) is recognised as one of the most important fruits globally, attributed to its remarkable nutritional and economic importance and its augmented production, trade, and consumption (Jahurul et al., 2015; Yahia, de Jesús Ornelas-Paz, Brecht, García-Solís, & Celis, 2023).

Mango is presented in a multitude of forms, encompassing fresh fruit, puree, syrup-packed slices, leather, canned slices, chutney, juice concentrate, ready-to-drink juice, wine, jams, jellies, pickles, smoothies, chips, and powder (Jahurul et al., 2015; Owino & Ambuko, 2021), each of which undergoes consumer evaluation regarding its organic authenticity. It is imperative to ensure that products marketed as organic are verifiably organic to uphold consumer trust and adhere to regulatory compliance (Hamzaoui Essoussi & Zahaf, 2009; Pivato, Misani, & Tencati, 2008).

Traditionally, the distinction between organic and conventional products has depended on certification methodologies, some of which can be time-consuming and costly (Leitner & Vogl, 2020; Zorn, Lippert, & Dabbert, 2009). Conventional foods are also distinguished from their organic counterparts through the presence of undesirable contaminating substances. Unfortunately, the detection of these contaminants is fraught with challenges.

It necessitates sophisticated laboratory facilities, such as high-performance liquid chromatography (HPLC) and gas chromatography, which are often inaccessible within the food distribution network or proximal to the point of sale and involve the destruction of the sample under examination (López, Arazuri, García, Mangado, & Jarén, 2013; Teye & Amuah, 2022). Consumers' confidence constitutes a critical facet of the food market; thus, developing a cost-effective analytical system appropriate for field application within the food sector and reliably distinguishing organic from conventional products would be of immense value (Song, Wang, Maguire, & Nibouche, 2016).

These constraints have created the necessity for more efficient, nondestructive, and economically viable methods for authenticating organic products. As a result, portable near-infrared (NIR) spectrometers have emerged as a viable alternative, facilitating rapid, nondestructive analysis with minimal or simplified sample preparation while being compact, cost-effective, and user-friendly (Gullifa et al., 2023; Yin et al., 2019). In contrast to the previously detailed methodologies, NIR spectroscopy has surfaced as an innovative and user-friendly technique applied across diverse domains such as agriculture, food and beverages, petrochemicals, and biochemistry. The spectral data acquired from the NIR profile predominantly arises from the internal vibrations and absorptions of hydrogen-containing groups, including O–H, C–H, N–H, and other functional groups, in overtone and combination bands (Lin et al., 2019; Teye & Amuah, 2022). This data can be employed for both qualitative and quantitative predictions by applying chemometrics.

This technology has been successfully implemented in various sectors of food analysis, including the differentiation of organic and conventional products such as apples, cocoa, feed, green asparagus, milk, pineapple fruits, pineapple juice, and rice (Amuah et al., 2019; Anyidoho, Teye, & Agbemafle, 2021; Lamptey, Amuah, Boadu, Abano, & Teye, 2024; Liu et al., 2018; Sánchez, Garrido-Varo, Guerrero, & Pérez-Marín, 2013; Song et al., 2016; Tres, Van Der Veer, Perez-Marin, Van Ruth, & Garrido-Varo, 2012; Xiao, Liu, Zhang, Ma, & Ngadi, 2019). Some researchers have utilised NIR in various studies to assess mangoes, including determining their maturity levels (dos Santos Neto, de Assis, Casagrande, Júnior, & de Almeida Teixeira, 2017), detecting internal defects (Raghavendra, Guru, & Rao, 2021), verifying the authenticity of specific varieties (Jha et al., 2013), measuring key quality parameters such as total soluble solids (TSS) and pH (Lamptey, Teye, Abano, & Amuah, 2023).

In the present investigation, we explore the application of NIR spectroscopy to distinguish between organic and conventional mango products, including mango fruits, chips, and juice. The objective is to formulate a dependable and practical methodology for differentiating organic mango products from their conventional counterparts. The anticipated results of this investigation encompass the enhancement of methodologies for authenticating organic products, potentially providing advantages to both the food industry and consumers by safeguarding the authenticity of organic food items. This research addresses the difficulties encountered in authenticating organic produce, specifically emphasizing the necessity for expedited, more accessible, and non-invasive approaches.

5.1. Methodology

5.2.1. Sample Collection

In the present investigation, both organic and conventional varieties of mangoes were procured from farmers trained by the Ministry of Food and Agriculture within the Eastern region of Ghana. The selection consisted of 50 organic mangoes alongside 50 conventional counterparts. A thorough visual inspection of the fruits was conducted, and only the wholesome fruits were deemed suitable for this study. All fruits were meticulously labelled for identification purposes and subsequently transported to the laboratory of the School of Agriculture at the University of Cape Coast. The labelled fruits underwent an initial scanning using a portable near-infrared (NIR) spectroscopy device, operating within the wavelength range of 900 nm to 1700 nm. After scanning, each fruit was washed, peeled, deseeded, and sliced into chips or strips of irregular dimensions and configurations. A subset of the sliced mangoes was dehydrated employing a dehydrator set at a temperature of 65°C (Kayode, Joshua, & Oyetoro, 2023), while the remaining mangoes were processed using an electric juicer, with the extracted juice being meticulously packaged in bottles (Lamptey et al., 2024). Subsequently, the mango juice and the dehydrated mango chips were scanned using the portable NIR device.

5.2.2. Spectra Collection

The handheld NIR (Innospectra Co., Hsinchu, Taiwan) spectrometer (NIR-M-R1) was employed in reflectance mode to obtain spectral data of the mangoes, chips, and juice within the 900–1700 nm wavelength range. Before each scanning session, the device was calibrated per the manufacturer's guidelines to ensure accuracy and consistency. Scanning of the whole mango fruits was executed immediately following the labelling process. Each mango was secured in a stable holder, and the NIR spectrometer was utilised to scan three distinct points on the surface, thereby facilitating the acquisition of an average spectral reading. Following the juicing process, an analogous methodology was employed for the bottled mango juice, with the NIR spectrometer scanning through the juice's container to capture the spectral data. Ultimately, the dehydrated mango chips were subjected to scanning three times across various sections to account for potential variability in thickness or texture.

5.2.3. Initial Spectra Data Processing

In this investigation, four distinct pre-processing methodologies—multiplicative scatter correction (MSC), standard normal variate (SNV), first derivative (FD), and second derivative (SD)—were employed comparatively to control external influences, thereby ensuring that the resultant model was fundamentally based on the chemical fingerprint derived from the acquired spectral information (Teye, Elliott, Sam-Amoah, & Mingle, 2019). MC computes spectra by averaging and subsequently subtracting the mean from each spectrum. MSC effectively eliminates unwanted scattering effects and compensates for non-uniform scattering. SNV is a transformational spectral treatment technique that removes multiplicative interferences stemming from scatter, particle size, and light path length. Both methodologies linearize each spectrum to align with an ideal spectrum. The first derivative pre-processing technique reduces baseline effects by comparing the spectra from two consecutive measurement locations. The second derivative transformation algorithm distinguishes overlapping peaks and enhances resolution while eliminating additive and multiplicative baseline influences within the spectra. Please consult the following studies for additional insights regarding these pre-processing methods (Amuah et al., 2019; Lamptey et al., 2023; Teye, Amuah, McGrath, & Elliott, 2019; Teye, Elliott, et al., 2019).

5.2.4. Principal Component Analysis

After the preliminary pre-processing of the spectral data in this investigation, principal component analysis (PCA) was implemented as an unsupervised pattern recognition technique to elucidate data trends within a diminished dimensional space through score plots. PCA is a prevalent analytical method that diminishes the dimensionality of the data matrix by compressing the information into principal components, encompassing interpretable and significant variables. In the context of this analysis, the initial three principal components (PCs) were selected, as they encapsulate the most substantial information while exhibiting minimal or no redundancy (Teye & Amuah, 2022).

5.2.5. Data Partitioning

In order to construct a more robust and generalized model, the dataset was segmented into two distinct subsets: a training set and a test set. This investigation utilised repeated cross-validation to systematically partition the spectral data into training (calibration) and test subsets. A ten-fold repeated cross-validation methodology was employed. The data was stratified into 10 groups, with nine (9) groups functioning as calibration sets and the remaining group designated as the test set. The test set was subsequently modified until all groups had undergone testing (Ramírez-Morales, Rivero, Fernández-

Blanco, & Pazos, 2016). This procedure was applied to the fruits, chips, and juice categories.

5.2.6. Multivariate Classification Algorithms

The present study identified organic mango fruits, chips, and juice from their conventional counterparts. The identification methodologies employed to address this challenge were comparatively analysed, including linear discriminant analysis (LDA), random forest (RF), K-nearest neighbors (KNN), and partial least squares discriminant analysis (PLSDA). RF (random forest) is an ensemble technique predicated on tree classifiers that grows multiple classification trees to yield precise discrimination. It utilises two metrics of variable importance and data resemblance measures for graphical representation, multidimensional scaling, and clustering (Anyidoho et al., 2021). KNN (K-nearest neighbors) is a nonparametric and linear learning algorithm that evaluates the distance between samples from the calibration set and unknown samples (Thanh Noi & Kappas, 2017). The parameter K exerts a significant influence on the classification rate of the KNN model. LDA (linear discriminant analysis) constitutes a linear and parametric supervised pattern recognition technique to uncover a linear combination of features for linear classifiers (Anyidoho et al., 2021). PLS-DA (partial least squares discriminant analysis) represents a linear differentiation technique that combines the characteristics of partial least squares regression with a discriminative presentation (Lee, Liong, & Jemain, 2018). Figure 5.1 represents a flowchart of the classification process for distinguishing between organic and conventional mangoes.

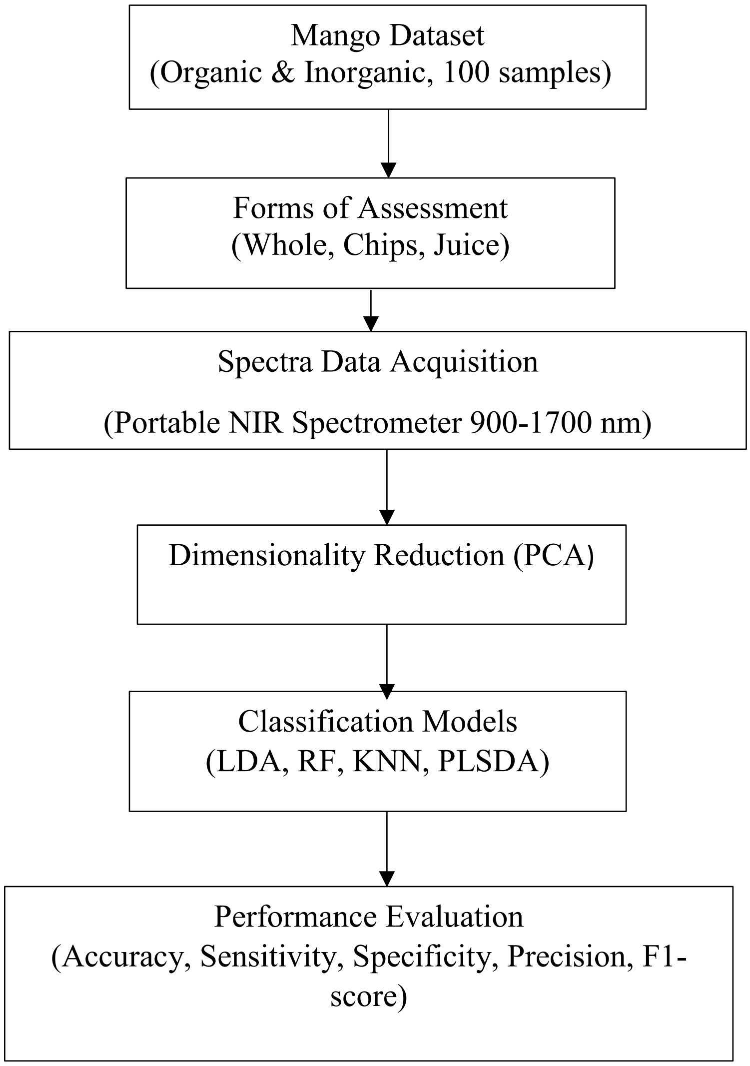


Figure 5.1. Flowchart illustrating the classification process for distinguishing between organic and conventional mangoes.

5.3. Results and Discussions

5.3.1. Spectra Analysis

The spectral characteristics of the mango samples examined in this research disclosed unique spectral fingerprints across the various forms of mango—namely, fresh fruits, dried chips, and juice—as depicted in the accompanying figures (see Figures 5.2A, B, and C). The observed variations may be ascribed to differences in surface texture, pigmentation, and moisture content among the distinct states (Keskin, Soysal, Arslan, Sekerli, & Celiktas, 2018; Omolola, Jideani, & Kapila, 2017). Figure 5.2A depicts the spectral profile corresponding to the mango fruits, Figure 5.2B represents the profile for dried mango chips, and Figure 5.2C illustrates the spectral profile of the mango juice. These visual representations show that each mango product manifests unique spectral patterns reflecting their respective physical and chemical attributes.

The average spectra of the mango fruits, chips, and juice were subsequently compared and are presented in Figure 5.4. Among the three forms analysed, the juice exhibited a more intense reflectance signal throughout the wavelength spectrum, particularly beyond 1400 nm. This signifies a pronounced water absorption band in near-infrared spectroscopy (Lammertyn, Peirs, De Baerdemaeker, & Nicolaı, 2000). This observation implies that the liquid state of the mango engages in distinctive interactions with near-infrared light, potentially due to its homogeneous nature and elevated moisture content relative to the other two forms (Londoño et al., 2017; Nyangena, Owino, Ambuko, & Imathiu, 2019; Othman & Mbogo, 2009).

The chips and fruits exhibit distinct peaks within the reflectance spectrum in the range of 1400-1500 nm, likely related to the specific absorption characteristics of water corresponding to their first overtone (Cayuela & Weiland, 2010). The 1400 - 1500 nm wavelength range is associated with CH_2, CH, and ROH, indicating the presence of oil content (Iqbal, Herodian, & Widodo, 2019). Noteworthy peaks observed around 1200 nm and 1450 nm in the fruit samples can be attributed to the second overtone of C-H and the first overtone of O-H, respectively (Eldin & Akyar, 2011), which indicate the presence of sugars and various carbohydrates within the fruit (Li et al., 2021).

Furthermore, the absorption band identified around 970-980 nm in the dried mango chips reflects a diminished moisture content in these dehydrated products (Ishikawa, Ueno, & Fujii, 2017). Figures 5.3A, B, and C present the mean spectra of the organic and conventional mango fruits, chips, and juice, respectively. These mean spectra provide enhanced clarity regarding the distinctions between organic and conventional samples of each product. Figure 5.4 compares the mean spectral profiles across all three product categories (fruits, chips, and juice) for organic and conventional mangoes.

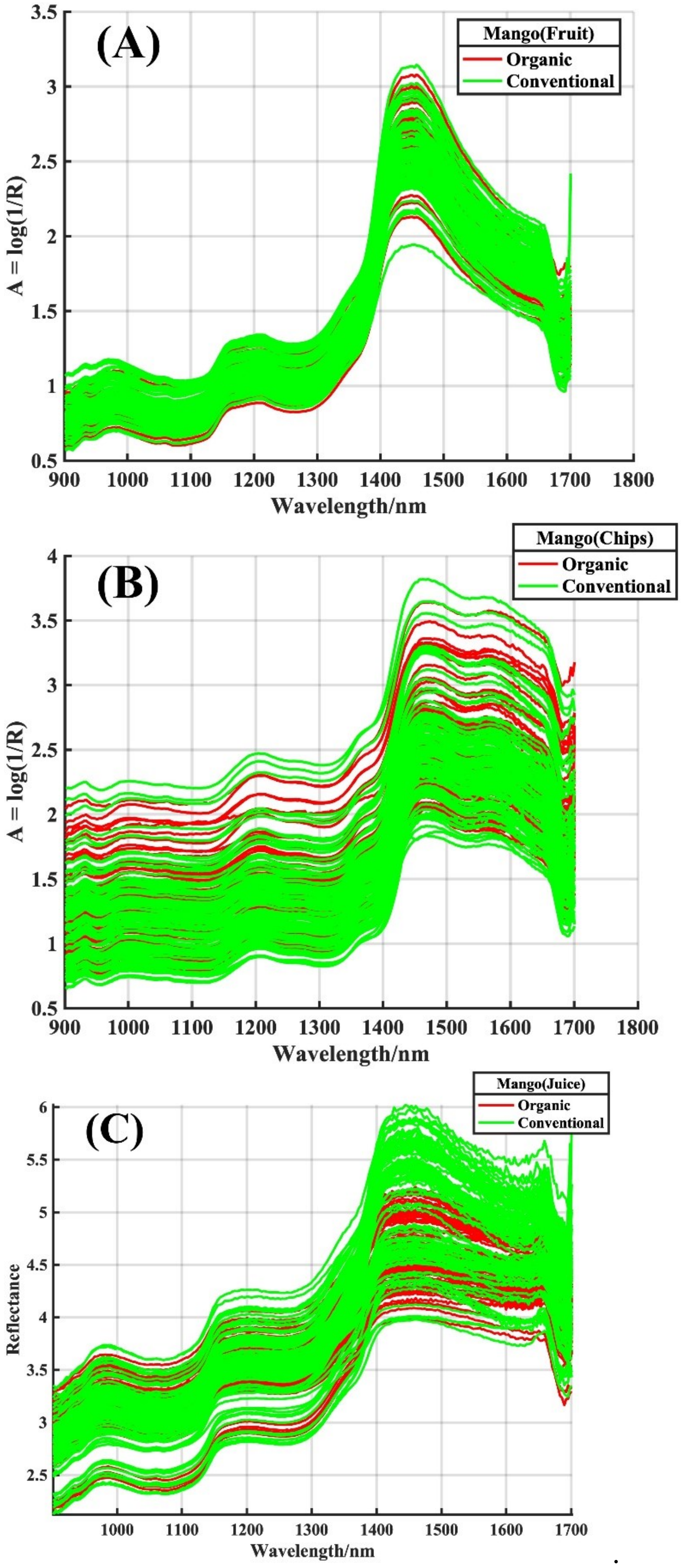


Figure 5.2. Raw Spectral Profile of Organic and Conventional Mango (A) Fruits (B) Chips (c) Juice

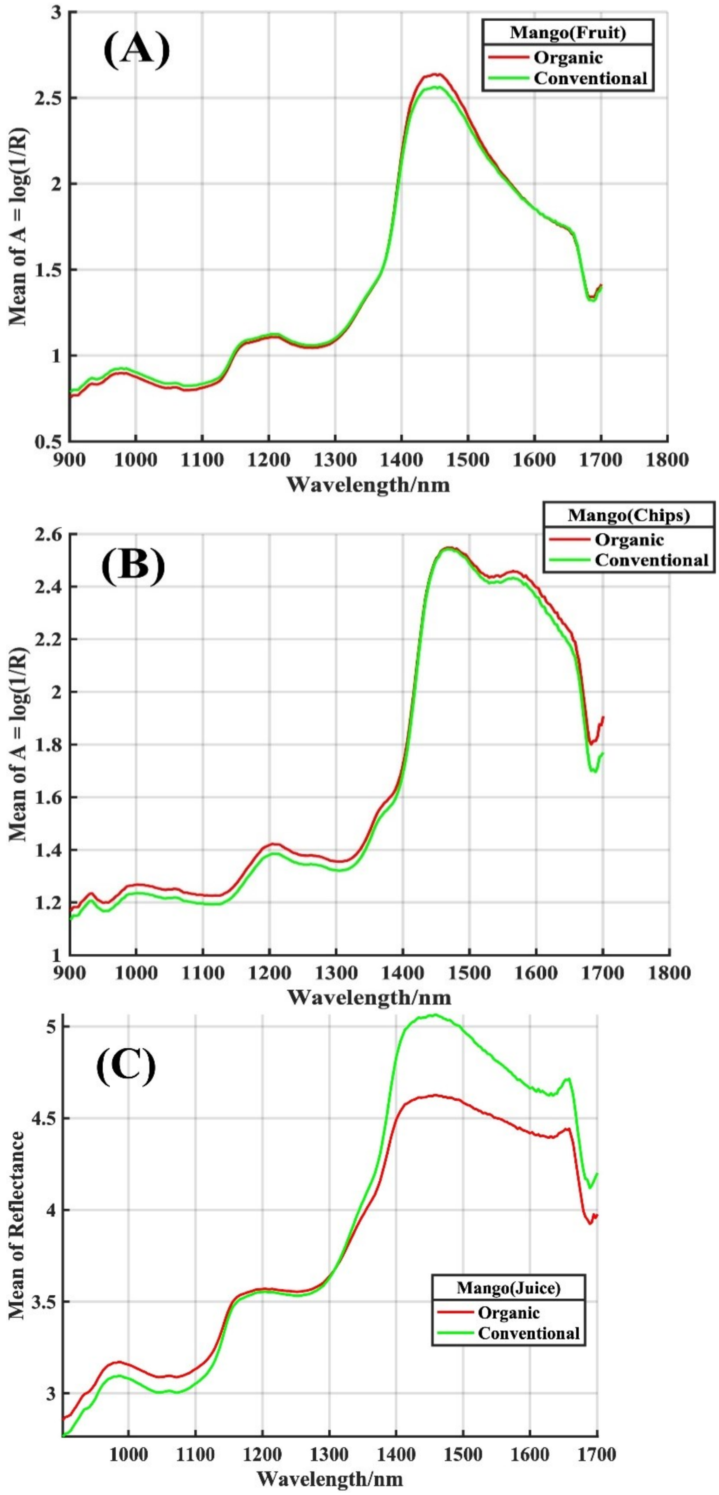


Figure 5.3. Mean Spectral Profile of Organic and Conventional Mango (A) Fruits (B) Chips (c) Juice

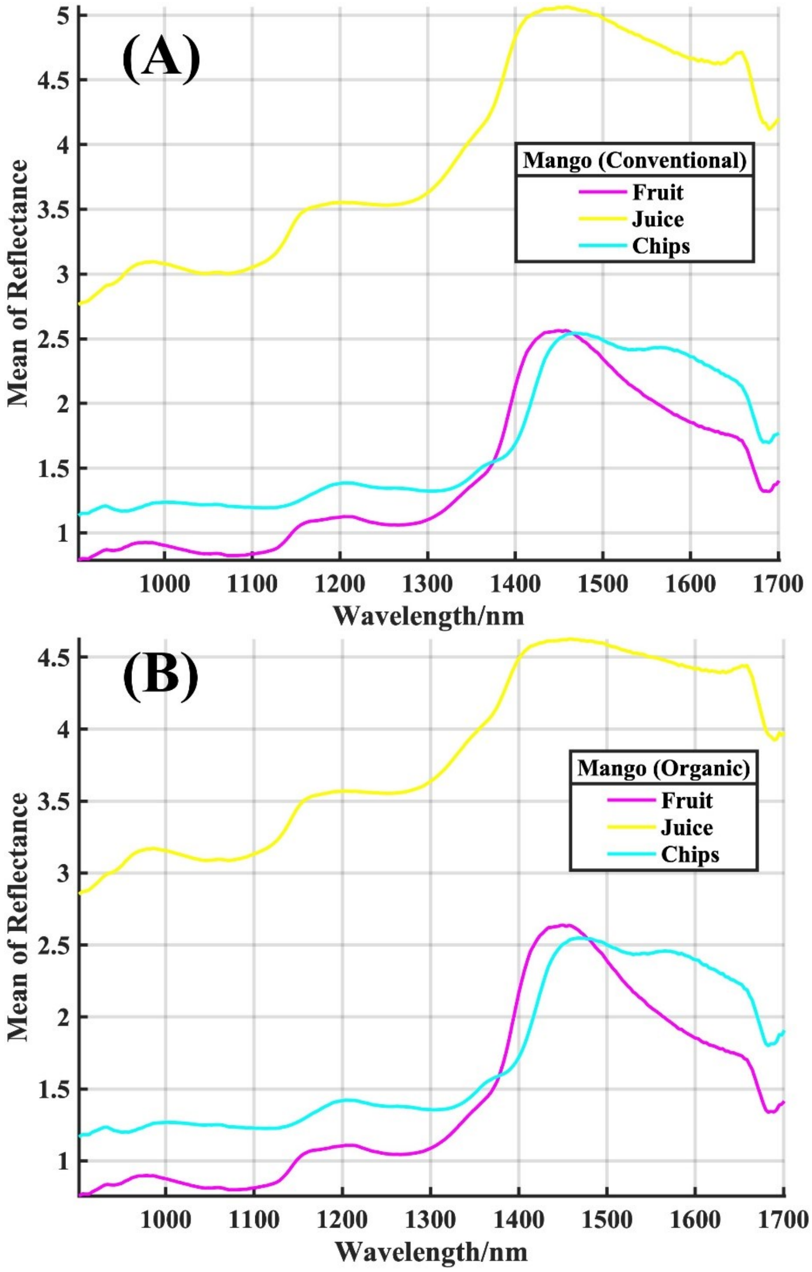
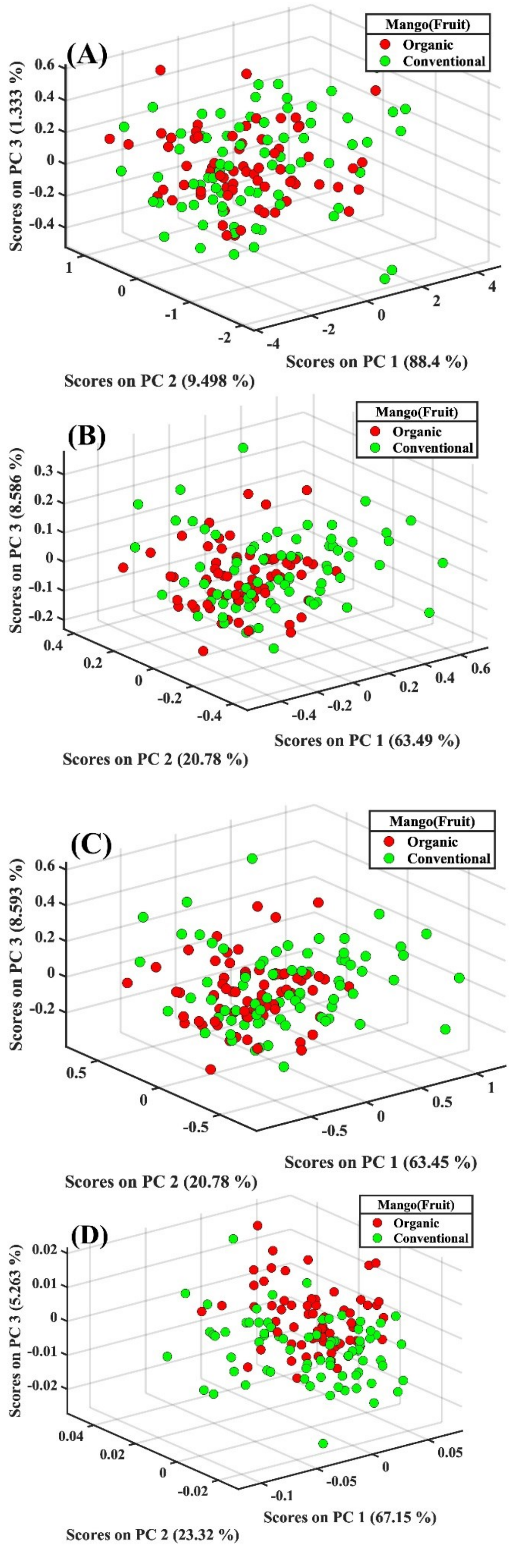


Figure 5.4. Comparison of Mean Spectra Profile of Organic and Conventional Mango Fruits, Chips, and Juice.

5.3.2. Principal Component Analysis

Principal component analysis (PCA) was performed as an unsupervised pattern recognition technique (Teye & Amuah, 2022) to reveal the inherent patterns and classifications within the spectral data of organic and conventional mango products. PCA score plots were developed for mango fruits, chips, and juice, with each subjected to various pre-processing methodologies: Raw, multiplicative scatter correction (MSC), standard normal variate (SNV), first derivative (FD), and second derivative (SD). The PCA score plots corresponding to the pre-processing methodologies of mango fruits are illustrated in Figure 5.5 (A-E). The FD pre-processing exhibited the most proficient outcome in PCA, achieving a variance explanation of 95.73%, succeeded by MSC at 92.86%, SNV at 92.82%, and SD at 92.66%. The three leading principal components (PCs) for FD, as depicted in Figure 5.5 (d), contributed to the maximal total variance (95.73%) of the dataset concerning organic and conventional mango fruits.

The PCA results for the pre-processing methodologies applied to mango chips are presented in Figure 5.6 (A-E). The MSC pre-processing outperformed the other methodologies, with SNV, FD, and SD following in decreasing order. As demonstrated in Figure 5.6 (B), the top three PCs for MSC encapsulated 91.35% of the variation within the dataset concerning mango chips. The PCA outcomes for mango juice regarding the pre-processing methodologies are illustrated in Figure 5.7 (A-E). The FD pre-processing yielded superior results to the other pre-processing techniques, achieving a variance explanation of 96.63%, with MSC at 91.88%, SNV at 91.73%, and SD at 90.09% in descending order. As indicated in Figure 5.7 (D), the top three PCs for FD pre-processing accounted for 96.63% of the variation within the dataset of mango juice.


(A)
Mango(Fruit)
Organic
Conventional
Scores on PC 3 (1.333 %)
Scores on PC 2 (9.498 %)
Scores on PC 1 (88.4 %)
(B)
Mango(Fruit)
Organic
Conventional
Scores on PC 3 (8.586 %)
Scores on PC 2 (20.78 %)
Scores on PC 1 (63.49 %)
(C)
Mango(Fruit)
Organic
Conventional
Scores on PC 3 (8.593 %)
Scores on PC 2 (20.78 %)
Scores on PC 1 (63.45 %)
(D)
Mango(Fruit)
Organic
Conventional
Scores on PC 3 (5.263 %)
Scores on PC 2 (23.32 %)
Scores on PC 1 (67.15 %)

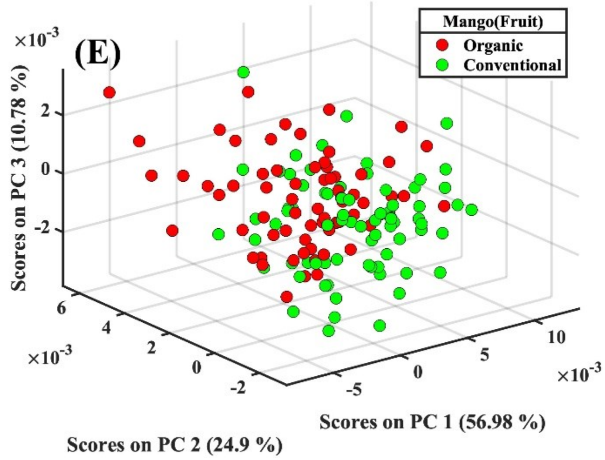
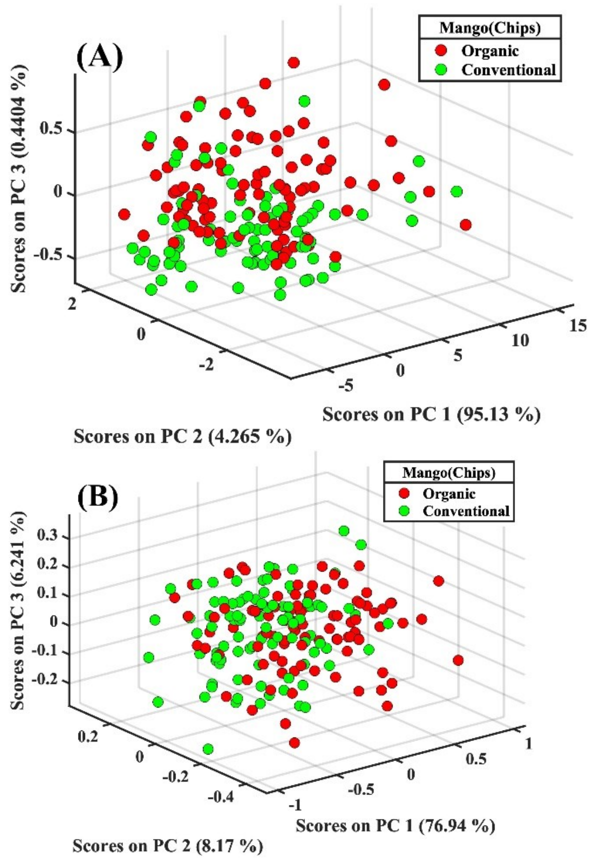


Figure 5.5 PCA Score Plot of the First Three PCs of Organic and Conventional Mango Fruits Pre-Processed – a) RAW b) MSC c) SNV d) FD e) SD.



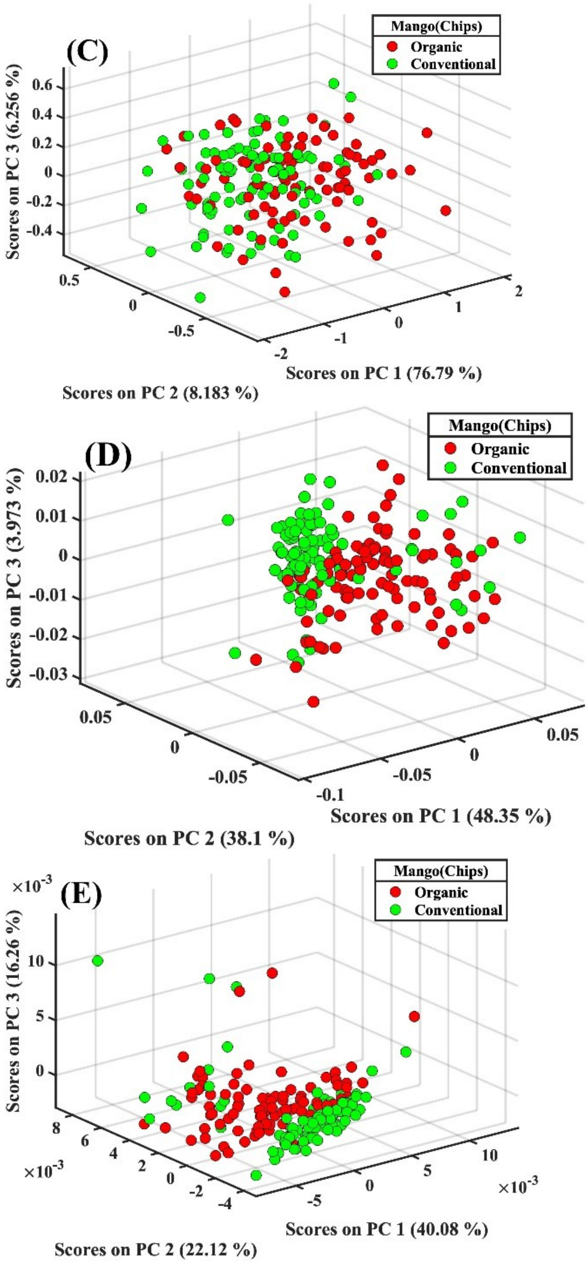
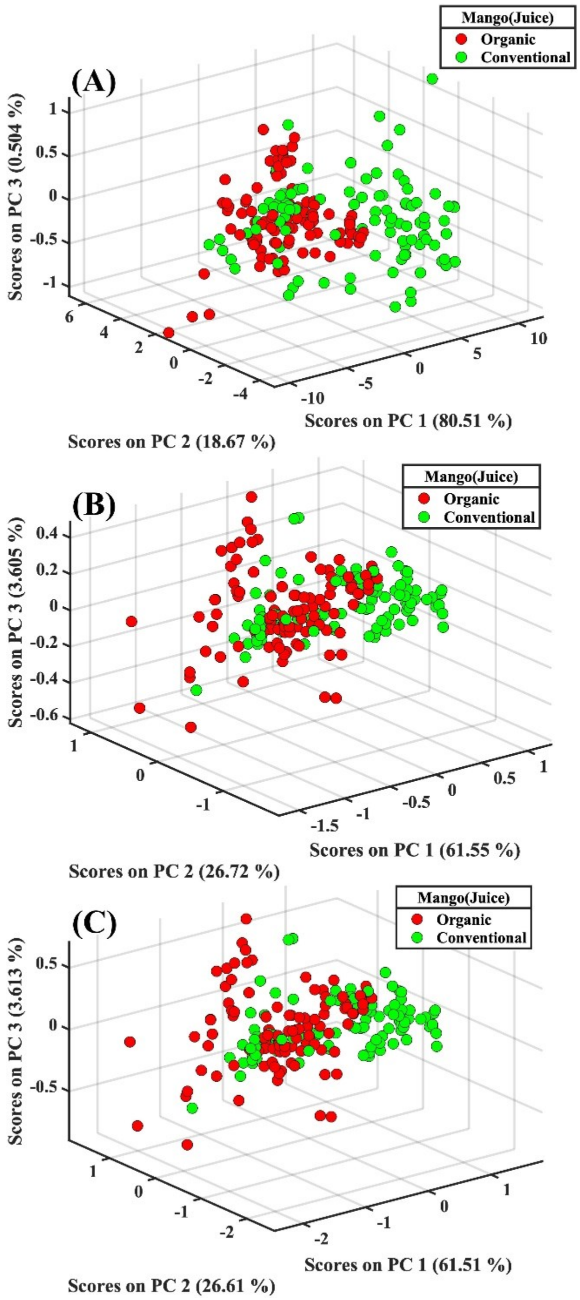


Figure 5.6. PCA Score Plot of the First Three PCs of Organic and Conventional Mango Chips Pre-Processed- a) RAW b) MSC c) SNV d) FD e) SD.


Mango(Juice)
Organic
Conventional
(A)
Scores on PC 3 (0.504 %)
Scores on PC 1 (80.51 %)
Scores on PC 2 (18.67 %)
(B)
Scores on PC 3 (3.605 %)
Scores on PC 1 (61.55 %)
Scores on PC 2 (26.72 %)
(C)
Scores on PC 3 (3.613 %)
Scores on PC 1 (61.51 %)
Scores on PC 2 (26.61 %)

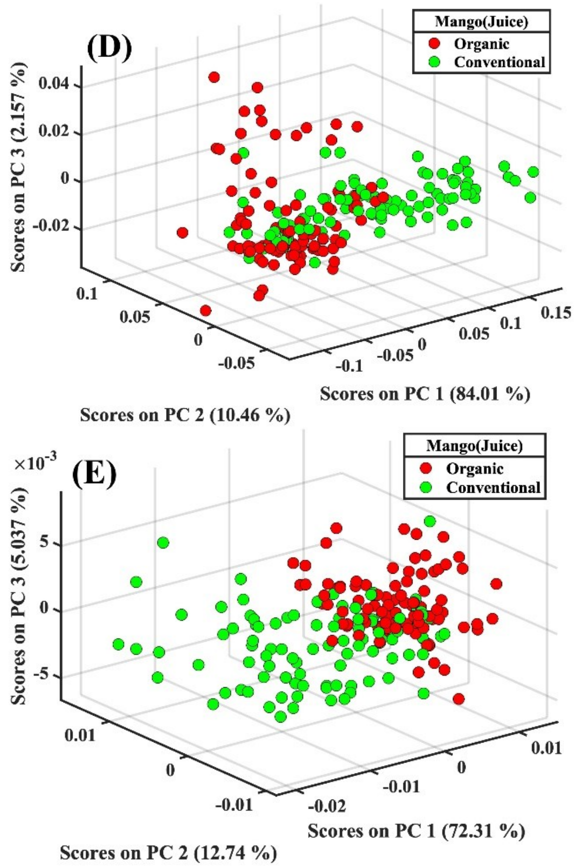


Figure 5.7. PCA Score Plot of the First Three PCs of Organic and Conventional Mango Juice Pre-Processed- a) RAW b) MSC c) SNV d) FD e) SD.

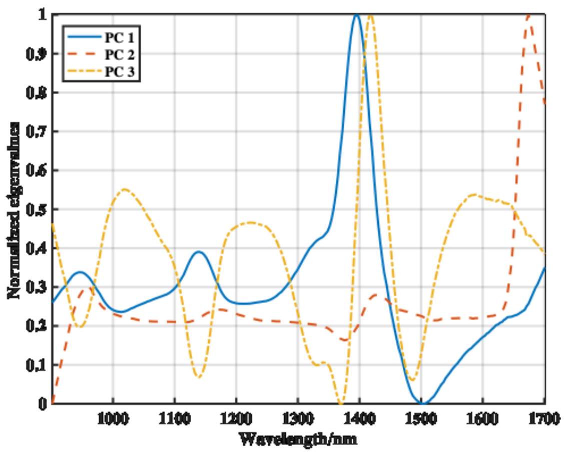


Figure 5.8 (A). Principal Component Loadings with Three Principal Components for Mango Fruits

Figure 5.8 (A) illustrates the principal component loadings corresponding to three principal components, elucidating the clustering trend in the PCA score plot. Notably, the wavelengths at 958 nm and 1416 nm are associated with the presence of water in the fruit (Lu, 2001; Seki, Murakami, Ma, Tsuchikawa, & Inagaki, 2024), whereas the peak at 1588 nm is indicative of substantial absorption attributed to the N-H stretching first overtone (Lee, Jang, Lee, & Kim, 2019). The peak observed at 1610 nm is correlated with the carbohydrate content (Teena, Manickavasagan, Ravikanth, & Jayas, 2014), while the peak at 945 nm is likely representative of the third overtone of CH_2 stretching in sugars (Wu, He, & Feng, 2008). The peak at 1140 nm has been ascribed to the second overtone of C-H (Oliveira, Cruz-Tirado, Roque, Teófilo, & Barbin, 2020).

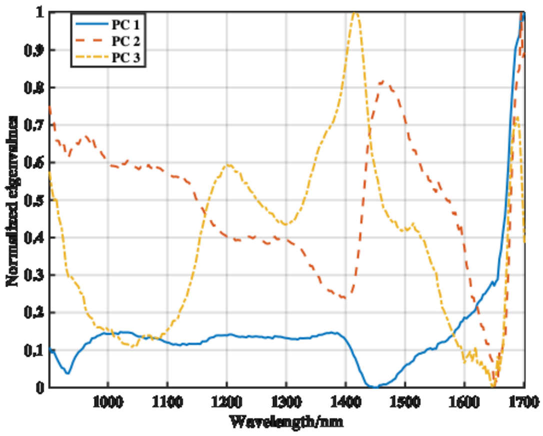


Figure 5.8 (B). Principal Component Loadings with Three Principal Components for Mango Chips

Figure 5.8 (B) summarizes the principal component (PC) loadings for three distinct principal components, elucidating the critical wavelengths that explain the clustering dynamics observed in the PCA score plot. Among these wavelengths, the 909 nm peak is correlated with detecting firmness and glucose levels (Hu, Ma, Liu, Wu, & Ouyang, 2017; Mubarok, Sutari, & Hadiwijaya, 2021), whereas the 962 nm peak indicates O–H and NH_2 stretch overtones (Wedding et al., 2011). The peak observed at 973 nm is a valuable indicator for estimating soluble solids content (Omar, Atan, & MatJafri, 2012). The peaks between 1457 and 1647 nm are associated with the structural constituents of protein, starch, and water (Hu et al., 2017). Finally, the spectral regions spanning 1585 to 1695 nm and 1006 to 1152 nm correspond to the molecular signatures of water, starch, and chitin (Hu et al., 2017).

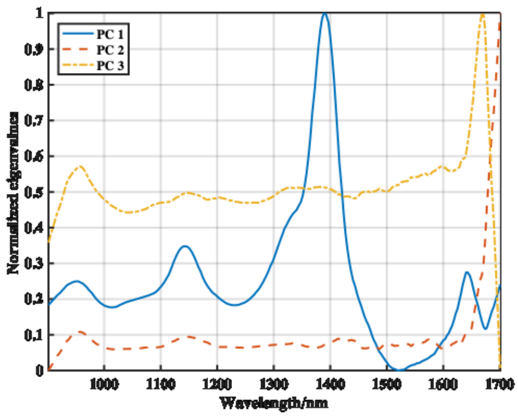


Figure 5.8 (C). Principal Component Loadings with Three Principal Components for Mango Juice

The principal component loadings highlighting three key principal components related to mango juice are illustrated in Figure 5.8(C). Within these spectral wavelengths, 1145 nm is associated with the second overtone of the C-H bond, while the range from 1255 nm to 1344 nm corresponds to the first overtone of the C-H combination. Additionally, the wavelengths between 1504 nm and 1523 nm are associated with the first overtone of the N-H bond (Eldin & Akyar, 2011).

5.3.3. Classification Models for Organic and Conventional Mango Products

The findings pertaining to classifying organic and conventional mango products (including fruits, chips, and juice) are systematically presented in Tables 1, 2, and 3. Multivariate analytical techniques serve as instrumental methodologies for extracting pertinent qualitative or quantitative insights from complex datasets, thus rendering them exceedingly beneficial in food analysis (Rocha, Prado, & Blonder, 2020). Concerning mango fruits (Table 1), all four computational algorithms (LDA, RF, KNN, and PLSDA) attained classification accuracies ranging from 57.50% to 100% on the training datasets. Upon application to the testing datasets, the overall classification efficacy exhibited considerable variability, with the RF model emerging as the most advantageous, achieving a predictive accuracy of 88.76% when utilizing second derivative (SD) pre-processing.

The KNN model demonstrated relatively stable performance in calibration and validation predictions, particularly when employing FD pre-processing (88.03% for training and 86.57% for testing). Notwithstanding the RF model's exemplary training results (100% accuracy), it manifested diminished classification rates on the testing datasets under certain pre-processing methodologies, such as MSC (73.00%). Comparable findings were reported by Hidalgo et al. (2023), where the RF model exhibited outstanding performance on training datasets yet recorded lower classification rates on testing datasets. For mango fruits, sensitivity metrics varied from 53.69% to 100%, while specificity metrics ranged from 53.69% to 100%. The RF model utilizing SD pre-processing demonstrated the highest sensitivity (88.90%) and specificity (88.90%) on the testing dataset, thereby indicating a well-balanced performance in accurately identifying organic and conventional mango fruits.

Regarding the analysis of mango chips (refer to Table 2), the implemented algorithms attained accuracies ranging from 65.77% to 100% across the training samples. The PLSDA model exhibited consistent performance in calibration and validation predictions, particularly when utilizing SD pre-processing, which yielded accuracies of 84.25% for training and 80.20% for testing. For the classification of mango chips, the observed sensitivities varied from 57.94% to 100%, while the specificities likewise demonstrated a range from 57.94% to 100%.

In the analysis of mango juice (as illustrated in Table 3), the algorithms secured accuracies from 73.71% to 100% on the training samples. When evaluated on the test samples, the overall classification efficacy was typically superior to that of fruits and chips, with the Random Forest (RF) model emerging as the most effective, achieving a prediction rate of 87.53% utilizing RAW spectral data. The PLSDA model exhibited exceptional performance, particularly with FD pre-processing, recording accuracies of 89.29% for training and 90.21% for testing. The KNN model maintained consistent performance in calibration and validation predictions when employing MSC and SNV pre-processing techniques, achieving 90.38% for training and 86.08% for testing. For mango juice, the sensitivities ranged from 71.39% to 100%, with the specificities also spanning from 71.39% to 100%. Notably, the PLSDA model utilizing FD pre-processing demonstrated the highest sensitivity (90.17%) and specificity (90.17%) on the test set, reflecting a robust and balanced capacity to accurately identify both organic and conventional mango juice.

Sensitivity, defined as the true positive rate or recall, quantitatively assesses the model's proficiency in accurately identifying organic mango products among all actual organic samples, while specificity, or the true negative rate, assesses the model's ability to identify conventional mango products among all actual conventional samples correctly; this metric is considered optimal when both sensitivity and specificity attain a level of 100% (Lamptey et al., 2023; Rocha et al., 2020; Yegon, 2023). These metrics are fundamental for a comprehensive evaluation of classification models, offering an in-depth understanding of their efficacy in differentiating between organic and conventional mango products (including fruits, chips, and juice). Precision serves as an index within information retrieval, signifying the proportion of relevant items, or predicted presences, that are truly present (Liu, Berry, Dawson, & Pearson, 2005).

The F1-score, representing the harmonic mean of precision and recall (Dalianis & Dalianis, 2018), provides a metric that harmonizes false positives and false negatives. Throughout the analysis of all three mango products, the RF model consistently exhibited the highest precision and F1-scores, particularly when integrated with appropriate pre-processing methodologies (SD for fruits, FD for chips, and RAW for juice). Nonetheless, the PLSDA model demonstrated competitive efficacy, especially in the classification of mango juice. Other researchers have similarly employed accuracy, precision, sensitivity, and F1-score to evaluate the performance of various classification models (Boadu, Teye, Lamptey, Amuah, & Sam-Amoah, 2024).

Table 5.1: Performance of Classification Models in Differentiating Organic Mango Fruits from Conventional Ones using Pre-Processing Techniques with 10-Fold Cross-Validation for both Training and Test Sets

Model	Classification metrics	RAW		MSC		SNV		FD		SD	
		TRAINING SET	TESTING SET	TRAINING SET	TESTING SET	TRAINING SET	TESTING SET	TRAINING SET	TESTING SET	TRAINING SET	TESTING SET
LDA	Accuracy	57.50	53.43	79.34	75.90	79.34	75.90	69.80	68.14	68.62	69.57
	Sensitivity	57.55	53.69	79.10	75.92	79.10	75.92	69.32	68.04	68.17	69.38
	Specificity	57.55	53.69	79.10	75.92	79.10	75.92	69.32	68.04	68.17	69.38
	Precision	57.57	54.08	79.44	77.63	79.44	77.63	70.02	68.74	68.72	70.39
	F1-score	0.58	0.52	0.79	0.75	0.79	0.75	0.69	0.67	0.68	0.69
RF	Accuracy	100.00	65.62	99.92	73.00	100.00	73.71	100.00	86.57	100.00	88.76
	Sensitivity	100.00	65.12	99.92	73.15	100.00	74.08	100.00	86.61	100.00	88.90
	Specificity	100.00	65.12	99.92	73.15	100.00	74.08	100.00	86.61	100.00	88.90
	Precision	100.00	65.56	99.93	73.74	100.00	75.15	100.00	87.94	100.00	89.71
	F1-score	1.00	0.65	1.00	0.72	1.00	0.73	1.00	0.86	1.00	0.89
KNN	Accuracy	66.36	50.10	79.97	72.33	79.97	72.33	88.03	86.57	87.79	84.43
	Sensitivity	66.36	50.45	79.99	72.35	79.99	72.35	88.11	86.70	87.84	84.70
	Specificity	66.36	50.45	79.99	72.35	79.99	72.35	88.11	86.70	87.84	84.70
	Precision	66.33	50.29	79.94	72.93	79.94	72.93	88.03	87.20	87.77	85.77
	F1-score	0.66	0.50	0.80	0.72	0.80	0.72	0.88	0.86	0.88	0.84
PLSDA	Accuracy	70.58	56.95	82.86	75.14	83.88	75.86	81.30	74.43	83.72	79.52
	Sensitivity	70.20	56.93	83.08	75.51	84.16	76.13	81.16	74.29	83.69	79.52
	Specificity	70.20	56.93	83.08	75.51	84.16	76.13	81.16	74.29	83.69	79.52
	Precision	70.68	57.28	83.07	77.63	84.20	78.16	81.33	75.23	83.72	80.78
	F1-score	0.70	0.56	0.83	0.75	0.84	0.75	0.81	0.74	0.84	0.79

Note: LDA (linear discriminant analysis), RF (random forest), KNN (K-nearest neighbors), and PLSDA (partial least squares discriminant analysis), MSC (multiplicative scatter correction), SNV (standard normal variate), FD (first derivative), and SD (second derivative).

Table 5.2: Performance of Classification Models in Differentiating Organic Mango Chips from Conventional Ones using Pre-Processing Techniques with 10-Fold Cross-Validation for both Training and Test Sets

Model	Classification Metrics	RAW		MSC		SNV		FD		SD	
		TRAINING SET	TESTING SET	TRAINING SET	TESTING SET	TRAINING SET	TESTING SET	TRAINING SET	TESTING SET	TRAINING SET	TESTING SET
LDA	Accuracy	75.05	73.27	65.77	66.75	65.89	66.75	76.65	77.02	77.78	78.07
	Sensitivity	75.05	73.28	65.74	66.61	65.87	66.61	76.77	77.17	77.96	78.33
	Specificity	75.05	73.28	65.74	66.61	65.87	66.61	76.77	77.17	77.96	78.33
	Precision	75.05	73.67	65.75	67.24	65.88	67.24	76.92	78.53	78.36	80.11
	F1-score	0.75	0.73	0.66	0.66	0.66	0.66	0.77	0.77	0.78	0.78
RF	Accuracy	100.00	76.43	100.00	65.12	100.00	65.64	100.00	76.93	99.94	77.98
	Sensitivity	100.00	76.28	100.00	65.17	100.00	65.67	100.00	77.06	99.94	78.06
	Specificity	100.00	76.28	100.00	65.17	100.00	65.67	100.00	77.06	99.94	78.06
	Precision	100.00	77.94	100.00	65.67	100.00	66.00	100.00	77.71	99.94	79.25
	F1-score	1.00	0.76	1.00	0.65	1.00	0.65	1.00	0.77	1.00	0.78
KNN	Accuracy	86.34	70.00	82.77	69.94	83.72	67.16	92.87	76.35	88.29	74.33
	Sensitivity	86.21	69.78	82.83	70.06	83.79	67.39	92.84	76.28	88.28	74.28
	Specificity	86.21	69.78	82.83	70.06	83.79	67.39	92.84	76.28	88.28	74.28
	Precision	86.76	70.46	82.86	70.56	83.87	68.43	92.90	76.76	88.38	74.97
	F1-score	0.86	0.70	0.83	0.70	0.84	0.67	0.93	0.76	0.88	0.74
PLSDA	Accuracy	69.34	58.27	77.54	70.47	77.60	70.47	83.42	79.59	84.25	80.20
	Sensitivity	69.19	57.94	77.57	70.44	77.63	70.44	83.33	79.61	84.22	80.22
	Specificity	69.19	57.94	77.57	70.44	77.63	70.44	83.33	79.61	84.22	80.22
	Precision	69.52	59.61	77.60	71.04	77.66	71.04	83.59	81.07	84.27	81.38
	F1-score	0.69	0.57	0.78	0.70	0.78	0.70	0.83	0.79	0.84	0.80

Note: LDA (linear discriminant analysis), RF (random forest), KNN (K-nearest neighbors), and PLSDA (partial least squares Discriminant analysis), MSC (multiplicative scatter correction), SNV (standard normal variate), FD (first derivative), and SD (second derivative).

Table 5.3. Performance of Classification Models in Differentiating Organic Mango Juice from Conventional Ones using Pre-Processing Techniques with 10-Fold Cross-Validation for both Training and Test Sets

Model	Classification Metrics	RAW		MSC		SNV		FD		SD	
		TRAINING SET	TESTING SET	TRAINING SET	TESTING SET	TRAINING SET	TESTING SET	TRAINING SET	TESTING SET	TRAINING SET	TESTING SET
LDA	Accuracy	75.09	74.79	76.34	74.24	76.23	74.76	73.71	72.26	79.95	79.87
	Sensitivity	75.14	74.61	76.40	74.11	76.29	74.61	73.79	72.17	80.07	79.78
	Specificity	75.14	74.61	76.40	74.11	76.29	74.61	73.79	72.17	80.07	79.78
	Precision	75.41	76.46	76.67	75.53	76.58	76.13	74.30	73.47	81.51	82.36
	F1-score	0.75	0.74	0.76	0.74	0.76	0.74	0.74	0.72	0.80	0.79
RF	Accuracy	100.00	87.53	100.00	84.58	100.00	83.03	100.00	78.84	100.00	76.13
	Sensitivity	100.00	87.28	100.00	84.50	100.00	82.89	100.00	78.78	100.00	76.11
	Specificity	100.00	87.28	100.00	84.50	100.00	82.89	100.00	78.78	100.00	76.11
	Precision	100.00	88.66	100.00	85.32	100.00	83.68	100.00	79.54	100.00	76.79
	F1-score	1.00	0.87	1.00	0.84	1.00	0.83	1.00	0.79	1.00	0.76
KNN	Accuracy	86.20	80.39	90.38	86.08	90.38	86.08	86.20	77.79	82.19	71.53
	Sensitivity	86.28	80.44	90.42	86.00	90.42	86.00	86.24	77.83	82.27	71.39
	Specificity	86.28	80.44	90.42	86.00	90.42	86.00	86.24	77.83	82.27	71.39
	Precision	87.22	82.76	90.57	86.56	90.57	86.56	86.51	78.17	83.32	72.50
	F1-score	0.86	0.80	0.90	0.86	0.90	0.86	0.86	0.78	0.82	0.71
PLSDA	Accuracy	84.88	83.97	81.84	79.39	81.73	80.39	89.29	90.21	86.14	85.11
	Sensitivity	84.95	83.78	81.94	79.39	81.84	80.33	89.36	90.17	86.24	85.00
	Specificity	84.95	83.78	81.94	79.39	81.84	80.33	89.36	90.17	86.24	85.00
	Precision	85.60	85.94	83.04	80.88	83.16	82.37	90.04	91.13	87.39	87.60
	F1-score	0.85	0.84	0.82	0.79	0.82	0.80	0.89	0.90	0.86	0.85

Note: LDA (linear discriminant analysis), RF (random forest), KNN (K-nearest neighbors), and PLSDA (partial least squares discriminant analysis), MSC (multiplicative scatter correction), SNV (standard normal variate), FD (first derivative), and SD (second derivative).

5.4. Conclusion

The study successfully demonstrated the application of portable near-infrared spectroscopy for differentiating organic from conventional mango fruits, chips, and juice. By employing various pre-processing techniques and multivariate classification algorithms, the research identified Random Forest (RF) as the most effective model, particularly when paired with Second Derivative (SD) and First Derivative (FD) pre-processing. There is potential for these models to be imported into mobile phones for effective all-around applications. This model can be integrated into mobile phones for versatile and efficient applications. The findings underscore the efficacy of portable NIR devices in nondestructive, rapid authentication of organic products, providing a valuable tool for ensuring product integrity in the food industry.

Conflict of Interest Statement

The authors have declared no conflicts of interest for this article.

Data Availability Statement

The data supporting the findings of this study are included in this article.

Acknowledgments

We express our profound gratitude for the assistance rendered by Gideon Kwesi Nakotey, David Yaro, the CSIR-Food Research Institute, the World Food Programme, and the School of Agriculture at the University of Cape Coast (UCC).

Credit Authorship Contribution Statement:

Conceptualization: FPL, CLYA, VGB; ET, Formal analysis: FPL, CLYA, ET; Investigation: FPL, VGB, ET; Funding acquisition: FPL, ET; Methodology: FPL, CLYA, EEA, VGB, ET; Writing – original draft: FPL, ET; Data curation: FPL, CLYA, EEA, VGB, ET; Funding acquisition: FPL, ET; Resources: FPL, CLYA, VGB, ET; Supervision: CLYA, EEA, ET; Writing – review & editing: FPL, CLYA, EEA, VGB, ET.

CHAPTER SIX

SMART CLASSIFICATION OF ORGANIC AND INORGANIC PINEAPPLE JUICE USING DUAL NIR SPECTROMETERS COMBINED WITH CHEMOMETRIC TECHNIQUES.

Francis Padi Lamptey[a,b,e], Charles L. Y. Amuah[c], Vida Gyimah Boadu[a,d,e], Ernest Ekow Abano[a], *Ernest Teye[a,e]

[a] *University of Cape Coast, College of Agriculture and Natural Sciences, School of Agriculture, Department of Agricultural Engineering, Cape Coast, Ghana*

[b] *Cape Coast Technical University, School of Applied Sciences and Technology, Department of Food Science and Postharvest Technology, Cape Coast, Ghana*

[c] *University of Cape Coast, College of Agriculture and Natural Sciences, School of Physical Sciences, Department of Physics, Laser and Fibre Optics Centre, Cape Coast, Ghana*

[d] *Akenten Appiah-Menka University of Skills Training and Entrepreneurial Development,*

Department of Hospitality and Tourism Education, Kumasi, Ghana

[e] *Africa Centre of Excellence for Food Fraud and Safety, AfriFoodinTegrity Centre, University of Cape Coast, Cape Coast, Ghana*

Tel: 233-243170302

**Email:ernest.teye@ucc.edu.gh*

Published, Applied Food Research (Scopus / Sci, IF = 4.5)

Statement of Contribution of Joint Authorship

Lamptey F. P.: **(Candidate)**

Conceptualised the topic, established methodology, data collection and analysis, preparation of tables and figures, writing and compilation of the original manuscript.

Teye, E.: **(Principal Supervisor)**

Conceptualised the topic, established methodology, supervised and edited the manuscript and co-author of manuscript.

Abano E. E.: **(Co-Supervisor)**

Conceptualised the topic, supervised and edited the manuscript and co-author of manuscript.

Amuah C. L. Y.: **(Research Team member)**

Validation, visualization and data curation and co-author of manuscript

Boadu V.G.: **(Research Team member)**

Validation, visualization and data curation and co-author of manuscript

Abstract

The global demand for organic foods, driven by health benefits and consumer preferences, necessitates reliable methods for distinguishing organic products from their inorganic counterparts. This study investigates the application of dual handheld near infrared (NIR) spectroscopy devices, SCiO and Tellspec, combined with chemometric techniques for the nondestructive differentiation of organic and inorganic pineapple juices. The objective was to establish a rapid and robust method to differentiate organic pineapple juice from inorganic juice using unique spectral data from the two devices. Eighty-four pineapple juice samples were analysed with preprocessing techniques, including mean centering, multiplicative scatter correction, standard normal variate, first derivative, and second derivative applied to the spectral data. Partial least squares discriminant analysis (PLS-DA) was employed for classification, and variable importance in projection (VIP) was used for optimal wavelength selection. The results demonstrated that the Tellspec scanner, particularly with second derivative preprocessing, achieved high accuracy in differentiating organic from inorganic pineapple juice. The fusion of data from both SCiO (740–1070 nm) and Tellspec (900–1700 nm) scanners, without preprocessing, coupled with the PLS-DA model, achieved perfect classification accuracy, sensitivity, and specificity (100%) in both training and testing sets. This study highlights the potential of integrating dual handheld NIR spectroscopy with chemometrics to effectively and accurately classify organic and inorganic pineapple juices. The findings support using these advanced techniques for quality assurance and authentication in the food industry.

Keywords: Consumer preferences, Juice differentiation, NIR spectroscopy, Nondestructive, Partial least squares discriminant analysis

6.1 Introduction

Pineapple is a prominent fruit that is abundantly cultivated in Ghana. However, a substantial amount of it is wasted during periods of maximum productivity despite huge variations in quality and demand (Beausang, Hall, & Toma, 2017). The fruit can undergo processing to create value-added products, including juice, concentrates, jam, candy, wine, dried pineapple, canned pineapple, dehydrated products, and frozen fruits (Hamzah et al., 2021; Chaudhary, Kumar, Singh, Kumar, & Kumar, 2019). This helps to prevent any decrease in the quality or quantity of the food after it has been harvested and extends its shelf life (Saloni, Chauhan, & Tiwari, 2017). The US Food and Drug Administration requires a product to consist entirely of fruit juice to be classified as fruit juice (Heyman et al., 2017). The global juice industry is projected to produce a revenue of US$116.80 billion in 2023, and the market is projected to grow annually at 3.65% (Statista, 2023).

Consumers prefer pineapple juice due to its distinctive fragrance and taste derived from amino acids, amines, phenolic chemicals, and furanone (Shaik & Chakraborty, 2022). Pineapple juice contains various minerals, particularly manganese, amino acids, carbohydrates, vitamins, and polyphenols. The drink is considered functional since it possesses health-promoting features such as anti-inflammatory, anti-atherosclerotic, anti-aging, and other therapeutic capabilities (Khalid, Suleria, & Ahmed, 2016). Fruit juice processing involves selection, extraction, de-aeration, filtration, preservation, and packaging, with or without preservatives, and can be stored neatly in rubber bottles or other neat packaging (Adebayo, Unigbe, & Atanda, 2014). Pineapple juice is available in various types, including single-strength, concentrated, blended, and clear juices. Techniques like pasteurization, ultrafiltration, and freeze drying are used to improve its quality and shelf life. Advanced processing techniques like ultraviolet irradiation and reverse osmosis, sonication also help preserve antioxidant compounds, vitamins, and minerals (Hoque, Talukdar, Roy, Hossain, & Zzaman, 2024; Khalid et al., 2016). Nevertheless, fruit juices are one of the primary food categories that are experiencing a rise in fraudulent incidents (Codex Alimentarius, 2016).

Organic food is gaining popularity among customers due to its claimed health benefits and environmental sustainability compared to inorganically produced food (Gomiero, 2018). This is because of the abundance of cumulative evidence indicating that organic food exhibits reduced levels of pesticide residue, superior taste, healthiness, reduced risk of allergic diseases, and overweight and obesity (Mie et al., 2017; Puska, Kurki, Lähdesmäki, Siltaoja, & Luomala, 2018). Consumers are thus inclined to pay the higher price that organic foods demand (Amuah et al., 2019). The high cost of organic pineapple juice makes it susceptible to food fraud and adulteration. Hence, inorganic pineapple juice is falsely branded as organic to mislead consumers into purchasing it at inflated prices.

The methods currently used to guarantee the integrity and quality of organic products are typically labor-intensive, costly, time-consuming, involve destructive methods, call for highly skilled workers, and are frequently inapplicable in nations with limited resources (Anyidoho et al., 2021). Chemometric analysis using a portable NIR spectrophotometer could be beneficial in distinguishing genuine and authentic organic pineapple juice from inorganic juice. This would provide a quick, nondestructive, and less expensive method for differentiating organic pineapple juice from inorganic ones for quality assurance and control. Ehsani, Yazdanpanah, and Parastar (2023) employed a dual handheld NIR spectrometer, namely Tellspec (with a wavelength range of 900–1700 nm) and Neospectra (with a wavelength range of 1350–2550 nm), as a screening technique to verify the authenticity of orange juice. No studies in the literature utilized classification models to differentiate organic pineapple juice from inorganic ones utilizing SCiO and Tellspec scanners at the time of this research. Other researchers have successfully utilized the SCiO scanner to conduct research on various food products, including mango (Lamptey et al., 2023), chili pepper (Essuman, Teye, Sam-Amoah, & Amuah, 2023), coffee (Boadu, Teye, Amuah, & Sam-Amoah, 2022), rice (Teye & Amuah, 2022; Teye, Amuah, McGrath, & Elliott, 2019), and palm oil (MacArthur, Teye, & Darkwa, 2020). The Tellspec scanner has also been used by other researchers on food and food products such as cocoa (Anyidoho, Teye, & Agbemafle, 2020), lime (Jahani et al., 2020), and pork (Lam et al., 2023). Hence, the objective of this work was to use two handheld NIR spectrometers and chemometric techniques to nondestructively identify organic pineapple juice from inorganic ones.

6.2. Materials and Methods

6.2.1 Pineapple Juice Preparation

For this study, eighty-four (84) sugarloaf pineapple fruits in various stages of maturity (unripe, ripe, and overripe) were purchased from the Ministry of Food and Agriculture registered pineapple producers in the Central Region of Ghana. Fruits were chosen based on their consistent size and colour, while damaged or diseased fruits were rejected. These fruits comprised organic (42 pieces) and inorganic (42 pieces) pineapples. The fruits were then sent to the University of Cape Coast's A. G. Carson Teaching and Research Laboratory. The pineapples were sorted, washed, peeled, and cut into pieces. The cut fruits were squeezed using an electric juicer (Kenwood Excel Juicer, JE850) (Santos et al., 2021). The obtained pineapple juice was meticulously packaged in bottles (Adebayo et al., 2014). After juicing, we obtained 42 bottles (300ml per bottle) of pineapple juice made from organic pineapples and 42 bottles made from inorganic pineapples.

6.2.2 Reference Measurement (TSS/°Brix and pH).

The total soluble solids (TSS) contents were determined using a digital refractometer (model: PAL-1, ◦Brix range of 0–35%; Atago, Tokyo, Japan) according to established protocols (Abarra et al., 2018; Ehsani et al., 2023). The measurements were recorded in degrees Brix, with three replicates taken for each value. The fruit juice was dropped onto the prism of the refractometer to measure TSS. The refractometer was cleaned with distilled water after each measurement. The pH of the scanned pineapple juice was determined using a digital pH meter, and the average readings were recorded three times.

6.2.3 Instrumentation

This research utilized two scanners, namely SCiO (Consumer Physics, Tel Aviv, Israel) and Tellspec (Tellspec Inc., Toronto, Canada), to analyse the composition parameters of various food and food products. The scanners were explicitly employed for scanning juices. Figure 6.1 displays an image depicting the SCiO and Tellspec scanners and an illustration of the NIR spectrum acquisition scheme. Tellspec is a compact and mobile device that utilizes NIR spectroscopy, bioinformatics, learning algorithms, and a mobile application to provide real-time information. This cutting-edge device uses advanced technology to examine food, analyse it at the molecular level, and provide comprehensive information regarding calories, macronutrients, ingredients, and chemical composition. Tellspec utilizes artificial intelligence algorithms and an extensive food database to operate within the 900–1700 nm wavelength range. The device uses photon emission to analyse the reflected photons and build a spectrum revealing the chemical components in the targeted food. The spectrum is transmitted to the cloud for analysis, and the findings, encompassing macronutrients, calories, and allergies, are shown on a smartphone. The Tellspec scan can penetrate up to a depth of 15-20 mm beneath the surface of the food, which is contingent upon the food's translucency to infrared light (Kapse, Kausley, & Rai, 2022; Tellspec, 2020).

The SCiO is a compact NIR spectrometer that analyses food and transmits data on its components to a smartphone. It accurately measures the nutritional content of vegetables, fruits, dairy products, and meat. The sensor comprises a miniaturized infrared spectrometer and micro-optical technology, functioning within the wavelength range of 740 nm to 1070 nm. It is integrated with cloud-based technology. The micro-spectrometer captures and analyses reflected light by separating it into its constituent spectra. The spectrum data is sent to the cloud through the mobile application, where it undergoes analytical processing and is matched against a material database. Following cloud-based data processing, the mobile app displays the results (Consumer Physics, 2020). The examined spectrum unveils the chemical makeup, offering up-to-the-minute quality data for well-informed decision-making (Kapse et al., 2022).

6.2.4 Spectra Acquisition

Each pineapple juice sample was analysed separately utilizing the SCiO and Tellspec scanners (Figure 6.1). Each mango juice spectrum was collected in the reflectance mode using the SCiO and Tellspec devices. The SCiO device had a spectra range of 740nm-1070nm, while the Tellspec device had a 900–1700 nm range. Both devices recorded the spectra data with a resolution of 1nm. The scanners were linked individually to a mobile phone via Bluetooth, and the data was saved in the cloud, which was later downloaded remotely onto a computer. Both scanners were calibrated before the scanning process. The samples were scanned thrice, each in a 100 ml glass beaker, and were rotated 120^0 after each scan as done by others (Lamptey et al., 2023; Teye, Elliott, et al., 2019).

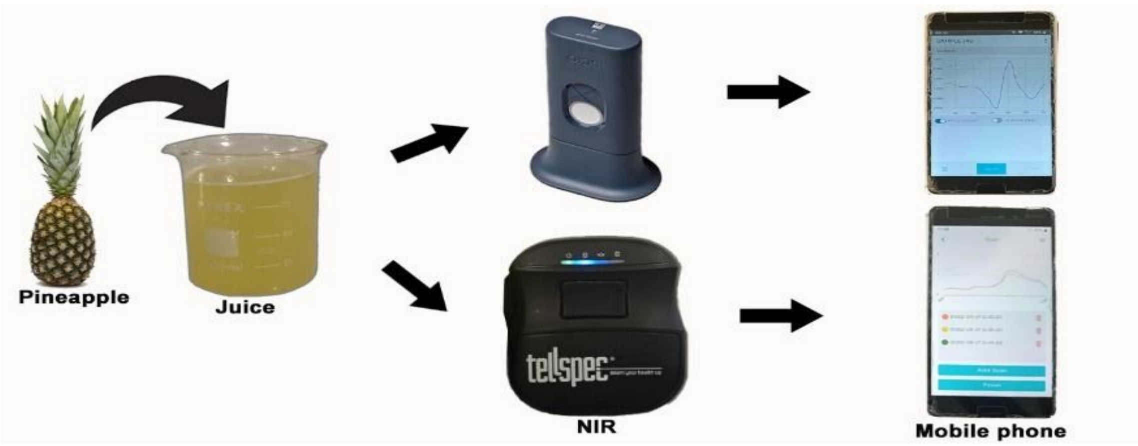


Figure 6.1. Schematic of the Procedure for Obtaining Spectra of Juice (from left, Sugarloaf Pineapple, Pineapple Juice, Dual Scanners: SCiO – up, and Tellspec - Down, and Spectra on a Mobile Phone after Scanning).

6.2.5 Software Device, Data Partitioning, and Spectra Preprocessing

MATLAB (2021a, MathWorks Inc., Natick, MA, USA) with Windows 10 Pro software package was used for computation, chemometric analysis, and generation of figures. The spectroscopic data from 84 organic and inorganic pineapple juice samples underwent preprocessing using appropriate procedures. These datasets were divided into a training set with data from 58 samples and a testing set with data from 26 samples. The training set, representing 69% of the data, was then used to develop models, while the testing set, constituting 31% of the data, was employed to evaluate the models' predictive performance. The raw spectra depicted in Figures 6.2a and 6.3a reveal relevant and undesirable information about pineapple juice, possibly due to light scattering, variability in spectra reproducibility, temperature fluctuations, or background noise (Jha & Garg, 2010). To enhance the chemical information while reducing background noise and enhancing the signal, five preprocessing methods were applied: mean centering (MC), multiplicative scatter correction (MSC), standard normal variate (SNV), first derivative (FD), and second derivative (SD).

Figures 6.2c-g and 6.3c-g display the spectra after preprocessing for the SCiO and Tellspec scanners. Mean centering (MC) is a fundamental preprocessing method that eliminates absolute absorbance values (absolute baseline) and amplifies the absorbance from each wavelength (Agelet & Hurburgh Jr, 2010). MSC and SNV are commonly used techniques to mitigate spectral aberrations caused by scattering. SNV eliminates scatter variation, while the second derivatives correct signal components exhibiting linear variations across different wavelengths (baseline slope). Before applying the

SD, NIR spectra underwent smoothing using the Savitzky-Golay algorithm, significantly improving linearity and rectifying offset in the data. Further details on these preprocessing strategies can be found in studies conducted by other researchers (Agelet & Hurburgh Jr, 2010).

6.2.6 Partial Least-Squares Discriminant Analysis (PLS-DA) for Classification

Classification methods are employed to organize samples based on specific criteria. PLS-DA, a supervised classification technique widely used in NIRs, is valuable for establishing robust connections among initial predictors. It aims to enhance the differentiation between clusters of observations (Kusumaningrum et al., 2018). PLS-DA is a multivariate approach that reduces the number of variables to classify samples, which is particularly useful when it is uncertain if group differences significantly affect overall sample variability (de Sousa Marques, Nicácio, Cidral, de Melo, & de Lima, 2013). It is a variant of traditional PLS regression, generating models for both training and testing. The training set constructs classification models and determines optimal parameters, while the testing set assesses model effectiveness on new samples (Lee & Choung, 2011). Training and testing accuracy are determined by correctly identifying sample numbers (Ambrose, Lohumi, Lee, & Cho, 2016). Data reduction generates latent variables correlating effectively with response variables. The PLS-DA model is represented by the equation provided in the research of other authors (Kandpal, Lohumi, Kim, Kang, & Cho, 2016; Kusumaningrum et al., 2018);

$$Y = XB + E \quad (1)$$

where Y is a matrix of the response variables that are related to the measured sample categories, X is an n × p matrix of the spectral variables for each measured sample category (n is the number of sample objects, and p is the number of variables), B is the matrix regression coefficient for the spectral variables, and E is the matrix of residuals. To find the relationship, the X and Y matrices are decomposed by the latent variables such that:

$$X = TP_{load}^{T} + E \qquad (2)$$

and

$$Y = UQ^{T} + F \qquad (3)$$

Here, T and U are the score matrices, and Pload and Q are the loading matrices.

Spectral data from 84 pineapple juice samples were preprocessed and divided into training (58 samples, two-thirds of the total) and testing (26 samples) subsets. Spectral data are organized with X representing spectral variables and Y indicating sample categories. Predicted scores near 1 denote organic juice, while those near 0 denote inorganic juice. A baseline of ±0.5 is established to assess classification performance. Scores at this baseline (0.5) are ambiguous, indicating uncertainty in classification as organic or inorganic juice.

6.2.7 Optimal Variable Selection

The variable selection process in this study followed other researchers' previous methodologies (Kusumaningrum et al., 2018). Optimal wavelength selection aimed to pinpoint specific wavebands containing crucial information while excluding undesired ones from the spectral dataset. To achieve this, we employed the variable importance in projection (VIP) technique on the training set. We aimed to identify the most relevant variables for constructing a classification model distinguishing organic from inorganic pineapple juice using the fewest possible wavebands.

VIP, an acronym commonly associated with multivariate models like PLS and PLS-DA, aids in understanding influential variables (Andersen & Bro, 2010). Lohumi, Lee, and Cho (2015) defined VIP as a composite metric quantifying a variable's contribution to characterizing two sets of data: the dependent (Y) and independent (X) variables.

The VIP measure (vj) is precisely calculated as:

$$v_J = \sqrt{p \sum_{a=1}^{A} [SS_a (w_{aj}/\|w_a\|^2)] / \sum_{a=1}^{A} SS_a} \qquad (4)$$

Here, p represents the number of variables, SSa is the sum of squares explained by the a-th component, and ∑SSa is the total sum of squares explained by the dependent variable. Thus, vj weights measure each variable's contribution based on the variance explained by each PLS component, where $(\frac{w_{aj}}{\|w_a\|^2})$ represents the importance of the j-th variable (Mehmood, Liland, Snipen, & Sæbø, 2012). The VIP value (vj) provides significant information about variables that contribute to describing dependent variables from independent ones (Andersen & Bro, 2010). Consequently, the VIP technique selects effective wavelengths that differentiate organic juice from inorganic juice. We set a threshold value of 1.25, balancing classification accuracy with the fewest optimal variables. Consequently, we developed the PLS-DA model using wavebands with VIP values exceeding this threshold (Ambrose et al., 2016; Kandpal et al., 2016).

6.2.8 Evaluation of Model Performance

The assessment of model performance was conducted by analyzing sensitivity, specificity, receiver operating characteristic (ROC), and area under curve (AUC) (Pourdarbani et al., 2020). Sensitivity refers to the percentage of correctly categorized samples within a certain class, whereas specificity refers to the percentage of samples from the other class that the model correctly rejects. A class model is considered perfect when it achieves both 100% sensitivity and 100% specificity, as stated by (de Sousa Marques et al., 2013).

6.3. Results And Discussions

6.3.1 Reference Measurement

Fruit juices are beverages that are low in fat, devoid of alcohol and lactose, and are highly sought after by customers due to their nutritional value (Ephrem, Najjar, Charcosset, & Greige-Gerges, 2018). They are naturally abundant in bioactive chemicals that benefit health, can help reduce disease risk, and play a significant role in human nutrition (Ephrem et al., 2018). pH and TSS are essential physicochemical parameters that determine the quality of juice. The pH of fruit juices has a notable impact on the stability of bioactive components (Grobelna, Kalisz, & Kieliszek, 2019), while the Total Soluble Solids (TSS) is a widely employed quality control indicator in the fruit industry (Wojdyło, Teleszko, & Oszmiański, 2014). Table 6.1 displays the pH and TSS range and the mean and standard deviation of the organic and inorganic pineapple juice. The pH of the organic juice ranged from 3.39 to 4.66, while the pH of the inorganic juice ranged from 3.87 – 4.32. The TSS content of the organic pineapple juice ranged from 12- 18.5 (^{0}Brix), while the TSS content of the inorganic pineapple juice ranged from 14.1 – 19.7 (^{0}Brix).

Amuah et al. (2019) recorded a similar TSS range for organic and inorganic pineapple fruits. The variation in these quality indicators could be attributed to disparities in the maturity of the fruits used for the juice (Amuah et al., 2019; Pauziah, Malip, Norhayati, Tham, & Ibrahim, 2012) and the growing conditions under which they were cultivated (Bilalis et al., 2018; Subedi & Walsh, 2011; Wojdyło et al., 2014).

6.3.2 Analysis of NIR Spectra.

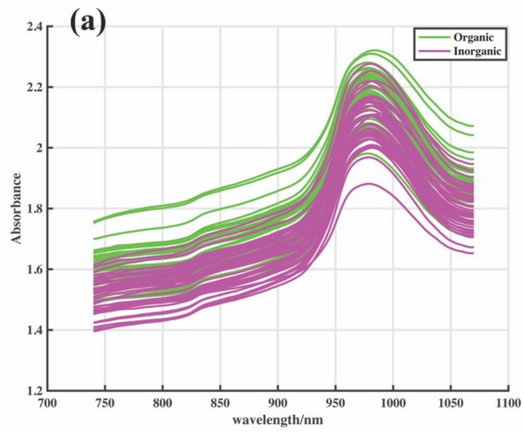
The two NIR spectrometers' raw spectra of the 84 pineapple juice samples are shown in Figures 6.2(a) and 6.3(a). The spectra exhibited a significant range of baseline shifts caused by background information, particle size impact, temperature fluctuation, and noise (Jha & Garg, 2010). The considerable band overlap in the raw spectra posed a challenge in accurately identifying certain bands. Therefore, chemometric preprocessing analysis was utilized to extract the valuable characteristics of the dataset (Kusumaningrum et al., 2018) and construct a dependable model while preserving the similarities and differences among the original observations. The preprocessing methods utilized were MC, MSC, SNV, FD, and SD. Of all the chemometric analyses used, the MSC analysis demonstrated effective categorization when the SCiO scanner was used, and its corresponding spectra can be found in Figure 6.2 (d). This resulted in distinct and easily identifiable clusters, as demonstrated in the average spectral profile depicted in Figure 6.2 (b).

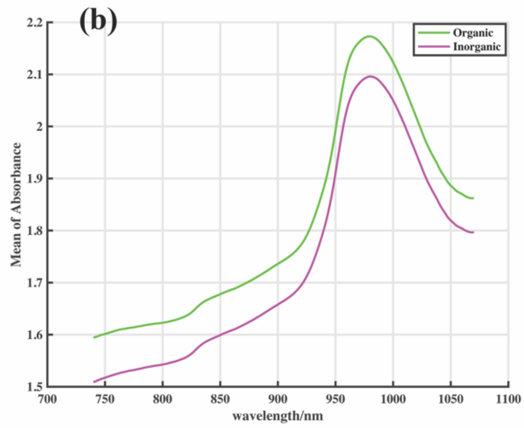
Significant spectral disparities exist between organic and inorganic fruit juice. In Figure 6.2 (b), organic pineapple juice's spectra exhibit more reflectance than inorganic pineapple juice. The NIR spectra consist of wide bands corresponding to overtones and combinations of vibrational modes, such as C-H, O-H, N-H, and S-H chemical bonds (Czarnecki, Morisawa, Futami, & Ozaki, 2015). Sugars and carbohydrates consist mostly of carbon-hydrogen (C-H), oxygen-hydrogen (O-H), carbon-carbon (C-C), and carbon-oxygen (C-O) bonds (Chen et al., 2021). The primary spectral features of water and sugar are represented by the O-H and C-H bands, respectively (Ehsani et al., 2023). The figures demonstrate prominent peaks within the wavelength range of 960 nm to 1050 nm, as seen in Figures 6.2(a) to 6.2(g). The wavelength range mentioned corresponds to the second overtone of the O-H and N-H bonds, characteristic of H_2O, ROH, ArOH (OH bond on the aromatic group), and NH_2 functional groups (Amuah et al., 2019).

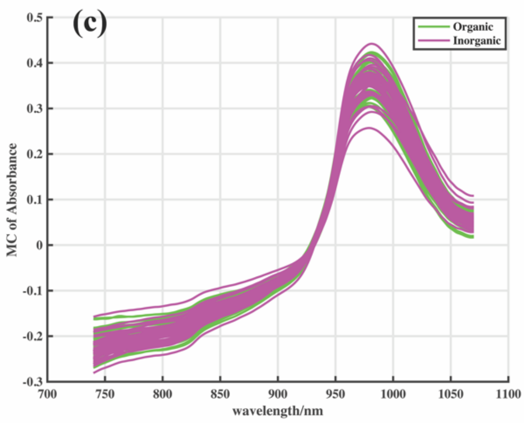
These groups know the primary components of water, glucose, sucrose, and cellulose found in pineapples. After undergoing FD preprocessing, significant peaks were detected at wavelengths of 820nm - 845nm and 950nm. Substantial peaks were observed at wavelengths of 820nm - 845nm, 930nm - 950nm, 965nm - 985nm, and 1020nm - 1035nm. The observed peaks correspond to the third overtone of ArOH, H_2O, RNH_2, ROH, CH, and CH_2 (Ehsani et al., 2023; Lamptey et al., 2023). However, a distinct division was observed when the MSC preprocessing technique was employed among all the preprocessing methods. These results were obtained using the SCiO scanner.

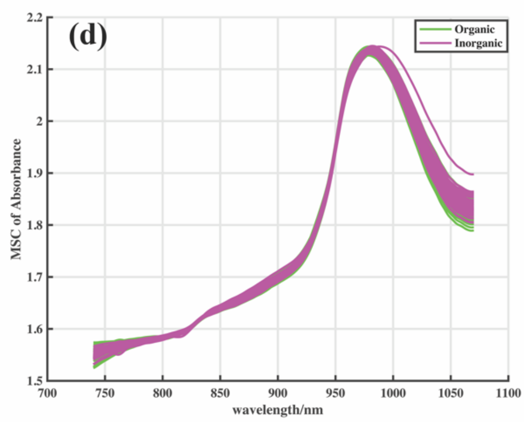
Nevertheless, when the Tellspec scanner was employed, significant peaks were detected at wavelengths of 920nm – 1020nm, 1120nm – 1220nm, 1400nm – 1500nm, and 1610-1630nm, after undergoing preprocessing with MC, as depicted in Figure 6.3 (c). The primary peaks found after preprocessing with Multiplicative Scatter Correction (MSC) were at wavelengths of 920nm – 1020nm, 1120nm - 1220nm, 1400nm – 1500nm, and 1600nm -1650nm, as depicted in Figure 6.3 (d). The prominent peaks detected after applying the SNV preprocessing technique were found within the wavelength ranges of 920nm – 1020nm, 1120nm – 1220nm, 1400nm - 1500nm, and 1600nm – 1680nm, as illustrated in Figure 6.3 (e). The significant peaks detected after applying the FD to the preprocessed data were found at wavelengths of 920nm – 940nm, 1110nm – 1180nm, 1300nm – 1410nm, and 1620nm -1630nm, as illustrated in Figure 6.3 (f).

The SD preprocessing strategy that achieved the most effective spectral smoothing and resulted in good classification exhibited peaks at specific wavelengths: 1110nm - 1125nm, 1280nm - 1320nm, 1370nm - 1380nm, and 1620nm - 1640nm, as illustrated in Figure 6.3 (g). This resulted in distinct and easily identifiable clusters, as demonstrated in the average spectral profile depicted in Figure 6.3 (b). These bands correspond to the following functional groups: CH_2, ArOH, H_20, third overtones, CH_3, CH_2, CH, ArOH, ROH, RNH_2, second overtones, and ArCH, CH_3, CH_2 first overtone (Anyidoho et al., 2021; Ehsani et al., 2023). An observable absorption band at approximately 1440 nm could be linked to the first overtone of starch, moisture, and sugars (Anyidoho et al., 2021).


(a)
Organic
Inorganic
Absorbance
wavelength/nm


(b)
Organic
Inorganic
Mean of Absorbance
wavelength/nm


(c)
Organic
Inorganic
MC of Absorbance
wavelength/nm


(d)
Organic
Inorganic
MSC of Absorbance
wavelength/nm

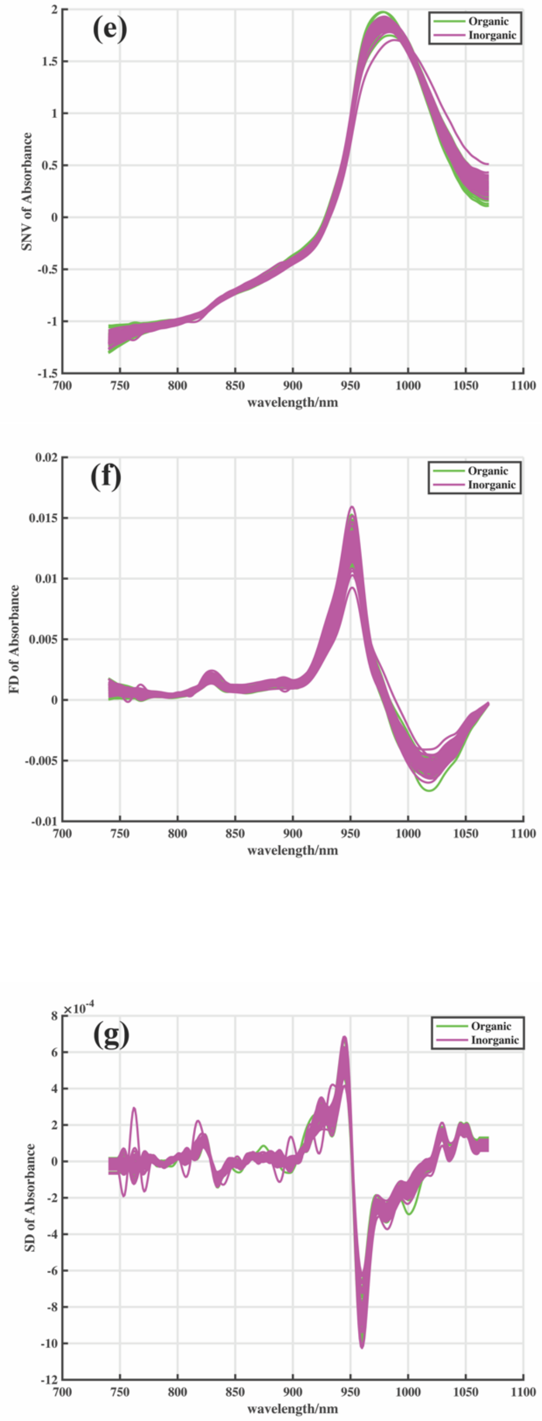
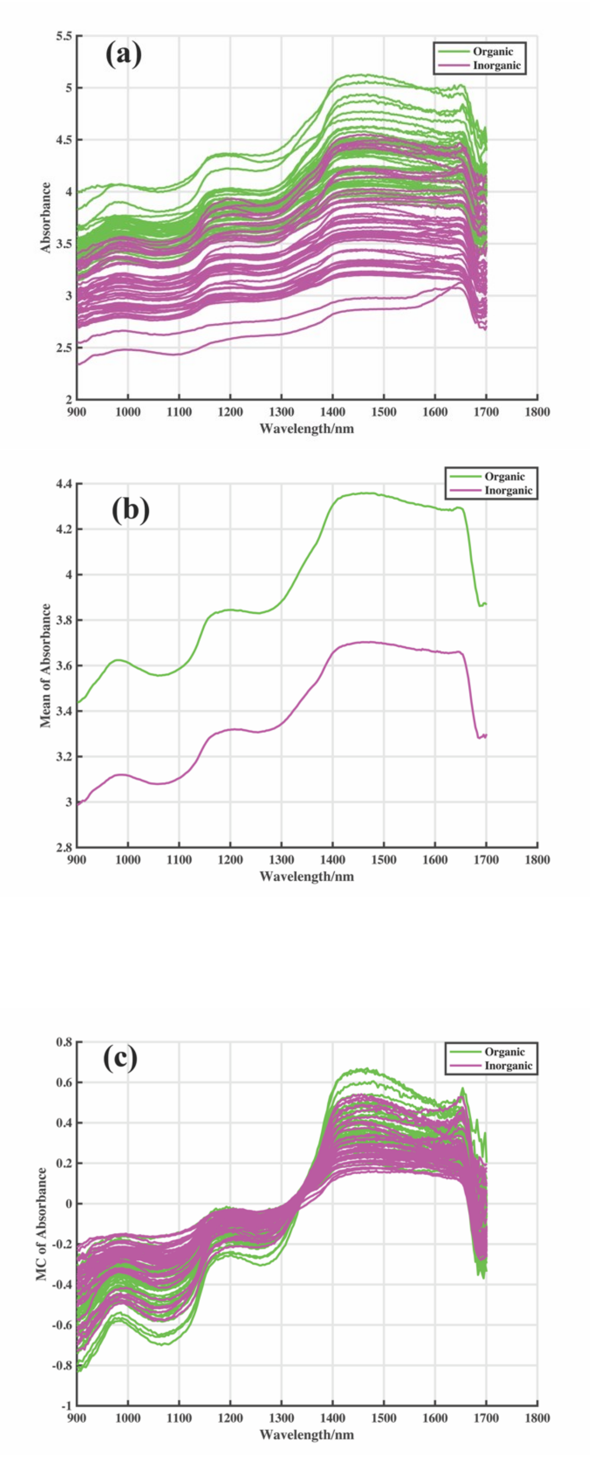
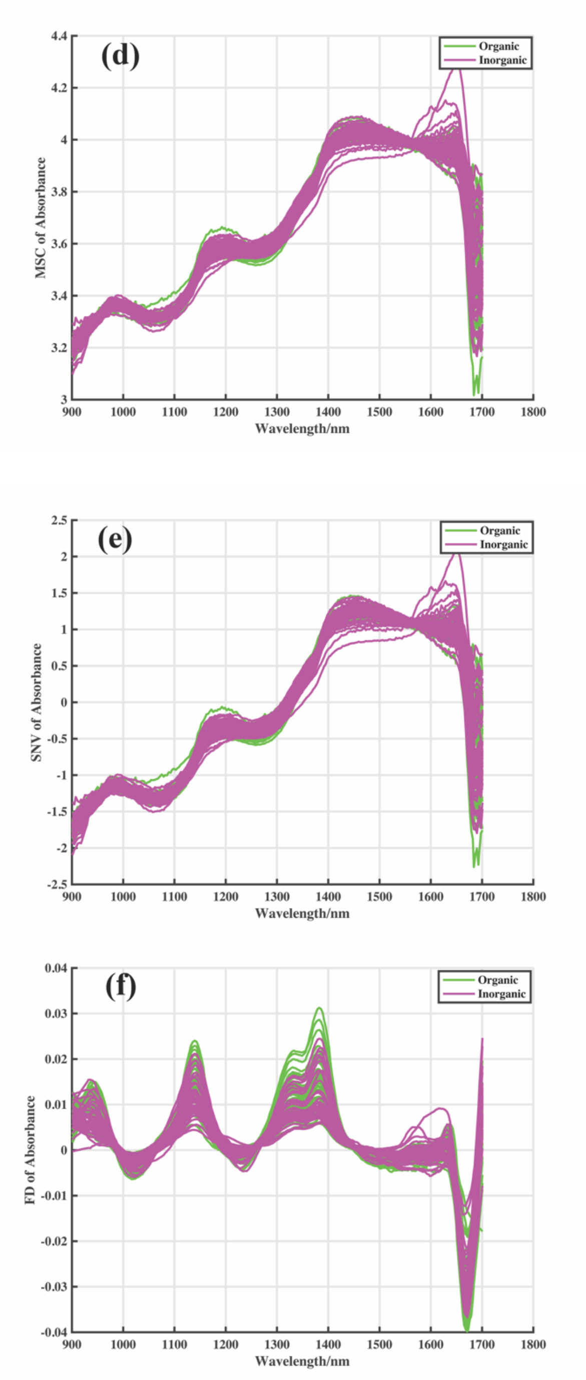


Figure 6.2. Spectra of Pineapple Juice Based on Handheld NIR SCiO ^{TM} (740 - 1070 nm): (a) Raw, (b) Mean, (c) MC (d) MSC, (e) SNV, (f) FD, and (g) SD.


(a)
Organic
Inorganic
Absorbance
Wavelength/nm
(b)
Organic
Inorganic
Mean of Absorbance
Wavelength/nm
(c)
Organic
Inorganic
MC of Absorbance
Wavelength/nm


(d)
Organic
Inorganic
MSC of Absorbance
Wavelength/nm
(e)
Organic
Inorganic
SNV of Absorbance
Wavelength/nm
(f)
Organic
Inorganic
FD of Absorbance
Wavelength/nm

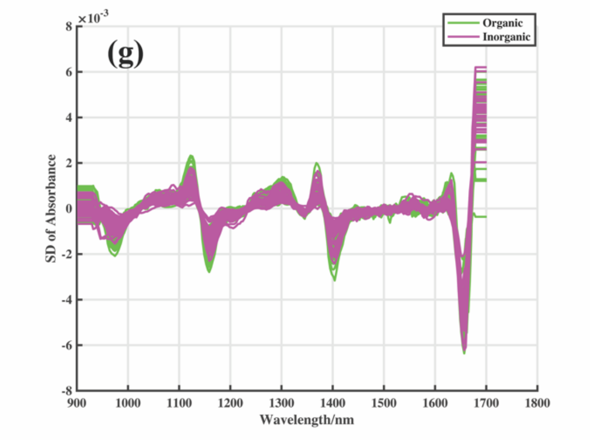


Figure 6.3. Spectra of Pineapple Juice Based on Handheld NIR Tellspec (900–1700 nm): (a) Raw, (b) Mean, (c) MC, (d) MSC, (e) SNV, (f) FD, and (g) SD.

6.3.3 Classification using PLS-DA Model

The PLS-DA model was utilized, leveraging the entire wavelength range to formulate classification equations for unidentified samples in real-world scenarios. This method involved supervised learning, where prior knowledge of class membership was employed to guide the analysis (Kusumaningrum et al., 2018). This study established two distinct categories, organic and inorganic fruit juice, for classification purposes. The complete dataset was partitioned into training and testing sets. Depending on the preprocessing method, the PLS-DA models classified the pineapple fruit juices with varying accuracies. In NIR spectral analysis, preprocessing is essential to adjust the original variables, ensuring they meet the dataset's general assumptions. Suitable preprocessing techniques can reduce baseline shifts and non-linear effects, enhancing the classification models (Kusumaningrum et al., 2018). Consequently, we applied the PLS-DA model with different preprocessing methods to evaluate how each technique, algorithm, and dataset influenced the classification model's accuracy. The classification accuracies from the PLS-DA training and testing datasets are presented in Tables 6.2, 6.3, and 6.4.

The spectral profiles of the organic and inorganic pineapple juices were accurately categorized in all preprocessing model sets, utilizing the complete spectral range. The study accurately classified organic and inorganic pineapple juices using spectral profiles, achieving training accuracy ranging from 70.4% to 87.0% when scanned with the SCiO scanner and 86.0% to 96.0% when scanned with the Tellspec scanner. The SD preprocessing set had the highest accuracy when scanned by both scanners, while the MC processing set demonstrated the lowest accuracy. Testing accuracy ranged from 69.6% to 87.0% when analysed using the SCiO scanner and from 90.0% to 100.0% when analysed using the Tellspec scanner. The FD preprocessing set was most accurate when scanned using the SCiO scanner (Figure 6.4).

In contrast, the SD preprocessing set achieved the highest accuracy when scanned using the Tellspec scanner (Figure 6.5). Samples scanned with the SCiO scanner and preprocessed with FD outperformed the other preprocessing methods, while samples scanned with the Tellspec scanner and preprocessed with SD showed the best performance. The unprocessed absorbance data performed better than those that underwent preprocessing techniques (Figure 6.6), demonstrating that the handheld NIR approach can effectively and accurately distinguish organic pineapple juice from inorganic ones.

Figures 6.7, 6.8, and 6.9 show the ROC curve for the pineapple juice preprocessed with FD for SCiO, SD for Tellspec, and raw (no preprocessing) for the combination of both scanners, respectively. The evaluation of various models included sensitivity and specificity values for the training and testing sets and the area under the ROC curve. Sensitivity and specificity values were determined to optimize their sum. High specificity is often preferred in diagnostic models to reduce false positives (de Sousa Marques et al., 2013). The study utilized the SCiO scanner for PLS-DA models with first derivative spectra for organic and inorganic pineapple juice and the Tellspec scanner for PLS-DA models with second derivative spectra.

In the case where the SCiO scanner was used, it achieved sensitivity and specificity values of 0.74 and 0.85 for training and testing sets, respectively. When the Tellspec scanner was used, it achieved sensitivity and specificity values of 0.92 and 1.0 for training and testing sets. Employing NIR spectra based on combining both scanners without preprocessing resulted in 1.0 for training and 1.0 for testing sets. Ehsani et al. (2023) also achieved perfect sensitivity and specificity of 100% in both training and testing sets when they combined results from two spectrometers (Tellspec, 900-1700 nm, and Neospectra, 1350-2550 nm) in their study. The findings of the PLS-DA models indicate the viability of employing chemometric techniques for the differentiation of organic and inorganic pineapple juice.

Table 6.1: Reference Measurement of pH and TSS

	Subset	Range	Mean	Std
Organic	pH	3.39 - 4.66	4.14	0.25
	TSS	12- 18.5	16.66	1.44
Inorganic	pH	3.87 – 4.32	4.08	0.11
	TSS	14.1 – 19.7	17.95	1.03

Note: TSS (Total Soluble Solids, and Std (Standard Deviation).

Table 6.2: Training and Testing Results of the PLS-DA Model Using Spectral Data from the Handheld NIR Spectrometer (SCIO^{TM}) with Various Preprocessing Methods.

	Training set (n = 58)				Testing set (n = 26)			
Preprocessing	Accuracy	Sensitivity	Specificity	AUC	Accuracy	Sensitivity	Specificity	AUC
RAW	0.741	0.741	0.741	0.875	0.696	0.636	0.750	0.822
MC	0.704	0.667	0.741	0.759	0.783	0.812	0.750	0.833
MSC	0.815	0.741	0.889	0.47	0.826	0.909	0.750	0.864
SNV	0.796	0.704	0.889	0.842	0.826	0.909	0.750	0.852
FD	0.796	0.740	0.852	0.870	0.870	0.909	0.883	0.883
SD	0.870	0.852	0.889	0.966	0.565	0.455	0.667	0.659

Note: AUC (Area Under Curve), MC (Mean Centering), MSC (Multiplicative Scatter Correction), SNV (Standard Normal Variate), FD (First Derivative), and SD (Second Derivative).

Table 6.3: Training and Testing Results of the PLS-DA Model using Spectral Data from the Handheld NIR Spectrometer (Tellspec) with Various Preprocessing Methods.

	Training set (n = 58)				Testing set (n = 26)			
Preprocessing	Accuracy	Sensitivity	Specificity	AUC	Accuracy	Sensitivity	Specificity	AUC
RAW	0.940	0.920	0.960	0.998	0.900	0.800	1.000	0.995
MC	0.860	0.880	0.840	0.963	0.950	0.900	1.000	1.000
MSC	0.900	0.800	1.000	0.970	0.950	0.900	1.000	0.925
SNV	0.920	0.840	1.000	0.970	0.950	0.900	1.000	0.925
FD	0.960	0.920	1.000	0.994	0.950	0.900	1.0000	0.9400
SD	0.960	0.920	1.000	0.997	1.000	1.000	1.000	1.000

Note: AUC (area under curve), MC (mean centering), MSC (multiplicative scatter correction), SNV (standard normal variate), FD (first derivative), and SD (second derivative).

Table 6.4: Training and Testing Results of the PLS-DA Model using Combined Spectral Data from Handheld NIR Spectrometers (Tellspec and SCiO^{TM}) for Classification.

	Training set (n = 58)				Testing set (n = 26)			
Preprocessing	Accuracy	Sensitivity	Specificity	AUC	Accuracy	Sensitivity	Specificity	AUC
RAW	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
MC	0.983	1.000	0.966	1.000	0.846	0.923	0.769	0.959
MSC	0.914	0.931	0.897	0.981	0.885	0.846	0.923	0.976
SNV	0.931	0.931	0.931	0.988	0.885	0.846	0.923	0.970
FD	0.966	0.966	0.966	0.998	1.000	1.000	1.000	1.000
SD	1.000	1.000	1.000	1.000	0.885	0.846	0.923	0.979

Note: AUC (area under curve), MC (mean centering), MSC (multiplicative scatter correction), SNV (standard normal variate), FD (first derivative), and SD (second derivative).

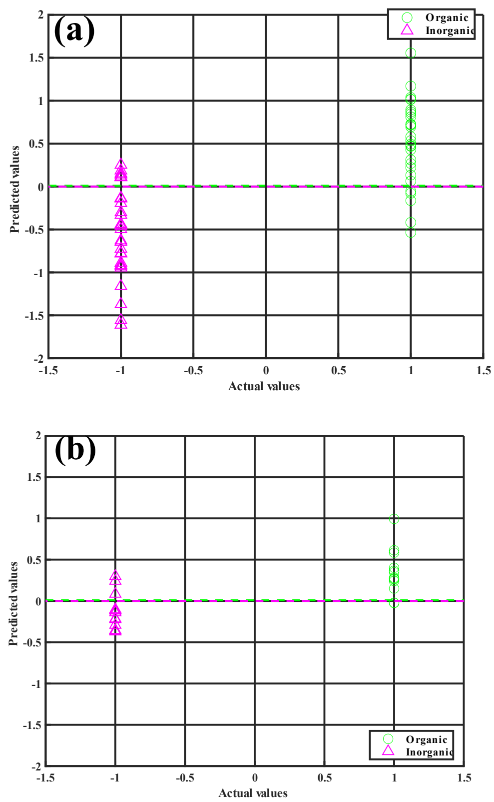


Figure 6.4. Classification of Organic and Inorganic Pineapple Juice Utilizing PLS-DA with FD Preprocessing Based on Handheld SCiO ^{TM} (740 - 1070 nm): (a) Training and (b) Testing.

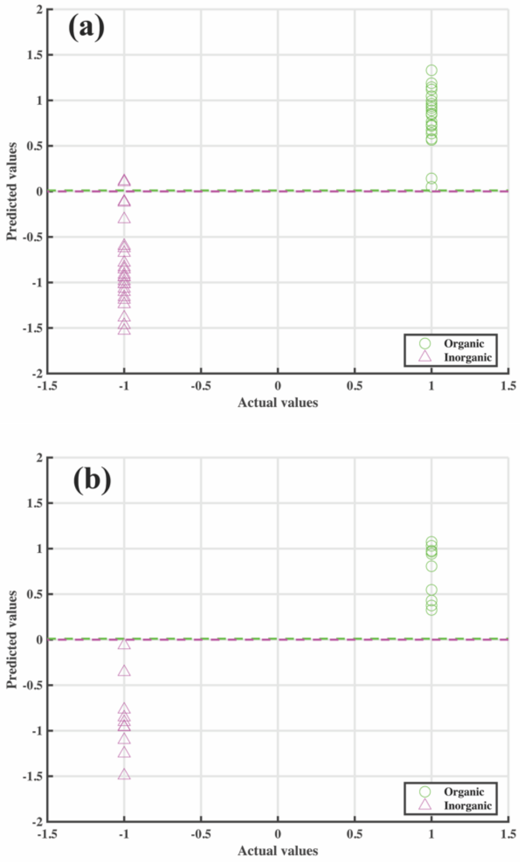


Figure 6.5. Classification of Organic and Inorganic Pineapple Juice Utilizing PLS-DA with SD Preprocessing Based on Tellspec: (a) Training (b) Testing.

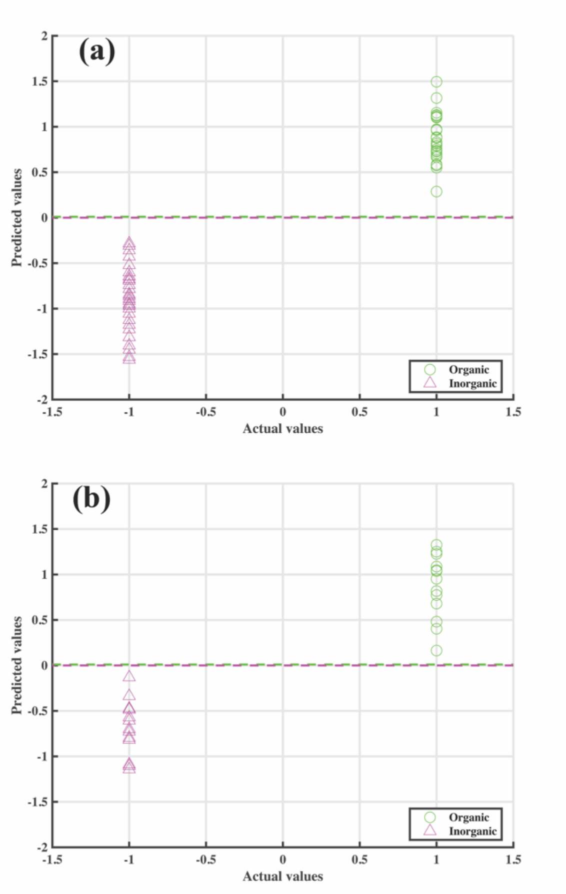


Figure 6.6. Classification Results for Organic and Inorganic Pineapple Juice Utilizing PLS-DA with RAW Preprocessing Based on Combining Handheld Tellspec and SCiO^{TM}: (a) Training and (b) Testing.

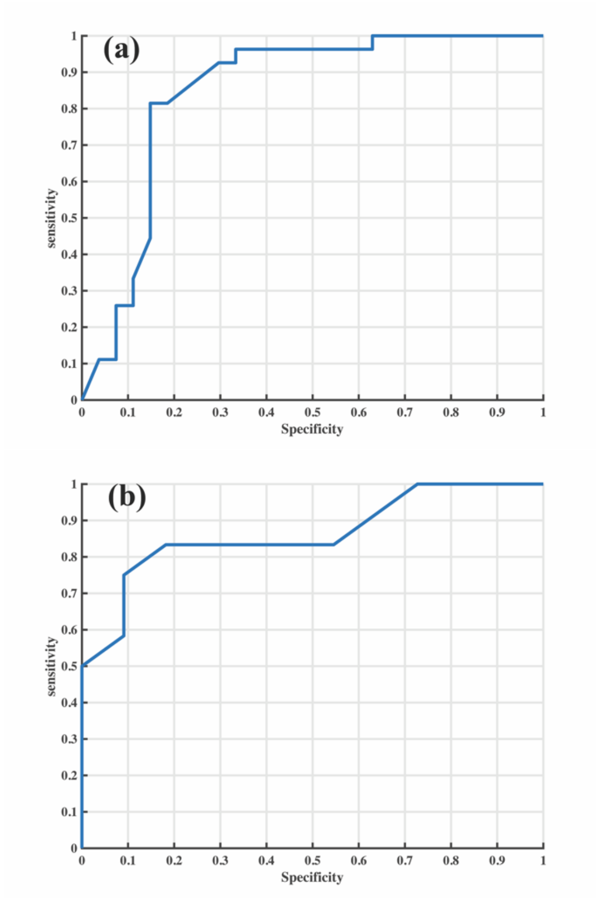
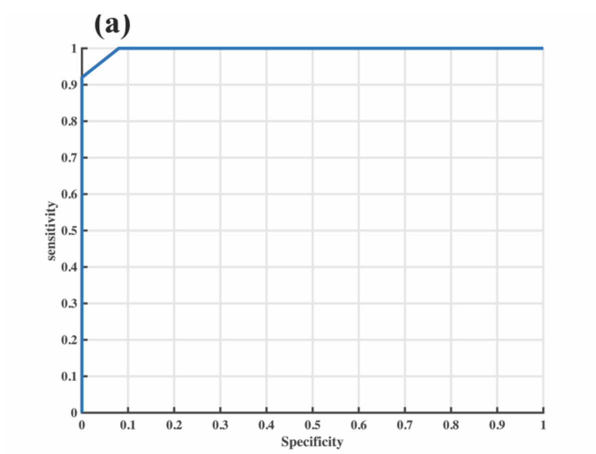


Figure 6.7. ROC Curve for Pineapple Juice with FD Preprocessing of NIR Spectra Based on Handheld NIR SCiO ^{TM} (740 - 1070 nm (a) Training and (b) Testing.



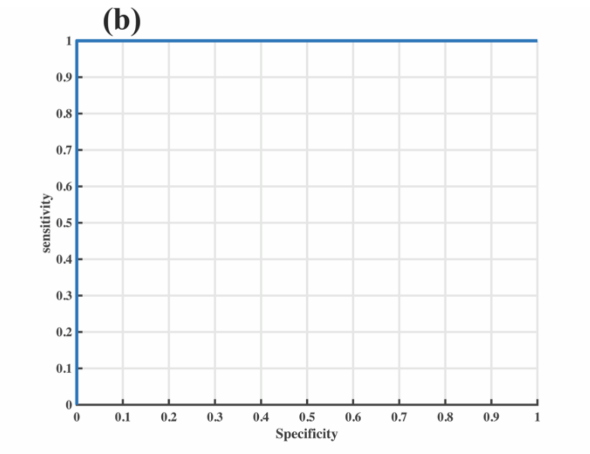


Figure 6.8. ROC Curve for Pineapple Juice with SD Preprocessing of NIR Spectra Based on Tellspec: (a) Training (b) Testing.

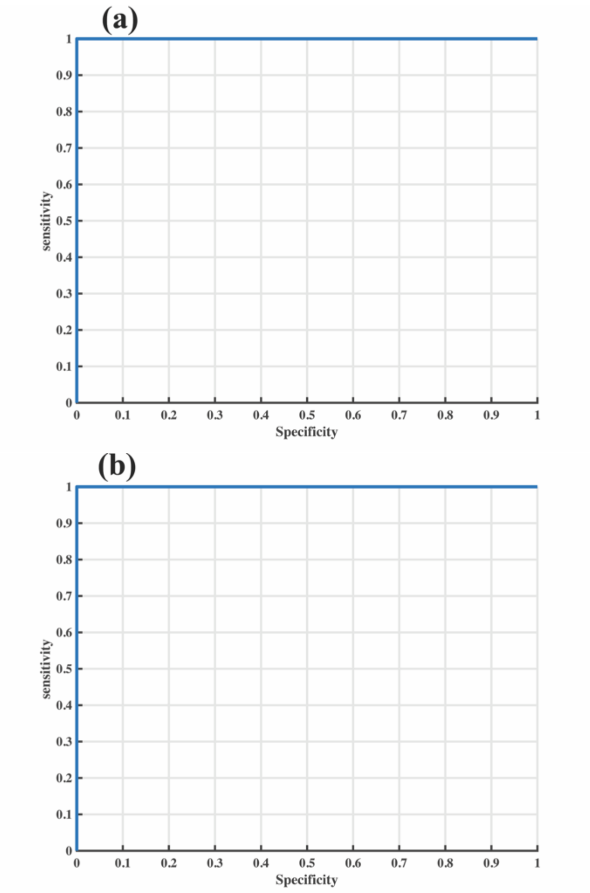


Figure 6.9. ROC Curve for Pineapple Juice with Raw Preprocessing of NIR Spectra Based on the Combination of Handheld Tellspec and SCiO^{TM}: (a) Training and (b) Testing.

6.3.4 Variable Importance Projection Model

The VIP is a comprehensive metric quantifying a variable's contribution to characterizing independent and dependent datasets (Peng, Shi, Song, Chen, & Gao, 2014). The effectiveness of the PLS-VIP technique can depend on the selected cut-off value. Nevertheless, the "greater-than-one rule" frequently identifies relevant predictors. In this study, a threshold of 1.25 was applied (Wang, He, & Wang, 2015). The peaks observed in the VIP score plot offer valuable insights into the organic and inorganic components that contribute to the unique characteristics of pineapple juice. Subsequent peaks were selected after analyzing multiple NIR spectra.

Upon utilizing the SCiO scanner, the peaks chosen were 804 and 964 nm, corresponding to the second overtone O-H stretching vibrations associated with water and the third overtone of C-H functional group, respectively (Hao, Wang, & Zhang, 2021; Siedliska, Baranowski, Zubik, Mazurek, & Sosnowska, 2018) (Figure 6.10). Upon utilizing the Tellspec scanner, the peaks at the wavelengths of 971 nm, 1013 nm, 1142 nm, 1192 nm, and 1671 nm were specifically chosen (Figure 6.11). These peaks correspond to the N-H and C-H regions corresponding to the second overtone. Various substances, including water, sugars, carbohydrates, organic acids, polyphenolic compounds, certain vitamins, and specific amino acids, can all contribute to this absorption process (Włodarska, Szulc, Khmelinskii, & Sikorska, 2019). The peak wavelengths of 810 nm, 827 nm, 852 nm, 868 nm, 1013 nm, 1142 nm, and 1218 nm were chosen when combined with the SCiO and Tellspec scanners (Figure 6.12). This peak corresponds to the third overtone area of N-

H, the second overtone region of N-H, the second overtone region of C-H, and the absorption of theaflavins (Panigrahi, Bhol, & Das, 2016; Wu et al., 2012).

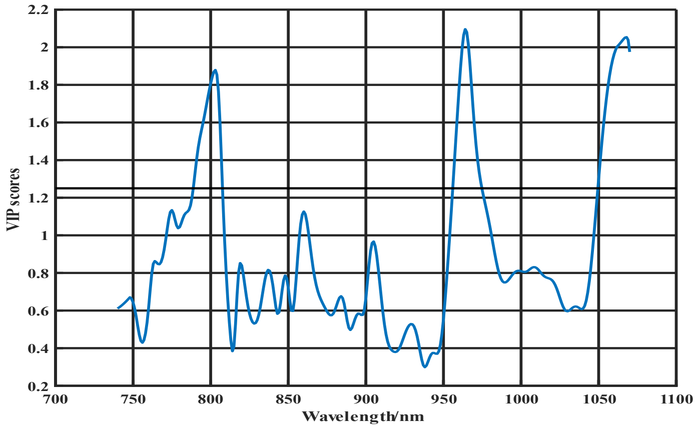


Figure 6.10 VIP Calculated Utilizing SCiO^TM. The Line Indicates the Threshold Value.

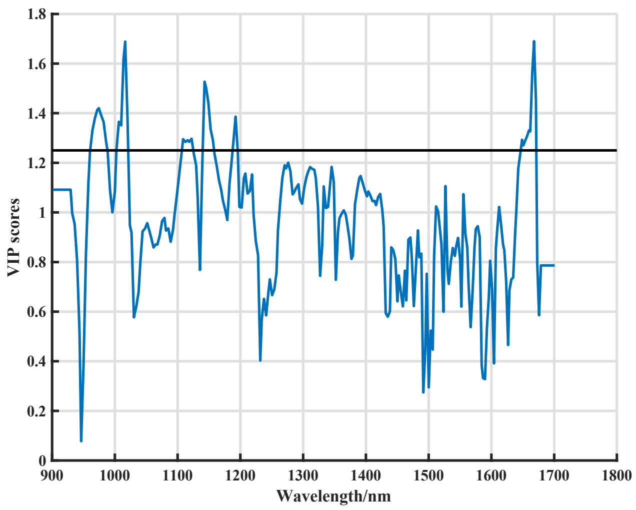


Figure 6.11. VIP Calculated Utilizing Tellspec. The Line Indicates the Threshold Value.

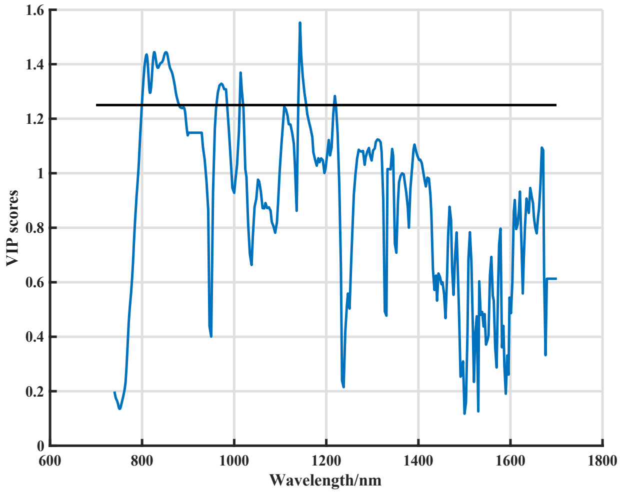


Figure 6.12. VIP Calculated Utilizing a Combination of both Tellspec and SCiO^{TM}. The Line Indicates the Threshold Value.

6.4. Conclusion

This study classified organic and inorganic pineapple juice using NIR, a nondestructive method, and PLSDA, a multivariate analytical technique. The results from the Tellspec scanner performed better than those from the SCiO scanner. The findings of the PLS-DA models indicate the viability of employing chemometric techniques for the differentiation of organic and inorganic pineapple juice. Employing NIR spectra based on combining both scanners for the PLSDA models, without any preprocessing for organic and inorganic pineapple juice, resulted in sensitivity and specificity values of 1.0 for the training set 1.0 and 1.0 for the testing set, respectively. The results of this study can pave the way for food control authorities to develop a nondestructive, rapid, and on-site strategy to differentiate organic juice from inorganic juice without generating chemical waste. NIR spectroscopy and chemometrics also offer significant benefits in terms of time, efficiency, and cost. The rapid analysis provided by NIR reduces the time required for testing, increases the efficiency of quality control processes, and minimizes costs associated with chemical reagents and labor.

Acknowledgments

We appreciate the help provided by Gideon Kwesi Nakotey, David Yaro, CSIR-Food Research Institute, World Food Programme, and School of Agriculture, UCC.

Conflict of Interest Statement

The authors have declared no conflicts of interest for this article.

Credit Authorship Contribution Statement:

Conceptualization: FPL, CLYA, VGB; ET, Formal analysis: FPL, CLYA, ET; Investigation: FPL, VGB, ET; Funding acquisition: FPL, ET; Methodology: FPL, CLYA, EEA, VGB, ET; Writing – original draft: FPL, ET; Data curation: FPL, CLYA, EEA, VGB, ET; Funding acquisition: FPL, ET; Resources: FPL, CLYA, VGB, ET; Supervision: CLYA, EEA, ET; Writing – review & editing: FPL, CLYA, EEA, VGB, ET.

CHAPTER SEVEN

COMPARATIVE ANALYSIS OF EXPIRED AND UNEXPIRED COMMERCIAL FRUIT JUICES: PHYSICOCHEMICAL AND MICROBIAL PROPERTIES.

*Francis Padi Lamptey[1,2,4], *Ernest Teye[1,4], Charles L. Y. Amuah[3,4], Ernest Ekow Abano[1]

[1]Department of Agricultural Engineering, School of Agriculture, College of Agriculture and Natural Sciences, University of Cape Coast, Cape Coast, Ghana

[2]Department of Food Science and Postharvest Technology, Cape Coast Technical University, P. O. Box DL 50, Cape Coast, Ghana.

[3]Department of Physics, Laser and Fibre Optics Centre, School of Physical Sciences, College of Agriculture and Natural Sciences, University of Cape Coast, Cape Coast, Ghana

[4]University of Cape Coast, Africa Centre of Excellence for Food Fraud and Safety Research; AfriFoodinTegrity Centre, Cape Coast, Ghana

Correspondence: ernest.teye@ucc.edu.gh/francis.lamptey@cctu.edu.gh

Submitted: International Journal of Food Science

Statement of Contribution of Joint Authorship

Lamptey F. P.: (Candidate)

Conceptualised the topic, established methodology, data collection and analysis, preparation of tables and figures, writing and compilation of the original manuscript.

Teye, E.: (Principal Supervisor)

Conceptualised the topic, established methodology, supervised and edited the manuscript and co-author of manuscript.

Abano E. E.: (Co-Supervisor)

Conceptualised the topic, supervised and edited the manuscript and co-author of manuscript.

Amuah C. L. Y.: (Research Team member)

Validation, visualization and data curation and co-author of manuscript

Abstract

Fruit juices are widely consumed for their nutritional benefits, yet they can be prone to quality degradation over time, especially after expiration. This study compares the physicochemical properties and microbial load in expired and unexpired commercial fruit juices from apples, grapes, and pineapples. Eight juice samples (four expired and four unexpired) were analysed for key parameters such as pH, titratable acidity (TA), total soluble solids (TSS), colour, and vitamin C content, as well as microbial assessments, including aerobic plate count (APC) and yeast/mold counts. Expired juices showed a significant reduction in TA (e.g., apple juice dropped from 0.60% to 0.12%), vitamin C (with a 57.6% decrease in pineapple juice), and colour lightness (L* value dropped from 10.19 to 0.81 in apple juice). Microbial analysis revealed a marked increase in yeast/mold counts in expired juices, with apple juice showing a rise from 40.00 to 97.50 CFU/mL. No pathogenic bacteria, including *Salmonella* and *Staphylococcus aureus*, were detected in any sample. These findings highlight the degradation in physicochemical properties and increased microbial load in expired juices, posing potential health risks. The study emphasizes the importance of monitoring expiration dates and ensuring proper storage conditions to maintain juice quality and consumer safety.

Keywords: Expiration date; Food safety; Fruit juice; Microbial load; Vitamin C degradation

7.1 Introduction

Consumers today are increasingly health-conscious, seeking beverages that provide excellent nutritional value and bioactive compounds, such as fibre, vitamins, minerals, antioxidants, probiotics, omega-3 fatty acids, and prebiotics (Bhardwaj, Nandal, Pal, & Jain, 2014; Giri, Sakhale, & Nirmal, 2023). Fruit juice, an unfermented beverage derived from the edible parts of fresh, mature fruits, has become a staple in many households, offering a convenient source of essential nutrients like vitamins and minerals (Nowak, Gośliński, Wojtowicz, & Przygoński, 2018).

While fruit juices are nutritious, they can also serve as a vehicle for foodborne illnesses. Contamination during production can occur due to factors such as water quality, pH, hygiene practices, raw materials, temperature, and environmental vectors (Feroz, 2018). Most fruit juice products are labelled with expiration dates, guiding consumers and food handlers on safe consumption periods. Consuming expired products can pose significant health risks (Ankiel & Samotyja, 2020), as spoilage leads to microbial growth, gas buildup, discolouration, and cloudiness, which compromise the safety and quality of the product (Feroz, 2018).

Expired fruit juices may exhibit significant differences in physicochemical properties and microbial safety compared to unexpired counterparts, potentially posing health risks. Light, temperature, and microbial contamination can alter these properties, impacting juice quality and storage stability (Kaddumukasa, Imathiu, Mathara, & Nakavuma, 2017). The physicochemical properties and microbial safety of expired fruit juices may significantly differ from those of their unexpired counterparts, potentially posing health risks to consumers. Bacterial growth in juice is influenced by pH, storage temperature, packaging, sugar content, preservatives, and production methods (Ahmed, Das, & Uddin, 2018). Processed juice undergoes biochemical changes during storage, necessitating proper processing, preservation, and chemical preservatives. Proper processing and preservation techniques, such as pasteurization and cold storage, are crucial for maintaining juice quality (Singh & Sharma, 2017). Feroz (2018) attempted to identify changes in orange juice after the expiration date printed on the product packaging.

This study investigates the changes in commercial fruit juices (apple, grape, and pineapple) after expiration. By comparing the physicochemical, microbial, and heavy metal properties of expired and unexpired juices, we aim to understand the potential risks associated with consuming expired products and provide recommendations for consumer safety. Physicochemical characteristics, including pH, acidity, total soluble solids, and sugar content, are critical indicators of juice quality and tend to degrade over time, affecting taste, texture, and safety. Additionally, microbial contamination, especially from bacteria, yeasts, and molds, may increase in expired juices, posing a risk of foodborne illnesses.

7.2 Methodology

7.2.1 Sample Collection

Eight commercially available fruit juice samples were obtained from Cape Coast, Ghana retail shops. The samples included both expired and unexpired products to allow for comparative analysis. The samples included expired and unexpired juices to enable comparative analysis, with details on the number of samples and juice types shown in Table 1. Samples of the juices are displayed in Figure 7.1.

The expired juice samples had exceeded their expiration date by more than eight months, while the unexpired samples had a remaining shelf life of at least eight months. All the fruit juices were pasteurized and shelf-stable, packaged in commercially sealed containers. After purchase, the samples were transported to the laboratories of the Department of Laboratory Technology, University of Cape Coast, and the Department of Food Science and Postharvest Technology, Cape Coast Technical University, for microbial and physicochemical analysis.

Table 7.1: Details of Expired and Unexpired Commercial Fruit Juice Samples Purchased for Analysis

Juice Type	Category	Number of samples
Apple Juice	Expired	2
	Unexpired	2
Grape Juice	Expired	1
	Unexpired	1
Pineapple Juice	Expired	1
	Unexpired	1
Total		8

7.2.2 Total Soluble Solids, pH, and Titratable Acidity Determination

Total soluble solids (°Bx) were measured using a digital refractometer (Atago, Tokyo, Japan). The samples were thoroughly mixed and used directly for determination (Hamid & Hamid, 2015; Masithoh, Haff, & Kawano, 2016).

A pH meter was used to determine the pH of the juices. The standardized electrode tip of the pH meter was immersed in the solution, and the sample was stirred gently using a rod to give a constant pH value (Hamid & Hamid, 2015). TA (%) was determined using methods from other researchers (Masithoh et al., 2016) with slight modifications. The juice samples were titrated with 0.1 N NaOH, using phenolphthalein as an indicator. The titration was performed on 2 mL of each juice sample, and TA was calculated as citric acid equivalent. The average values from triplicate measurements were used for the final calculations.

7.2.3 Determination of Vitamin C

Vitamin C was estimated in the samples using the spectrophotometric method (Khadka & Pathak, 2023). In this method, ascorbic acid was oxidised to dehydroascorbic acid at 37^0C with bromine water in an acetic acid solution coupled with 2,4-dinitrophenylhydrazine (DNPH). This solution produced a red-coloured complex with 85% H_2SO_4. The absorbance of that complex was determined at 530 nm.

7.2.4 Determination of Colour

Hunter Colorimeter (Make: HunterLab, Reston, Virginia) was used to measure the colour of the juices. The L, a, and b values were recorded as the mean of three replicates in which a low number of L value (0–50) indicates darkness and a high number (51–100) indicates lightness, positive a value indicates red and negative a value indicates green and positive b value indicates yellow and negative b value indicates blue (Das, Goud, Das, & Sahu, 2021).

7.2.5 Microbial Analysis

Aerobic plate count, Total coliform Count, *Staphylococcus aureus* count, *Salmonella sp.,* Yeast/Molds, and fugal acidophiles were determined using the Standard plate count method (Chouhan, 2015). By ISO standards (ISO-4833, 2013), the samples were analysed for Aerobic plate count, Total coliform Count, *Staphylococcus aureus* count, Yeast and Molds in colony-forming units (CFU) per ml using the pour plate method/technique. Culture media consisting of Buffered Peptone Water (Microgen), Nutrient Agar (Oxoid), Eosin Methylene Blue Agar (Oxoid), Mannitol Salt Agar (Oxoid), *Salmonella/Shigella* agar and Potato Dextrose agar (Oxoid) were prepared according to the manufacturer's instructions. Using Buffered Peptone Water (Oxoid) as recovery diluent, 180ml of the buffered peptone water was prepared in triplicate for each sample and sterilized utilizing autoclaving along with all prepared media and petri dishes at 121°C, 15psi for 15 minutes. The sample was allowed to thaw at room temperature and adequately homogenized, 20g of the test sample was weighed aseptically into the recovery diluent (Buffered Peptone Water) and incubated in a water bath at 37ºC for 30 minutes. The test sample was serially diluted to 10^{-3} in sterile buffered peptone water.

For Aerobic plate count, triplicate dilutions of 1ml of 10^{-2} samples were plated on Nutrient agar. Each triplicate dilutions were incubated at 35±2°C for 48 hours to observe for aerobic plate count (CFU/ml) for each sample. For Total Coliform count, triplicate dilutions of 1ml of 10^{-1} sample were plated on Eosin Methylene Agar. Each triplicate dilution was incubated at 35±2°C for 48 hours to observe each sample's Total Coliform count (CFU/ml).

For *Staphylococcus aureus* count, triplicate dilutions of 1ml of 10^{-1} sample were plated on Mannitol Salt Agar. Each triplicate dilution was incubated at 35±2°C for 48 hours to observe each sample for Staphylococcus aureus count (CFU/ml).

For Total *Salmonella* count, triplicate dilutions of 1ml of 10^{-1} sample were plated on *Salmonella/Shigella* agar. Each triplicate dilution was incubated at 35±2°C for 48 hours to observe each sample's Total Salmonella count (CFU/ml).

For Yeast and Molds, triplicate dilutions of 1 ml of 10^{-1} sample were plated on Potato Dextrose agar supplemented with ampicillin. Each triplicate dilution was incubated at room temperature for 7 days to observe for yeast and mold counts (CFU/ml) for the sample. Fungal acidurics were assessed according to (Vantarakis, Affifi, Kokkinos, Tsibouxi, & Papapetropoulou, 2011).

Figure 7.1: Samples of Fruit Juice: A. Expired Apple Juice, B. Unexpired Apple Juice, C. Unexpired Grape Juice D. Expired Grape Juice E. Unexpired Pineapple Juice, F. Expired Pineapple Juice, G. Unexpired Apple Juice, and H. Expired Apple Juice

7.3. Results and Discussions

This study investigated the physicochemical and microbial changes in commercial fruit juices (apple, grape, and pineapple) before and after expiration (Tables 7.2 - 7.5). The results reveal significant alterations in various parameters, which have important implications for juice quality, safety, and consumer health.

7.3.1 Physicochemical Changes

7.3.1.1 Titratable Acidity (TA) and pH

Titratable acidity and pH are interrelated concepts in food analysis that deal with acidity. These quantities are analytically determined separately, providing particular insights on food quality. For example, while pH is essential to assess the ability of a microorganism to grow in a specific food, titratable acidity is a better predictor than pH of how organic acids in the food impact flavor (Sadler & Murphy, 2010; Tyl & Sadler, 2017). The pH values of all expired and unexpired juice samples ranged from 3.44 to 3.74, which is consistent with the typical pH range for fruit juices (Bhardwaj & Pandey, 2011; Kaddumukasa et al., 2017). Across all juice types, a consistent decrease in titratable acidity was observed in expired samples. This reduction was particularly pronounced in apple juice (Table 7.1), where TA decreased from 0.60 to 0.12, and in pineapple juice (Table 7.4), with a drop from 1.02 to 0.12. The decrease in titratable acidity in juice might be due to the chemical reaction between organic constituents, which increased upon prolonged storage (Bhardwaj & Pandey, 2011). The loss of acidity affects the flavor profile and may compromise the juice's natural preservative properties, potentially allowing for increased microbial growth.

7.3.1.2 Total Soluble Solids (TSS)

The Total soluble solid values, measured in °Brix, showed variations between expired and unexpired samples. It remained relatively stable in apple juice (Sample 1) but showed significant changes in grape and pineapple juices (Tables 7.3 and 7.4, respectively). The decrease in TSS for grape juice (from 14.23 to 12.70) might be attributed to the action of microorganisms present in the juice (Chauhan, Singh, Tyagi, & Balyan, 2002), while the slight increase in pineapple juice (from 12.53 to 12.80) might be due to hydrolysis of polysaccharides into monosaccharides and increase in concentration of juice due to dehydration (Bhardwaj & Pandey, 2011). These changes can affect the sweetness and overall sensory qualities of the juices.

7.3.1.3 Colour Parameters (L*, a*, b*)

changes were evident across all samples, with alterations in lightness (L*) and redness/greenness (a*) values. The decrease in L* values for all expired sample juices suggests darkening (Das et al., 2021), which could be attributed to the oxidation of phenolic compounds, enzymatic browning, or Maillard reactions (Wibowo et al., 2015). The shift towards positive a* values in expired samples indicates a loss of green colour components and an increase in redness, potentially due to the degradation of chlorophyll and the formation of brown pigments (Das et al., 2021; Wibowo et al., 2015). These colour changes do not only affect the visual appeal of the juices but may also indicate chemical alterations that could impact nutritional quality.

7.3.1.4 Vitamin C Content

A substantial decrease in vitamin C content was observed in all expired samples, with reductions ranging from 35.8-57.6%. This loss of vitamin C directly impacts the juices' nutritional value and antioxidant properties (Bhardwaj & Pandey, 2011). Ascorbic acid is sensitive to heat and is oxidised quickly in the presence of oxygen. Hence, it might have been destroyed during the storage period due to its oxidation (Bhardwaj & Pandey, 2011).

7.3.1.5 Microbial Changes

All expired and unexpired juice samples showed no detectable levels of common pathogenic bacteria such as *Salmonella sp., Bacillus cereus, Staphylococcus aureus*, and *Escherichia coli* (Table 7.4). Tasnim et al. (2010) also detected no coliform or Salmonella in industrially processed packed juices in Bangladesh. In contrast, Vantarakis et al. (2011) also indicated that fruit juice samples sold in retail markets in Greece recorded negative for *Salmonella spp., Bacillus cereus, total coliforms, and E. coli*. Kaddumukasa et al. (2017) also documented that lower pH generally inhibits bacterial growth. This absence of pathogens indicates good manufacturing practices and effective preservation methods the juice producers employ. All expired juice samples showed a marked increase in APC, indicating a general increase in microbial load. This growth could be due to the degradation of preservatives over time or changes in the juice environment (such as decreased acidity) that favor microbial proliferation.

The most dramatic increases were observed in YM counts, particularly in apple (Sample 1) and grape juices, where counts rose from negligible levels to 97.50 and 90.00, respectively. This substantial growth of fungi in expired juices raises significant food safety concerns. Other researchers also reported an increased microbial population (total plate counts and yeast and mold counts) during the storage of sugarcane juice (Chauhan et al., 2002). The appearance of acidophilic microorganisms in expired apples (Sample 1) and grape juices, where none were detected in fresh samples, indicates a shift in the microbial ecosystem of the juices. Ampofo-Asiama and Quaye (2018) documented that lower pH favors the growth of acidophilic microorganisms.

Table 7.2.: Physicochemical Properties and Microbial Load of Expired and Unexpired Apple Juice Samples (Sample 1)

Parameter(s)	Authentic	Expired	p-value
TA (%)	0.60±0.01	0.12±0.00	0.00
pH	3.65±0.01	3.57±0.01	0.00
TSS (ºB)	12.53±0.06	12.50±0.10	0.64
L	10.19±0.80	0.81±1.04	0.00
A	-2.5±2.12	5.04±2.46	0.02
B	9.24±2.89	6.08±1.61	0.17
Vitamin C (mg/100ml)	25.2±0.04	10.93±0.65	0.00
APC (CFU/ml)	50.50±0.50	61.00±1.00	0.00
YM (CFU/ml)	40.00±0.00	97.50±2.50	0.00
Acidophiles (CFU/ml)	0.00±0.00	54.50±0.00	0.00

Note: TA (Titratable acidity), TSS (Total soluble solids), APC (Aerobic plate count), YM (Yeast/mold counts) and CFU (Colony forming unit).

Table 7.3: Physicochemical Properties and Microbial Load of Expired and Unexpired Grape Juice Samples (Sample 2)

Parameter(s)	Authentic	Expired	p-value
TA (%)	0.42±0.04	0.17±0.01	0.00
pH	3.60±0.01	3.55±0.01	0.02
TSS (ºB)	14.23±0.06	12.70±0.10	0.00
L	2.44±2.04	0.06±0.11	0.11
A	4.20±2.31	4.45±0.90	0.87
B	5.42±2.28	12.72±3.99	0.05
Vitamin C (mg/100ml)	20.91±0.01	13.43±0.27	0.00
APC (CFU/ml)	15.00±5.00	75.00±0.00	0.00
YM (CFU/ml)	0.00±0.00	90.00±0.00	0.00
Acidophiles (CFU/ml)	0.00±0.50	50.50±0.00	0.00

Note: TA (Titratable acidity), TSS (Total soluble solids), APC (Aerobic plate count), YM (Yeast/mold counts) and CFU (Colony forming unit).

Table 7.4: Physicochemical Properties and Microbial Load of Expired and Unexpired Pineapple Juice Samples (Sample 3)

Parameter(s)	Authentic	Expired	p-value
TA (%)	1.02±0.00	0.12±0.01	0.00
pH	3.66±0.01	3.74±0.01	0.00
TSS (ºB)	12.53±0.06	12.80±0.0.	0.00
L	3.37±0.61	2.35±0.16	0.05
A	-6.12±1.20	4.08±0.93	0.00
B	4.28±2.55	2.55±1.31	0.35
Vitamin C (mg/100ml)	29.22±0.00	12.40±0.07	0.00
APC (CFU/ml)	66.50±0.50	75.00±5.005	0.04
YM (CFU/ml)	0.00±0.00	41.00±1.00	0.00
Acidophiles (CFU/ml)	0.00±0.00	0.00±0.00	

Note: TA (Titratable acidity), TSS (Total soluble solids), APC (Aerobic plate count), YM (Yeast/mold counts) and CFU (Colony forming unit).

Table 7.5: Physicochemical Properties and Microbial Load of Expired and Unexpired Apple Juice Samples (Sample 4)

Parameter(s)	Authentic	Expired	p-value
TA (%)	0.55±0.01	0.44±0.00	0.000
pH	3.63±0.01	3.44±0.01	0.000
TSS (oB)	12.50±0.00	14.20±0.00	0.000
L	9.99±1.70	2.61±0.61	0.002
a	-2.39±1.86	3.33±1.22	0.011
b	5.16±1.07	4.0.00±1.21	0.281
Vitamin C (mg/100ml)	17.53±0.00	8.83±0.00	0.00
APC (CFU/ml)	40.00±0.00	55.50±0.00	0.000
YM (CFU/ml)	11.50±0.00	22.50±2.50	0.003
Acidophiles (CFU/ml)	0.00±0.00	0.00±0.00	

Note: TA (Titratable acidity), TSS (Total Soluble Solids), APC (Aerobic Plate Count), YM (Yeast/Mold Counts) and CFU (Colony Forming Unit).

7.4 Conclusion

The comparative analysis of expired and unexpired commercial fruit juices demonstrated significant changes in physicochemical properties and microbial load after expiration. Expired juices exhibited reduced titratable acidity, vitamin C content, and alterations in colour, which impact the products' sensory qualities. Furthermore, the increase in microbial counts, particularly yeast and molds, raises concerns about the safety of consuming expired juices. While no pathogenic bacteria were detected, the presence of acidophilic microorganisms in expired samples emphasizes the risks associated with prolonged storage. These results underscore the need for strict adherence to expiration dates and proper preservation methods to maintain the safety and quality of fruit juices for consumers.

CHAPTER EIGHT

SUMMARY, GENERAL CONCLUSIONS AND RECOMMENDATION

8.1 **Summary**

This research focused on the developing handheld near-infrared (NIR) spectroscopy technique for assessing the quality and integrity of mangoes fruits and products, as well as analyzing the effects of expiration on commercial fruit juices safety. Handheld NIR spectroscopy combined with effective selection of multivariate data modeling revealed some useful information. The study investigated several aspects of fruit quality, including variety differentiation, ripeness, ripening agents, and organic authenticity. The first objective was developed using NIR spectroscopy to identify different mango varieties and predict their maturity based on pH and total soluble solids (TSS). The study involved scanning 198 mangoes and applying various preprocessing techniques and multivariate models, achieving high accuracy in training and prediction sets.

The second objective focused on detecting ethephon, a ripening agent, in artificially ripened mangoes. Different ethephon concentrations were analysed using NIR spectroscopy, with different classification models such as neural networks and random forests. Both models delivered a high accuracy in predicting the concentration levels. The third objective examined the distinction between organic and inorganic mangoes in three forms: raw fruits, chips, and juice. Random Forest models combined with specific preprocessing techniques accurately classified organic and inorganic products. The fourth objective applied dual handheld NIR spectrometers (SCiO and Tellspec) comparatively to differentiate organic and inorganic pineapple juice. The results showed perfect classification accuracy using partial least squares discriminant analysis (PLS-DA) models. The fifth objective compared expired and unexpired fruit juices regarding physicochemical properties and microbial load, highlighting significant quality degradation in expired samples, including reduced acidity, vitamin C, and increased microbial contamination.

8.1.1 Key Findings

The following findings regarding the research objectives that guided the study were revealed concerning the first research objective. NIR spectroscopy in the 740–1070 nm range provided essential insights into determining mango fruit quality and variety. For mango variety identification, the LDA-SVM classifier coupled with various preprocessing techniques such as RAW, MC, SNV, FD, and SD achieved 100% accuracy in the training set and 97.44% in the prediction set. This demonstrates that the combination of these techniques offers excellent classification results.

In the quantification of total soluble solids (TSS) and pH in mangoes, the synergy partial least square (Si-PLS) model proved the most effective, with an r^2 of 0.63 and an RMSEP of 1.83 for TSS and an r^2 of 0.81 and an RMSEP of 0.49 for pH. These findings show that handheld NIR devices combined with suitable chemometric tools can rapidly and non-destructively assess mango maturity and quality.

The second objective was to detect ethephon, a ripening agent used in mangoes, using NIR spectroscopy in combination with chemometric methods. The neural network (NN) model and MSC preprocessing achieved 100% accuracy in the training and test sets for classifying ethephon concentrations. For quantitative detection, partial least squares (PLS) regression with SNV preprocessing yielded an r^2 of 0.996 and an RMSEP of 0.068 in the test set, demonstrating a robust capability to predict ethephon concentration. These findings reveal that NIR spectroscopy can effectively monitor banned artificial ripening agents in mangoes' ripening processes.

The third objective aimed to differentiate organic and inorganic mangoes in three forms (raw fruits, chips, and juices) using NIR spectroscopy with various preprocessing tools and classification algorithms. The random forest algorithm was applied for classification, yielding different accuracy levels across the training and test sets. The algorithm achieved 100% accuracy in the training set and 88.76% in the test set for mango fruits. For mango chips, it attained 99.94% accuracy in the training set and 77.98% in the test set when preprocessed using the second derivative. For mango juice, the algorithm reached 100% accuracy in the training set and 87.53% in the test set without preprocessing. The study demonstrated that portable NIR spectroscopy is a reliable, non-invasive method for authenticating organic mango products.

The fourth objective sought to differentiate organic pineapple juice from its inorganic counterparts using dual handheld NIR devices (SCiO and Tellspec) with chemometric models such as partial least squares-discriminant analysis (PLS-DA). Tellspec, paired with second derivative preprocessing, achieved the highest classification accuracy. When the data from both devices were combined, perfect classification accuracy (100%) was achieved for both training and testing sets, with no preprocessing required. This demonstrates the potential for dual NIR spectroscopy devices for rapid, on-site differentiation of organic pineapple juices.

The fifth objective was to compare expired and unexpired commercial fruit juices (apple, grape, and pineapple). Expired juices showed a significant reduction in TA (e.g., apple juice dropped from 0.60% to 0.12%), vitamin C (with a 57.6% decrease in pineapple juice), and colour lightness (L* value dropped from 10.19 to 0.81 in apple juice). Microbial analysis revealed a marked increase in yeast/mold counts, with apple juice showing the most significant rise. Although no pathogenic bacteria were found, the results highlight the degradation in quality and potential health risks of expired juices, emphasizing the importance of monitoring expiration dates to maintain consumer safety and quality.

8.2 Conclusion

Findings from the first objective suggest that handheld near-infrared (NIR) spectroscopy combined with suitable chemometric techniques is a rapid, non-destructive method for identifying mango varieties and predicting their maturity. The use of various preprocessing methods (RAW, MC, SNV, FD, SD) and multivariate classification models, including support vector machine (SVM) and linear discriminant analysis (LDA), proved effective for identifying mango varieties with 100% accuracy in the training set and 97.44% in the prediction set. It can be concluded that handheld NIR spectroscopy, coupled with chemometric models, presents a reliable method for classifying mango varieties and assessing fruit maturity, which can be applied for on-site, real-time quality control.

In the second objective, the detection and quantification of ethephon concentrations in artificially ripened mangoes could be successfully achieved using handheld NIR spectroscopy combined with chemometric techniques.

The neural network (NN) model, particularly with MSC preprocessing, demonstrated 100% accuracy in training and test sets. Partial least squares (PLS) regression also exhibited excellent predictive performance, with an r^2 value of 0.996 in the test set. These results suggest that NIR spectroscopy, combined with appropriate chemometric models, can be employed as a non-invasive, rapid method for monitoring artificial ripening processes, supporting food safety regulations in the fruit industry.

In the third objective, the study demonstrated that handheld NIR spectroscopy can effectively distinguish between organic and inorganic mango fruits, chips, and juice. By using multivariate classification algorithms such as random forest (RF) and principal component analysis (PCA) combined with preprocessing methods (MSC, SNV, FD), the classification accuracies ranged from 65.12% to 100%. These findings validate the utility of NIR spectroscopy for authenticating organic mango products, providing a reliable, non-destructive solution for ensuring the integrity of food labeling.

In the fourth objective, organic and inorganic pineapple juices were differentiated using dual handheld NIR spectrometers (SCiO and Tellspec) combined with PLS-DA models. The fusion of spectral data from both devices resulted in 100% classification accuracy, sensitivity, and specificity in the training and testing sets. These results demonstrate the potential of combining NIR spectroscopy with advanced chemometric techniques for non-destructive, rapid differentiation of organic and inorganic pineapple juices, promoting quality assurance in the food industry.

In the fifth objective, the comparative analysis of expired and unexpired commercial fruit juices revealed significant degradation in physicochemical properties, such as decreased titratable acidity and vitamin C content, in expired samples. The study also recorded increased microbial load in expired juices, particularly yeast and mold. These findings highlight the need for strict adherence to expiration dates and proper storage practices to ensure the safety and quality of fruit juices.

8.3 Recommendation

Near-infrared (NIR) spectroscopy has proven to be a non-invasive analytical method that aids in the quality assessment and verification of fruits and their products. Portable NIR devices can be used on-site in farms or during post-harvest handling to assess the optimal harvest time, ensuring the fruits are harvested at peak ripeness for better quality and market value.

It can also be integrated into processing lines for instantaneous quality monitoring, bypassing destructive testing. It is beneficial for routine evaluations in food processing facilities to confirm product authenticity, especially for high-standard export products.

Training initiatives for farmers and producers in developing countries like Ghana on handheld NIR spectrometers can improve local food quality monitoring and product integrity.

Further research on the economic feasibility of NIRS adoption among Ghanaian farmers and processors is recommended to assess cost implications, infrastructure requirements, and potential barriers.

The food industry should focus on stricter expiration date monitoring, ensuring that products past their expiration date are removed from shelves to prevent potential health risks from expired juices.

REFERENCE LIST

Ab Razak, S., Ariffin, M. A. T., Mohamad, S. M. S., Azman, N. H. E. N., Hassan, M. A., & Sarip, J. (2020). Microsatellite markers for the molecular characterisation of potentially commercial mango (Mangifera Indica) progenies. *Malaysian Applied Biology, 49*(3), 81-85.

Abarra, M. S. J., Serrano, E. P., Sabularse, V. C., Mendoza, H. E. T., & Del Rosario, E. (2018). Determination of fruit ripeness degree of 'Carabao'mango (Mangifera indica L.) using digital photometry. *Philippine Journal of Science, 147*(2), 249-253.

Abd Salam, N. A., Saad, W. H. M., Manap, Z., Salehuddin, F., & Karim, S. A. A. (2018). Comparative Study of Different Near-Infrared (NIR) Wavelengths on Glucose Concentration Detection. *Journal of Telecommunication, Electronic and Computer Engineering, 10*(2-6), 59-64.

Abelha, M., Fernandes, S., Mesquita, D., Seabra, F., & Ferreira-Oliveira, A. T. (2020). Graduate employability and competence development in higher education—A systematic literature review using PRISMA. *Sustainability, 12*(15), 5900. doi:10.3390/su12155900.

Abu, M., Abbey, L. D., & Amey, N. K. (2021). Relation of harvesting time on physicochemical properties of Haden, Kent, Palmer and Keitt mango varieties for export and local markets. *Journal of horticulture and postharvest research, 4*(1), 87-100.

Aday, M. S., & Caner, C. (2014). Individual and combined effects of ultrasound, ozone and chlorine dioxide on strawberry storage life.

LWT-Food Science and Technology, 57(1), 344-351. doi:10.1016/j.lwt.2014.01.006

Adebayo, A., Unigbe, O., & Atanda, E. (2014). Fabrication and performance evaluation of a portable motorized pineapple juice extractor. *Innovative Systems Design and Engineering, 5*(8), 22-29.

Adeyemi, M., Bawa, M., & Muktar, B. (2018). Evaluation of the Effect of Calcium Carbide on Induce Ripening of Banana, Pawpaw and Mango cultivated within Kaduna Metropolis, Nigeriaâ€. *Journal of Chemical Society of Nigeria, 43*(2).

Adriaanse, L. S., & Rensleigh, C. (2013). Web of Science, Scopus and Google Scholar: A content comprehensiveness comparison. *The Electronic Library, 31*(6), 727-744. doi:10.1108/EL-12-2011-0174

Agelet, L. E., & Hurburgh Jr, C. R. (2010). A tutorial on near infrared spectroscopy and its calibration. *Critical Reviews in Analytical Chemistry, 40*(4), 246-260. doi:10.1080/10408347.2010.515468

Agulheiro-Santos, A. C., Ricardo-Rodrigues, S., Laranjo, M., Melgão, C., & Velázquez, R. (2022). Non-destructive prediction of total soluble solids in strawberry using near infrared spectroscopy. *Journal of the Science of Food Agriculture, 102*(11), 4866-4872. doi:10.1002/jsfa.11849

Agyekum, E. B., Odoi-Yorke, F., Abbey, A. A., & Ayetor, G. K. (2024). A review of the trends, evolution, and future research prospects of hydrogen fuel cells–A focus on vehicles. *International Journal of Hydrogen Energy, 72*, 918-939. doi:10.1016/j.ijhydene.2024.05.480

Ahmed, T., Das, K. K., & Uddin, M. A. (2018). The microbiological quality of commercial fruit juices-current perspectives. *Bangladesh Journal of Microbiology, 35*(2), 128-133.

Aili Hamzah, A. F., Hamzah, M. H., Che Man, H., Jamali, N. S., Siajam, S. I., & Ismail, M. H. (2021). Recent updates on the conversion of pineapple waste (Ananas comosus) to value-added products, future perspectives and challenges. *Agronomy, 11*(11), 2221.

Ajayi, I., Olawuyi, O., Ayodele, A., & Faneye, A. (2019). Molecular relationship among Mangifera indica L.(Mango) varieties using simple sequence repeat (SSR) marker. *Adv Biol Biotechnol, 22*(4), 1-16.

Alamar, P. D., Caramês, E. T., Poppi, R. J., & Pallone, J. A. (2020). Detection of fruit pulp adulteration using multivariate analysis: Comparison of NIR, MIR and data fusion performance. *Food analytical methods, 13*, 1357-1365. doi:10.1007/s12161-020-01755-x

Aleixandre-Tudó, J. L., Castelló-Cogollos, L., Aleixandre, J. L., & Aleixandre-Benavent, R. (2020). Bibliometric insights into the spectroscopy research field: A food science and technology case study. *Applied Spectroscopy Reviews, 55*(9-10), 873-906. doi:10.1080/05704928.2019.1694936

Aleixandre-Tudó, J. L., Castelló-Cogollos, L., Aleixandre, J. L., & Aleixandre-Benavent, R. (2022). Chemometrics in food science and technology: A bibliometric study. *Chemometrics and Intelligent Laboratory Systems, 222*, 104514. doi:10.1016/j.chemolab.2022.104514

Alolfe, M. A., Mohamed, W. A., Youssef, A.-B. M., Mohamed, A. S., & Kadah, Y. M. (2009). *Computer aided diagnosis in digital mammography using combined support vector machine and linear discriminant analyasis classification.* Paper presented at the 2009 16th IEEE International Conference on Image Processing (ICIP).

Ambrose, A., Lohumi, S., Lee, W.-H., & Cho, B. K. (2016). Comparative nondestructive measurement of corn seed viability using Fourier transform near-infrared (FT-NIR) and Raman spectroscopy. *Sensors and Actuators B: Chemical, 224*, 500-506.

Amodio, M. L., Ceglie, F., Chaudhry, M. M. A., Piazzolla, F., & Colelli, G. (2017). Potential of NIR spectroscopy for predicting internal quality and discriminating among strawberry fruits from different production systems. *Postharvest biology and technology, 125*, 112-121.

Ampofo-Asiama, J., & Quaye, B. (2018). Effect of storage temperature on the physicochemical, nutritional and microbiological quality of pasteurized soursop (Annona muricata L.) Juice. *African Journal of Food Science*. doi:10.5897/AJFS2018.1767

Amuah, C. L., Teye, E., Lamptey, F. P., Nyandey, K., Opoku-Ansah, J., & Adueming, P. O.-W. (2019). Feasibility study of the use of handheld NIR spectrometer for simultaneous authentication and quantification of quality parameters in intact pineapple fruits. *Journal of Spectroscopy, 2019*.

Andersen, C. M., & Bro, R. (2010). Variable selection in regression—a tutorial. *Journal of Chemometrics, 24*(11-12), 728-737. doi:10.1002/cem.1360

Anderson, N., Walsh, K., Subedi, P., & Hayes, C. (2020). Achieving robustness across season, location and cultivar for a NIRS model for intact mango fruit dry matter content. *Postharvest biology and technology, 168*, 111202. doi:10.1016/j.postharvbio.2020.111202

Andika, A., & Bidayati, U. (2024). Unraveling the complexity of the organic food market: Indonesian consumer perspective on price and product knowledge. *Asian Management and Business Review*, 73-89. doi:10.20885/AMBR.vol4.iss1.art5

Ankiel, M., & Samotyja, U. (2020). The role of labels and perceived health risk in avoidable food wasting. *Sustainability, 12*(20), 8725. doi:10. 3390/su12208725

Anwar, R., Mattoo, A. K., & Handa, A. K. (2018). Ripening and senescence of fleshy fruits. *Postharvest biology and nanotechnology*, 15-51. doi:10.1002/9781119289470.ch2

Anyidoho, E. K., Teye, E., & Agbemafle, R. (2020). Nondestructive authentication of the regional and geographical origin of cocoa beans by using a handheld NIR spectrometer and multivariate algorithm. *Analytical Methods, 12*(33), 4150-4158. doi:10.1039/D0AY00901F

Anyidoho, E. K., Teye, E., & Agbemafle, R. (2021). Differentiation of organic cocoa beans and conventional ones by using handheld NIR spectroscopy and multivariate classification techniques. *International Journal of Food Science, 2021*(1), 1844675. doi:10.1155/2021/18446 75

Anyidoho, E. K., Teye, E., Agbemafle, R., Amuah, C. L., & Boadu, V. G. (2021). Application of portable near infrared spectroscopy for classifying and quantifying cocoa bean quality parameters. *Journal of food processing and preservation, 45*(5), e15445.

Arendse, E., Fawole, O. A., Magwaza, L. S., Nieuwoudt, H. H., & Opara, U. L. (2017). Development of calibration models for the evaluation of pomegranate aril quality by Fourier-transform near infrared spectroscopy combined with chemometrics. *Biosystems Engineering, 159*, 22-32.

Arendse, E., Fawole, O. A., Magwaza, L. S., & Opara, U. L. (2018). Non-destructive prediction of internal and external quality attributes of fruit with thick rind: A review. *Journal of Food Engineering, 217*, 11-23. doi:10.1016/j.jfoodeng.2017.08.009

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of informetrics, 11*(4), 959-975. doi:10.1016/j.joi.2017.08.007

Asif, M. (2011). The role of fruits, vegetables, and spices in diabetes. *International journal of nutrition, pharmacology, neurological diseases, 1*(1), 27-35. doi:10.4103/2231-0738.77527

Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative science studies, 1*(1), 377-386. doi:10.1162/qss_a_00019

Bannor, R. K., Arthur, K. K., Oppong, D., & Oppong-Kyeremeh, H. (2023). A comprehensive systematic review and bibliometric analysis of food fraud from a global perspective. *Journal of Agriculture and Food Research, 14*, 100686. doi:10.1016/j.jafr.2023.100686

Baranska, M., Schütze, W., & Schulz, H. (2006). Determination of lycopene and β-carotene content in tomato fruits and related products: comparison of FT-Raman, ATR-IR, and NIR spectroscopy. *Analytical Chemistry, 78*(24), 8456-8461. doi:10.1021/ac061220j

Baranski, R., Baranska, M., & Schulz, H. (2005). Changes in carotenoid content and distribution in living plant tissue can be observed and mapped in situ using NIR-FT-Raman spectroscopy. *Planta, 222*, 448-457. doi:10.1007/s00425-005-1566-9

Beausang, C., Hall, C., & Toma, L. (2017). Food waste and losses in primary production: Qualitative insights from horticulture. *Resources, Conservation and Recycling, 126*, 177-185. doi:10.1016/j.resconrec.2017.07.042

Beghi, R., Giovanelli, G., Malegori, C., Giovenzana, V., & Guidetti, R. (2014). Testing of a VIS-NIR system for the monitoring of long-term apple storage. *Food and Bioprocess Technology, 7*, 2134-2143. doi:10.1007/s11947-014-1294-x

Bekele, M., Satheesh, N., & Sadik, J. (2020). Screening of Ethiopian mango cultivars for suitability for preparing jam and determination of pectin, sugar, and acid effects on physico-chemical and sensory properties of mango jam. *Scientific African, 7*, e00277. doi:10.1016/j.sciaf.2020.e00277

Bhardwaj, R. L., Nandal, U., Pal, A., & Jain, S. (2014). Bioactive compounds and medicinal properties of fruit juices. *Fruits, 69*(5), 391-412. doi:10.1051/fruits/2014027

Bhardwaj, R. L., & Pandey, S. (2011). Juice blends—a way of utilization of under-utilized fruits, vegetables, and spices: a review. *Critical reviews in food science and nutrition, 51*(6), 563-570. doi:10.1080/10408391003710654

Bhattacharjee, P., Warang, O., Das, S., & Das, S. (2022). Impact of Climate Change on Fruit Crops-A Review. *Current World Environment, 17*(2), 319. doi:10.12944/CWE.17.2.4

Bilalis, D., Krokida, M., Roussis, I., Papastylianou, P., Travlos, I., Cheimona, N., & Dede, A. (2018). Effects of organic and inorganic fertilization on yield and quality of processing tomato (Mill.). *Folia Horticulturae, 30*(2), 321-332. doi:10.2478/fhort-2018-0027

Boadu, V. G., Teye, E., Amuah, C. L., & Sam-Amoah, L. (2022). Rapid authentication of coffee bean varieties of different forms by using a pocket-sized spectrometer and multivariate data modelling. *Analytical Methods, 14*(46), 4756-4766. doi:10.1039/D2AY01480G

Boadu, V. G., Teye, E., Lamptey, F. P., Amuah, C. L. Y., & Sam-Amoah, L. (2024). Novel authentication of African geographical coffee types (bean, roasted, powdered) by handheld NIR spectroscopic method. *Heliyon, 10*(15). doi:10.1016/j.heliyon.2024.e35512

Bureau, S., Ruiz, D., Reich, M., Gouble, B., Bertrand, D., Audergon, J.-M., & Renard, C. M. (2009). Rapid and non-destructive analysis of apricot fruit quality using FT-near-infrared spectroscopy. *Food chemistry, 113*(4), 1323-1328. doi:10.1016/j.foodchem.2008.08.066

Camps, C., & Christen, D. (2009). Non-destructive assessment of apricot fruit quality by portable visible-near infrared spectroscopy. *LWT-Food*

Science and Technology, 42(6), 1125-1131. doi:10.1016/j.lwt.2009.01.015

Cayuela, J. A., & Weiland, C. (2010). Intact orange quality prediction with two portable NIR spectrometers. *Postharvest biology and technology, 58*(2), 113-120. doi:10.1016/j.postharvbio.2010.06.001

Cen, H., & He, Y. (2007). Theory and application of near infrared reflectance spectroscopy in determination of food quality. *Trends in Food Science Technology, 18*(2), 72-83.

Chaudhary, V., Kumar, V., Singh, K., Kumar, R., & Kumar, V. (2019). Pineapple (Ananas cosmosus) product processing: A review. *Journal of pharmacognosy and Phytochemistry, 8*(3), 4642-4652.

Chauhan, O., Singh, D., Tyagi, S., & Balyan, D. (2002). Studies on preservation of sugarcane juice. *International journal of food properties, 5*(1), 217-229. doi:10.1081/JFP-120015603

Chen, H., Qiao, H., Xu, L., Feng, Q., & Cai, K. (2019). A fuzzy optimization strategy for the implementation of RBF LSSVR model in vis–NIR analysis of pomelo maturity. *IEEE Transactions on Industrial Informatics, 15*(11), 5971-5979. doi:10.1109/TII.2019.2933582

Chen, L., & Opara, U. L. (2013). Texture measurement approaches in fresh and processed foods—A review. *Food Research International, 51*(2), 823-835. doi:10.1016/j.foodres.2013.01.046

Chen, W., Li, H., Zhang, F., Xiao, W., Zhang, R., Chen, Z., & Du, Y. (2021). Handheld short-wavelength NIR spectroscopy for rapid determination of sugars and carbohydrate in fresh juice with sampling error profile analysis. *Infrared Physics and Technology, 115*, 103732. doi:10.1016/j.infrared.2021.103732

Chen, X., Xue, J., Chen, X., Zhao, X., Ali, S., & Huang, G. (2023). Gaussian process regression for prediction and confidence analysis of fruit traits by near-infrared spectroscopy. *Food Quality and Safety, 7*, fyac068. doi:10.1093/fqsafe/fyac068

Cheng, W., Sørensen, K. M., Mongi, R. J., Ndabikunze, B. K., Chove, B. E., Sun, D.-W., & Engelsen, S. B. (2019). A comparative study of mango solar drying methods by visible and near-infrared spectroscopy coupled with ANOVA-simultaneous component analysis (ASCA). *Lwt, 112*, 108214. doi:10.1016/j.lwt.2019.05.112

Chia, K. S., & Suarin, N. A. S. B. (2022). *Neural network and extreme gradient boosting in near infrared spectroscopy.* Paper presented at the 2022 International Conference on Innovations and Development of Information Technologies and Robotics (IDITR).

Chouhan, S. (2015). Enumeration and identification of standard plate count bacteria in raw water supplies. *IOSR Journal of Environmental Science, Toxicology and Food Technology, 9*(2), 67-73. doi:10.9790/2402-09226773

Clément, A., Dorais, M., & Vernon, M. (2008). Nondestructive measurement of fresh tomato lycopene content and other physicochemical characteristics using visible– NIR spectroscopy. *Journal of agricultural and food chemistry, 56*(21), 9813-9818. doi:10.1021/jf801299r

Codex, Alimentarius. (2016). Critical and emerging issues in food safety and quality. Codex Alimentarius Commission. Assessed September 29, 2023.https://www.codexalimentarius.nl/sites/codexalimentarius.nl/files/bijlagen/cceuro_critical_and_emerging_issues_eu_final.pdf.

Consumer Physics. (2020). Powerful SCiO NIR microspectroscopy at your fingertips. Retrieved from https://www.consumerphysics.com/techno-logy/

Cortés, V., Blasco, J., Aleixos, N., Cubero, S., & Talens, P. (2019). Monitoring strategies for quality control of agricultural products using visible and near-infrared spectroscopy: A review. *Trends in Food Science and Technology, 85*, 138-148. doi:10.1016/j.tifs.2019.01.015

Cortés, V., Ortiz, C., Aleixos, N., Blasco, J., Cubero, S., & Talens, P. (2016). A new internal quality index for mango and its prediction by external visible and near-infrared reflection spectroscopy. *Postharvest Biology Technology, 118*, 148-158.

Cozzolino, D., & Corbella, E. (2003). Determination of honey quality components by near infrared reflectance spectroscopy. *Journal of apicultural research, 42*(1-2), 16-20.

Czarnecki, M. A., Morisawa, Y., Futami, Y., & Ozaki, Y. (2015). Advances in molecular structure and interaction studies using near-infrared spectroscopy. *Chemical reviews, 115*(18), 9707-9744. doi:10.1021/cr500013u

Dalianis, H., & Dalianis, H. (2018). *Evaluation metrics and evaluation*.

Das, A. B., Goud, V., Das, C., & Sahu, P. P. (2021). Development of colorimetric pH indicator paper using anthocyanin for rapid quality monitoring of liquid food. *Journal of Packaging Technology and Research, 5*, 41-49. doi:10.1007/s41783-020-00104-x

de Carvalho, L. C., Pereira, F. M. V., de Morais, C. d. L. M., de Lima, K. M. G., & de Almeida Teixeira, G. H. (2019). Assessment of macadamia kernel quality defects by means of near infrared spectroscopy (NIRS) and nuclear magnetic resonance (NMR). *Food Control, 106*, 106695. doi:10.1016/j.foodcont.2019.06.021

de Oliveira, G. A., Bureau, S., Renard, C. M.-G. C., Pereira-Netto, A. B., & de Castilhos, F. (2014). Comparison of NIRS approach for prediction of internal quality traits in three fruit species. *Food chemistry, 143*, 223-230. doi:10.1016/j.foodchem.2013.07.122

de Sousa Marques, A., Nicácio, J. T. N., Cidral, T. A., de Melo, M. C. N., & de Lima, K. M. G. (2013). The use of near infrared spectroscopy and multivariate techniques to differentiate Escherichia coli and Salmonella Enteritidis inoculated into pulp juice. *Journal of microbiological methods, 93*(2), 90-94. doi:10.1016/j.mimet.2013.02.003

Derviş, H. (2019). Bibliometric analysis using bibliometrix an R package. *Journal of scientometric research, 8*(3), 156-160. doi:10.5530/jscires.8.3.32

Devos, O., Ruckebusch, C., Durand, A., Duponchel, L., & Huvenne, J.-P. (2009). Support vector machines (SVM) in near infrared (NIR) spectroscopy: Focus on parameters optimization and model interpretation. *Chemometrics Intelligent Laboratory Systems, 96*(1), 27-33.

Dirks, M., & Poole, D. (2022). Automatic neural network hyperparameter optimization for extrapolation: Lessons learned from visible and near-infrared spectroscopy of mango fruit. *Chemometrics and Intelligent Laboratory Systems, 231*, 104685. doi:10.1016/j.chemolab.2022.104685

Dos Santos, C. A. T., Lopo, M., Páscoa, R. N., & Lopes, J. A. (2013). A review on the applications of portable near-infrared spectrometers in the agro-food industry. *Applied spectroscopy, 67*(11), 1215-1233. doi:10.1366/13-07228

dos Santos Neto, J. P., de Assis, M. W. D., Casagrande, I. P., Júnior, L. C. C., & de Almeida Teixeira, G. H. (2017). Determination of 'Palmer'mango maturity indices using portable near infrared (VIS-NIR) spectrometer. *Postharvest biology and technology, 130*, 75-80.

Egesel, C., & Kahrıman, F. (2012). Determination of quality parameters in maize grain by NIR reflectance spectroscopy. *Journal of Agricultural Sciences, 18*(1), 31-42.

Ehsani, S., Yazdanpanah, H., & Parastar, H. (2023). An innovative screening approach for orange juice authentication using dual portable/handheld NIR spectrometers and chemometrics. *Microchemical Journal, 194*, 109304. doi:10.1016/j.microc.2023.109304

Eisenstecken, D., Stürz, B., Robatscher, P., Lozano, L., Zanella, A., & Oberhuber, M. (2019). The potential of near infrared spectroscopy (NIRS) to trace apple origin: Study on different cultivars and orchard elevations. *Postharvest biology and technology, 147*, 123-131. doi:10.1016/j.postharvbio.2018.08.019

Eldin, A. B., & Akyar, I. (2011). Near infra red spectroscopy. *Wide spectra of quality control. InTech, Rijeka, Croatia*, 237-248. doi:10.5772/24208

Emelike, N., & Akusu, O. (2019). Quality attributes of jams and marmalades produced from some selected tropical fruits. *Journal of Food Processing Technology, 10*(5), 790.

Ephrem, E., Najjar, A., Charcosset, C., & Greige-Gerges, H. (2018). Encapsulation of natural active compounds, enzymes, and probiotics for fruit juice fortification, preservation, and processing: An overview. *Journal of Functional Foods, 48*, 65-84. doi:10.1016/j.jff.2018.06.021

Escribano, S., Biasi, W., Lerud, R., Slaughter, D., & Mitcham, E. (2017). Non-destructive prediction of soluble solids and dry matter content using NIR spectroscopy and its relationship with sensory quality in sweet cherries. *Postharvest biology and technology, 128*, 112-120. doi:10.1016/j.postharvbio.2017.01.016

Essuman, E. K., Teye, E., Sam-Amoah, L. K., & Amuah, C. L. (2023). Rapid and non-destructive prediction of chilli powder integrity by employing pocket-sized NIR spectrometer and chemometrics. *Infrared Physics Technology, 135*, 104961. doi:10.1016/j.infrared.2023.104961

Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., & Pappas, G. (2008). Comparison of PubMed, Scopus, web of science, and Google scholar: strengths and weaknesses. *The FASEB journal, 22*(2), 338-342. doi:10.1096/fj.07-9492LSF

Fan, G., Zha, J., Du, R., & Gao, L. (2009). Determination of soluble solids and firmness of apples by Vis/NIR transmittance. *Journal of Food Engineering, 93*(4), 416-420. doi:10.1016/j.jfoodeng.2009.02.006

Fan, S., Li, C., Huang, W., & Chen, L. (2017). Detection of blueberry internal bruising over time using NIR hyperspectral reflectance imaging with optimum wavelengths. *Postharvest biology and technology, 134*, 55-66. doi:10.1016/j.postharvbio.2017.08.012

Fan, S., Zhang, B., Li, J., Huang, W., & Wang, C. (2016). Effect of spectrum measurement position variation on the robustness of NIR spectroscopy models for soluble solids content of apple. *Biosystems Engineering, 143*, 9-19. doi:10.1016/j.biosystemseng.2015.12.012

Faniadis, D., Drogoudi, P., & Vasilakakis, M. (2010). Effects of cultivar, orchard elevation, and storage on fruit quality characters of sweet cherry (Prunus avium L.). *Scientia horticulturae, 125*(3), 301-304.

Feroz, F. (2018). Changes of microbial load in packet orange juice after expiration date. *Stamford Journal of Microbiology, 8*(1), 27-29. doi:10.3329/sjm.v8i1.42435

Fleming, G. D., Martinez, U., Mallea, M., & Guerra, J. (2014). Raman spectroscopy, DFT computations and SERS induced decomposition of (2-chloroethylphosphonic acid) ethephon. *J. Spectrosc. Dyn, 2014*, 16.

Fu, X., Ying, Y., Lu, H., & Xu, H. (2007). Comparison of diffuse reflectance and transmission mode of visible-near infrared spectroscopy for detecting brown heart of pear. *Journal of Food Engineering, 83*(3), 317-323. doi:10.1016/j.jfoodeng.2007.02.041

Funsueb, S., Thanavanich, C., Theanjumpol, P., & Kittiwachana, S. (2023). Development of new fruit quality indices through aggregation of fruit quality parameters and their predictions using near-infrared spectroscopy. *Postharvest biology and technology, 204*, 112438. doi:10.1016/j.postharvbio.2023.112438

Gabriëls, S. H., Mishra, P., Mensink, M. G., Spoelstra, P., & Woltering, E. J. (2020). Non-destructive measurement of internal browning in mangoes using visible and near-infrared spectroscopy supported by artificial neural network analysis. *Postharvest biology and technology, 166*, 111206. doi:10.1016/j.postharvbio.2020.111206

Ghooshkhaneh, N. G., Golzarian, M. R., & Mollazade, K. (2023). VIS-NIR spectroscopy for detection of citrus core rot caused by Alternaria alternata. *Food Control, 144*, 109320. doi:10.1016/j.foodcont.2022.109320

Giri, N. A., Sakhale, B. K., & Nirmal, N. P. (2023). Functional beverages: an emerging trend in beverage world. *Recent Frontiers of Phytochemicals*, 123-142. doi:10.1016/B978-0-443-19143-5.00002-5

Golic, M., & Walsh, K. (2006). Robustness of calibration models based on near infrared spectroscopy for the in-line grading of stonefruit for total soluble solids content. *Analytica chimica acta, 555*(2), 286-291. doi:10.1016/j.aca.2005.09.014

Gómez, A. H., He, Y., & Pereira, A. G. (2006). Non-destructive measurement of acidity, soluble solids and firmness of Satsuma mandarin using Vis/NIR-spectroscopy techniques. *Journal of Food Engineering, 77*(2), 313-319. doi:10.1016/j.jfoodeng.2005.06.036

Gomiero, T. (2018). Food quality assessment in organic vs. conventional agricultural produce: Findings and issues. *Applied Soil Ecology, 123*, 714-728. doi:10.1016/j.apsoil.2017.10.014

González-Caballero, V., Sánchez, M.-T., Fernández-Novales, J., López, M.-I., & Pérez-Marín, D. (2012). On-vine monitoring of grape ripening using near-infrared spectroscopy. *Food analytical methods, 5*, 1377-1385. doi:10.1007/s12161-012-9389-3

Grabska, J., Beć, K. B., Ueno, N., & Huck, C. W. (2023). Analyzing the quality parameters of apples by spectroscopy from Vis/NIR to NIR region: A comprehensive review. *Foods*, *12*(10), 1946.

Grassi, S., & Alamprese, C. (2018). Advances in NIR spectroscopy applied to process analytical technology in food industries. *Current Opinion in Food Science, 22*, 17-21. doi:10.1016/j.cofs.2017.12.008

Grobelna, A., Kalisz, S., & Kieliszek, M. (2019). The effect of the addition of blue honeysuckle berry juice to apple juice on the selected quality characteristics, anthocyanin stability, and antioxidant properties. *Biomolecules, 9*(11), 744. doi:10.3390/biom9110744

Gullifa, G., Barone, L., Papa, E., Giuffrida, A., Materazzi, S., & Risoluti, R. (2023). Portable NIR spectroscopy: The route to green analytical chemistry. *Frontiers in Chemistry, 11*, 1214825. doi:10.3389/fchem.2023.1214825

Guo, Y., Ni, Y., & Kokot, S. (2016). Evaluation of chemical components and properties of the jujube fruit using near infrared spectroscopy and chemometrics. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 153*, 79-86. doi:10.1016/j.saa.2015.08.006

Guo, Z., Chen, Q., Chen, L., Huang, W., Zhang, C., & Zhao, C. (2011). Optimization of informative spectral variables for the quantification of EGCG in green tea using Fourier transform near-infrared (FT-NIR) spectroscopy and multivariate calibration. *Applied Spectroscopy, 65*(9), 1062-1067. doi:10.1366/11-06287

Guru, D., Raghavendra, A., & Rao, M. K. (2021). *Post-harvest Handling of Mangoes: An Integrated Solution Using Machine Learning Approach.* Paper presented at the International Conference on Computer Vision and Image Processing.

Guthrie, J., Walsh, K., Reid, D., & Liebenberg, C. (2005). Assessment of internal quality attributes of mandarin fruit. 1. NIR calibration model development. *Australian Journal of Agricultural Research, 56*(4), 405-416. doi:10.1071/AR04257

Halepotara, F. H., Kanzaria, D. R., Rajatiya, J. H., Solanki, M. B., & Dodiya, K. (2019). Effect of heat unit and time duration required for maturation of mango (Mangifera indica L.) CV. Kesar. *Journal of Pharmacognosy Phytochemistry, 8*(1), 537-541.

Hamid, F., & Hamid, F. H. (2015). Manual of methods of analysis of foods. *Food safety and standards authority of India*.

Hamzaoui Essoussi, L., & Zahaf, M. (2009). Exploring the decision-making process of Canadian organic food consumers: Motivations and trust issues. *Qualitative Market Research: An International Journal, 12*(4), 443-459. doi:10.1108/13522750910993347

Hao, Y., Wang, Q., & Zhang, S. (2021). Online accurate detection of soluble solids content in navel orange assisted by automatic orientation correction device. *Infrared Physics Technology, 118*, 103871.

Hayati, R., Munawar, A. A., & Fachruddin, F. (2020). Enhanced near infrared spectral data to improve prediction accuracy in determining quality parameters of intact mango. *Data in Brief, 30*, 105571. doi:10.1016/j.dib.2020.105571

He, Y., Xiao, Q., Bai, X., Zhou, L., Liu, F., & Zhang, C. (2022). Recent progress of nondestructive techniques for fruits damage inspection: a review. *Critical reviews in food science and nutrition, 62*(20), 5476-5494.

Heyman, M. B., Abrams, S. A., Heitlinger, L. A., Cabana, M. d., Gilger, M. A., Gugig, R., . . . Corkins, M. R. (2017). Fruit juice in infants, children, and adolescents: current recommendations. *Pediatrics, 139*(6). doi:10.1542/peds.2017-0967

Hidalgo, M. J., Fechner, D. C., Marchevsky, E. J., & Pellerano, R. G. J. F. C. (2016). Determining the geographical origin of Sechium edule fruits by multielement analysis and advanced chemometric techniques. *210*, 228-234. doi:10.1016/j.foodchem.2016.04.120

Hidalgo, M. J., Gaiad, J. E., Goicoechea, H. C., Mendoza, A., Pérez-Rodríguez, M., & Pellerano, R. G. (2023). Geographical origin identification of mandarin fruits by analyzing fingerprint signatures based on multielemental composition. *Food Chemistry: X, 20*, 101040. doi:10.1016/j.fochx.2023.101040

Hong, Y., Liu, Y., Chen, Y., Liu, Y., Yu, L., Liu, Y., & Cheng, H. (2019). Application of fractional-order derivative in the quantitative estimation of soil organic matter content through visible and near-infrared spectroscopy. *Geoderma, 337*, 758-769. doi:10.1016/j.geoderma.2018.10.025

Hoque, M., Talukdar, S., Roy, K. R., Hossain, M. A., & Zzaman, W. (2024). Sonication and thermal treatment of pineapple juice: Comparative assessment of the physicochemical properties, antioxidant activities and microbial inactivation. *Food Science and Technology International, 30*(1), 37-48.

Hu, J., Ma, X., Liu, L., Wu, Y., & Ouyang, J. (2017). Rapid evaluation of the quality of chestnuts using near-infrared reflectance spectroscopy. *Food chemistry, 231*, 141-147. doi:10.1016/j.foodchem.2017.03.127

Huang, H., Yu, H., Xu, H., & Ying, Y. (2008). Near infrared spectroscopy for on/in-line monitoring of quality in foods and beverages: A review. *Journal of Food Engineering, 87*(3), 303-313. doi:10.1016/j.jfoodeng. 2007.12.022

Iqbal, Z., Herodian, S., & Widodo, S. (2019). *Development of partial least square (PLS) prediction model to measure the ripeness of oil palm fresh fruit bunch (FFB) by using NIR Spectroscopy.* Paper presented at the IOP Conference Series: Earth and Environmental Science.

Ishikawa, D., Ueno, G., & Fujii, T. (2017). Estimation Method of Moisture Content at the Meat Surface During Drying Process by NIR Spectroscopy and Its Application for Monitoring of Water Activity. *Japan Journal of Food Engineering, 18*(3), 135-143. doi:10.11301/ jsfe.17493

Islam, M. N., Mursalat, M., & Khan, M. S. (2016). A review on the legislative aspect of artificial fruit ripening. *Agriculture and Food Security, 5*, 1-10. doi:10.1186/s40066-016-0057-5

ISO-4833. (2013). *Microbiology of the food chain - Horizontal method for the enumeration of microorganisms*: International Organization for Standardization.

Izneid, B. A., & Al-kharazi, T. (2013). *Microcontroller based a nondestructive infrared spectroscopy instrument for assessment of mango quality using bio-optic techniques.* Paper presented at the 2013 IEEE International Conference on RFID-Technologies and Applications (RFID-TA).

Jahani, R., Yazdanpanah, H., van Ruth, S. M., Kobarfard, F., Alewijn, M., Mahboubi, A., . . . Salamzadeh, J. (2020). Novel application of near-infrared spectroscopy and chemometrics approach for detection of lime juice adulteration. *Iranian Journal of Pharmaceutical Research: IJPR, 19*(2), 34. doi:10.22037/ijpr.2019.112328.13686

Jahurul, M., Zaidul, I., Ghafoor, K., Al-Juhaimi, F. Y., Nyam, K.-L., Norulaini, N., . . . Omar, A. M. (2015). Mango (Mangifera indica L.) by-products and their valuable components: A review. *Food chemistry, 183*, 173-180. doi:10.1016/j.foodchem.2015.03.046

Janik, L., Cozzolino, D., Dambergs, R., Cynkar, W., & Gishen, M. (2007). The prediction of total anthocyanin concentration in red-grape homogenates using visible-near-infrared spectroscopy and artificial neural networks. *Analytica chimica acta, 594*(1), 107-118. doi:10.1016/j.aca.2007.05.019

Jha, S., Chopra, S., & Kingsly, A. (2007). Modeling of color values for nondestructive evaluation of maturity of mango. *Journal of Food Engineering, 78*(1), 22-26.

Jha, S., & Garg, R. (2010). Non-destructive prediction of quality of intact apple using near infrared spectroscopy. *Journal of food science andtechnology, 47*(2), 207-213. doi:10.1007/s13197-010-0033-1

Jha, S., Kingsly, A., & Chopra, S. (2006). Physical and mechanical properties of mango during growth and storage for determination of maturity. *Journal of Food Engineering, 72*(1), 73-76.

Jha, S. N., Jaiswal, P., Narsaiah, K., Gupta, M., Bhardwaj, R., & Singh, A. K. (2012). Non-destructive prediction of sweetness of intact mango using near infrared spectroscopy. *Scientia Horticulturae, 138*, 171-175. doi:10.1016/j.scienta.2012.02.031

Jha, S. N., Jaiswal, P., Narsaiah, K., Kumar, R., Sharma, R., Gupta, M., . . . Singh, A. K. (2013). Authentication of mango varieties using near-infrared spectroscopy. *Agricultural Research, 2*, 229-235. doi:10.1007/ s40003-013-0068-4

Jiang, H., Liu, G., Mei, C., Yu, S., Xiao, X., & Ding, Y. (2012). Rapid determination of pH in solid-state fermentation of wheat straw by FT-NIR spectroscopy and efficient wavelengths selection. *Analytical bioanalytical chemistry, 404*, 603-611. doi:10.1007/s00216-012-6128-y

Jiao, Y., Li, Z., Chen, X., & Fei, S. (2020). Preprocessing methods for near-infrared spectrum calibration. *Journal of Chemometrics, 34*(11), e3306. doi:https://doi.org/10.1002/cem.3306

Joshi, A. (2016). Comparison between Scopus & ISI Web of science. *Journal Global Values, 7*(1), 1-11.

Kaddumukasa, P. P., Imathiu, S. M., Mathara, J. M., & Nakavuma, J. L. (2017). Influence of physicochemical parameters on storage stability: Microbiological quality of fresh unpasteurized fruit juices. *Food Science and Nutrition, 5*(6), 1098-1105. doi:10.1002/fsn3.500

Kader, A. A. (1997). *Fruit maturity, ripening, and quality relationships.* Paper presented at the International Symposium Effect of Pre-& Postharvest factors in Fruit Storage 485.

Kandpal, L. M., Lohumi, S., Kim, M. S., Kang, J.-S., & Cho, B.-K. (2016). Near-infrared hyperspectral imaging system coupled with multivariate methods to predict viability and vigor in muskmelon seeds. *Sensors and Actuators B: Chemical, 229*, 534-544. doi:10.1016/j.snb.2016.02.015

Kapse, S., Kausley, S., & Rai, B. (2022). Portable food diagnostic devices and methods: A review. *Journal of Food Process Engineering, 45*(11), e14159. doi:10.1111/jfpe.14159

Kayode, R. M.-O., Joshua, V. A., & Oyetoro, M. O. (2023). Effects of drying methods on nutrients and organoleptic properties of dried pawpaw chips. *Croatian journal of food science and technology, 15*(1), 8-15. doi:10.17508/CJFST.2023.15.1.02

Keskin, M., Soysal, Y., Arslan, A., Sekerli, Y., & Celiktas, N. (2018). *Predicting drying temperature of infrared-dried pepper powders using FT-NIRS and chromameter.* Paper presented at the International Conference on Energy Research, November.

Kesse, S., Boakye-Yiadom, K. O., Farooq, M. A., Aquib, M., Filli, M. S., & Bo, W. (2019). Analysis of phosphorus as an impurity from the use of calcium carbide as an artificial ripening agent in banana (Musa acuminate). *Res. Pharm. Heal. Sci., 5*(1), 107-113. doi:10.32463/rphs.2019.v05i01.02

Khadka, D., & Pathak, K. (2023). Spectrophotometric Determination of Total Vitamin C Content in Different Fruits and Vegetables Consumed in Tansen, Palpa. *Tribhuvan Journal, 1*(1), 51-57.

Khalid, N., Suleria, H. A. R., & Ahmed, I. (2016). Pineapple juice. *Handbook of functional beverages and human health, 1*, 489-498.

Khanh Ninh, D., Doan, T.-N.-C., Khanh Ninh, C., Xuan Nguyen-Thi, T., & Le Thanh, N. (2021). *Fruit recognition based on near-infrared spectroscopy using deep neural networks.* Paper presented at the Proceedings of the 2021 5th International Conference on Machine Learning and Soft Computing.

Khumaidi, A., Purwanto, Y. A., Sukoco, H., & Wijaya, S. H. (2022). Using fuzzy logic to increase accuracy in mango maturity index classification: Approach for developing a portable near-infrared spectroscopy device. *Sensors, 22*(24), 9704. doi:10.3390/s22249704

Khumaidi, A., & Raafi'udin, R. (2022). Effects of Oversampling Smote and Spectral Transformations in the Classification of Mango Cultivars Using Near-Infrared Spectroscopy. *Int. J. Adv. Sci. Eng. Inf. Technol, 12*(3), 1047-1053. doi:10.18517/ijaseit.12.3.16001

Kou, X., Feng, Y., Yuan, S., Zhao, X., Wu, C., Wang, C., & Xue, Z. (2021). Different regulatory mechanisms of plant hormones in the ripening of climacteric and non-climacteric fruits: a review. *Plant Molecular Biology*, 1-21. doi:10.1007/s11103-021-01199-9

Kou, X., & Wu, M. (2018). Characterization of climacteric and non-climacteric fruit ripening. *Plant Senescence: Methods and Protocols*, 89-102. doi:10.1007/978-1-4939-7672-0_7

Kusumaningrum, D., Lee, H., Lohumi, S., Mo, C., Kim, M. S., & Cho, B. K. (2018). Non-destructive technique for determining the viability of soybean (Glycine max) seeds using FT-NIR spectroscopy. *Journal of the Science of Food and Agriculture, 98*(5), 1734-1742. doi:10.1002/jsfa.8646

Lakade, A. J., V, V., Ramasamy, R., & Shetty, P. H. (2019). NIR spectroscopic method for the detection of calcium carbide in artificial ripening of mangoes (Magnifera indica). *Food Additives and Contaminants: Part A, 36*(7), 989-995. doi:10.1080/19440049.2019.1605206

Lam, S., Rolland, D., Zawadski, S., Wei, X., Uttaro, B., & Juárez, M. (2023). Performance of a Handheld Near-Infrared Spectroscopy Device to Predict Pork Primal Belly Fat Iodine Value and Loin Lean Intramuscular Fat Content. *Foods, 12*(8), 1629. doi:10.3390/foods12081629

Lammertyn, J., Peirs, A., De Baerdemaeker, J., & Nicolaı, B. (2000). Light penetration properties of NIR radiation in fruit with respect to non-destructive quality assessment. *Postharvest biology and technology, 18*(2), 121-132. doi:10.1016/S0925-5214(99)00071-X

Lamptey, F. P., Amuah, C. L., Boadu, V. G., Abano, E. E., & Teye, E. (2024). Smart classification of organic and inorganic pineapple juice using dual NIR spectrometers combined with chemometric techniques. *Applied Food Research*, 100471. doi:10.1016/j.afres.2024.100471

Lamptey, F. P., Teye, E., Abano, E. E., & Amuah, C. L. (2023). Application of handheld NIR spectrometer for simultaneous identification and quantification of quality parameters in intact mango fruits. *Smart Agricultural Technology, 6*, 100357. doi:10.1016/j.atech.2023.100357

Larraín, M., Guesalaga, A. R., & Agosín, E. (2008). A multipurpose portable instrument for determining ripeness in wine grapes using NIR spectroscopy. *IEEE Transactions on Instrumentation and Measurement, 57*(2), 294-302. doi:10.1109/TIM.2007.910098

Lavanya, E. K., Rao, D. B., Edukondalu, L., Lakshmypathy, R., & Rao, V. S. (2019). Effect of ethephon and storage temperature on physico-chemical changes during ripening of Mango (Mangifera indica L.) Cv. Neelum. *Current Journal of Applied Science and Technology, 38*(6), 1-11. doi:10.9734/CJAST/2019/v38i630387

Lawson, T., Lycett, G. W., Mayes, S., Ho, W. K., & Chin, C. F. (2020). Transcriptome-wide identification and characterization of the Rab GTPase family in mango. *Molecular Biology Reports, 47*(6), 4183-4197.

Le Thi Thu, H., Tran, T., Trinh Thi Phuong, T., Le Thi Tuyet, T., Le Huy, H., & Vu Thi, T. (2021). Two decades of stem education research in middle school: A bibliometrics analysis in scopus database (2000–2020). *Education Sciences, 11*(7), 353. doi:10.3390/educsci11070353

Léchaudel, M., & Joas, J. (2007). An overview of preharvest factors influencing mango fruit growth, quality and postharvest behaviour.

Brazilian Journal of Plant Physiology, 19, 287-298. doi:10.1590/S1677-04202007000400004

Lee, H. J., Jang, D. H., Lee, Y. H., & Kim, Y. H. (2019). Near-infrared hyperspectral imaging for detection of bacterial fruit blotch in watermelon seedlings. *Horticultural science and technology, 37*(6), 719-732. doi:10.7235/HORT.20190072

Lee, J. H., & Choung, M.-G. (2011). Nondestructive determination of herbicide-resistant genetically modified soybean seeds using near-infrared reflectance spectroscopy. *Food Chemistry, 126*(1), 368-373. doi:10.1016/j.foodchem.2010.10.106

Lee, L. C., Liong, C.-Y., & Jemain, A. A. (2018). Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps. *Analyst, 143*(15), 3526-3539. doi:10.1039/C8AN00599K

Lee, S., Noh, T. G., Choi, J. H., Han, J., Ha, J. Y., Lee, J. Y., & Park, Y. (2017). *NIR spectroscopic sensing for point-of-need freshness assessment of meat, fish, vegetables and fruits.* Paper presented at the Sensing for Agriculture and Food Quality and Safety IX.

Lee, W.-H., Kim, M. S., Lee, H., Delwiche, S. R., Bae, H., Kim, D.-Y., & Cho, B.-K. (2014). Hyperspectral near-infrared imaging for the detection of physical damages of pear. *Journal of Food Engineering, 130*, 1-7. doi:10.1016/j.jfoodeng.2013.12.032

Leitner, C., & Vogl, C. R. (2020). Farmers' perceptions of the organic control and certification process in Tyrol, Austria. *Sustainability, 12*(21), 9160. doi:10.3390/su12219160

Li, J., Huang, W., Zhao, C., & Zhang, B. (2013). A comparative study for the quantitative determination of soluble solids content, pH and firmness of pears by Vis/NIR spectroscopy. *Journal of Food Engineering, 116*(2), 324-332. doi:10.1016/j.jfoodeng.2012.11.007

Li, L., Peng, Y., Li, Y., Yang, C., & Chao, K. (2020). Rapid and low-cost detection of moldy apple core based on an optical sensor system. *Postharvest Biology Technology, 168*, 111276. doi:10.1016/j. commatsci.2016.02.021

Li, X., Xie, B., Wu, M., Zhao, J., Xu, Z., & Liu, L. (2021). Visible-to-near-infrared optical properties of protein, lipid and carbohydrate in both solid and solution state at room temperature. *Journal of Quantitative Spectroscopy and Radiative Transfer, 259*, 107410. doi:10.1016/j.jqsrt. 2020.107410

Lin, L., He, Y., Xiao, Z., Zhao, K., Dong, T., & Nie, P. (2019). Rapid-detection sensor for rice grain moisture based on NIR spectroscopy. *Applied Sciences, 9*(8), 1654. doi:10.3390/app9081654

Liu, C., Berry, P. M., Dawson, T. P., & Pearson, R. G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography, 28*(3), 385-393. doi:10.1111/j.0906-7590.2005.03957.x

Liu, F., He, Y., Wang, L., & Sun, G. (2011). Detection of organic acids and pH of fruit vinegars using near-infrared spectroscopy and multivariate calibration. *Food and Bioprocess Technology, 4*, 1331-1340. doi:10.1007/s11947-009-0240-9

Liu, N., Parra, H. A., Pustjens, A., Hettinga, K., Mongondry, P., & Van Ruth, S. M. (2018). Evaluation of portable near-infrared spectroscopy for organic milk authentication. *Talanta, 184*, 128-135. doi:10.1016/j. talanta.2018.02.097

Liu, X., Le Bourvellec, C., Yu, J., Zhao, L., Wang, K., Tao, Y., . . . Hu, Z. (2022). Trends and challenges on fruit and vegetable processing: Insights into sustainable, traceable, precise, healthy, intelligent, personalized and local innovative food products. *Trends in Food Science and Technology, 125*, 12-25. doi:10.1016/j.tifs.2022.04.016

Liu, Y., Sun, X., & Ouyang, A. (2010). Nondestructive measurement of soluble solid content of navel orange fruit by visible–NIR spectrometric technique with PLSR and PCA-BPNN. *LWT-Food Science Technology, 43*(4), 602-607.

Liu, Z., Wang, S., Zhang, Y., Feng, Y., Liu, J., & Zhu, H. (2023). Artificial intelligence in food safety: A decade review and bibliometric analysis. *Foods, 12*(6), 1242. doi:10.3390/foods12061242

Lohumi, S., Lee, S., & Cho, B.-K. (2015). Optimal variable selection for Fourier transform infrared spectroscopic analysis of starch-adulterated garlic powder. *Sensors Actuators B: Chemical, 216*, 622-628.

Londoño, M. B. Z., Chaparro, D., Rojano, B. A., Arbelaez, A. F. A., Betancur, L. F. R., & Celis, M. E. M. (2017). Effect of storage time on physicochemical, sensorial, and antioxidant characteristics, and composition of mango (cv. Azúcar) juice. *Emirates Journal of Food and Agriculture, 29*(5), 367. doi:10.9755/ejfa.2016-09-1256

López, A., Arazuri, S., García, I., Mangado, J., & Jarén, C. (2013). A review of the application of near-infrared spectroscopy for the analysis of potatoes. *Journal of agricultural and food chemistry, 61*(23), 5413-5424. doi:10.1021/jf401292j

Lu, R. (2001). Predicting firmness and sugar content of sweet cherries using near–infrared diffuse reflectance spectroscopy. *Transactions of the ASAE, 44*(5), 1265. doi:10.13031/2013.6421

Ma, X., Luo, H., Zhang, F., & Gao, F. (2022). A bibliometric and visual analysis of fruit quality detection research. *Food Science and Technology, 42*, e72322. doi:10.1590/fst.72322

MacArthur, R. L., Teye, E., & Darkwa, S. (2020). Predicting adulteration of Palm oil with Sudan IV dye using shortwave handheld spectroscopy and comparative analysis of models. *Vibrational Spectroscopy, 110*, 103129. doi:10.1016/j.vibspec.2020.103129

Magwaza, L. S., & Opara, U. L. (2015). Analytical methods for determination of sugars and sweetness of horticultural products—A review. *Scientia Horticulturae, 184*, 179-192. doi:10.1016/j.scienta.2015.01.001

Magwaza, L. S., Opara, U. L., Nieuwoudt, H., Cronje, P. J., Saeys, W., & Nicolaï, B. (2012). NIR spectroscopy applications for internal and external quality analysis of citrus fruit—a review. *Food and Bioprocess Technology, 5*, 425-444. doi:10.1007/s11947-011-0697-1

Magwaza, L. S., & Tesfay, S. Z. (2015). A review of destructive and non-destructive methods for determining avocado fruit maturity. *Food and*

Bioprocess Technology, 8(10), 1995-2011. doi:10.1007/s11947-015-1568-y

Mahanti, N. K., & Chakraborty, S. K. (2020). Application of chemometrics to identify artificial ripening in sapota (Manilkara Zapota) using visible near infrared absorbance spectra. *Computers and Electronics in Agriculture, 175*, 105539. doi:10.1016/j.compag.2020.105539

Malegori, C., Marques, E. J. N., de Freitas, S. T., Pimentel, M. F., Pasquini, C., & Casiraghi, E. (2017). Comparing the analytical performances of Micro-NIR and FT-NIR spectrometers in the evaluation of acerola fruit quality, using PLS and SVM regression algorithms. *Talanta, 165*, 112-116. doi:10.1016/j.talanta.2016.12.035

Malvandi, A., Feng, H., & Kamruzzaman, M. (2022). Application of NIR spectroscopy and multivariate analysis for Non-destructive evaluation of apple moisture content during ultrasonic drying. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 269*, 120733. doi:10.1016/j.saa.2021.120733

Mancini, M., Mazzoni, L., Gagliardi, F., Balducci, F., Duca, D., Toscano, G., . . . Capocasa, F. (2020). Application of the non-destructive NIR technique for the evaluation of strawberry fruits quality parameters. *Foods, 9*(4), 441.

Mantanus, J., Ziemons, E., Lebrun, P., Rozet, E., Klinkenberg, R., Streel, B., . . . Hubert, P. (2009). Moisture content determination of pharmaceutical pellets by near infrared spectroscopy: method development and validation. *Analytica chimica acta, 642*(1-2), 186-192. doi:10. 1016/j.aca.2008.12.031

Marques, E. J. N., de Freitas, S. T., Pimentel, M. F., & Pasquini, C. (2016). Rapid and non-destructive determination of quality parameters in the 'Tommy Atkins' mango using a novel handheld near infrared spectrometer. *Food chemistry, 197*, 1207-1214. doi:10.1016/j.foodchem.2015.11.080

Martínez-Valdivieso, D., Font, R., Blanco-Díaz, M. T., Moreno-Rojas, J. M., Gómez, P., Alonso-Moraga, Á., & Del Río-Celestino, M. (2014). Application of near-infrared reflectance spectroscopy for predicting carotenoid content in summer squash fruit. *Computers and Electronics in Agriculture, 108*, 71-79. doi:10.1016/j.compag.2014.07.003

Masithoh, R. E., Haff, R., & Kawano, S. (2016). Determination of soluble solids content and titratable acidity of intact fruit and juice of satsuma mandarin using a hand-held near infrared instrument in transmittance mode. *Journal of Near Infrared Spectroscopy, 24*(1), 83-88. doi:10.1255/jnirs.1196

McGlone, V. A., Fraser, D. G., Jordan, R. B., & Künnemeyer, R. (2003). Internal quality assessment of mandarin fruit by vis/NIR spectroscopy. *Journal of Near Infrared Spectroscopy, 11*(5), 323-332. doi:10.1255/jnirs.383

Meena, S., Yadav, S., & Meena, M. K. (2020). Melatonin promotes post-harvest ripening and quality of fruits. *Food and Scientific Reports, 1*(7), 64-69.

Mehmood, T., Liland, K. H., Snipen, L., & Sæbø, S. (2012). A review of variable selection methods in partial least squares regression.

Chemometrics Intelligent Laboratory Systems, 118, 62-69. doi:10.1016/j.chemolab.2012.07.010

Mendoza, F., Lu, R., & Cen, H. (2012). Comparison and fusion of four nondestructive sensors for predicting apple fruit firmness and soluble solids content. *Postharvest biology and technology, 73*, 89-98. doi:10.1016/j.postharvbio.2012.05.012

Mie, A., Andersen, H. R., Gunnarsson, S., Kahl, J., Kesse-Guyot, E., Rembiałkowska, E., . . . Grandjean, P. (2017). Human health implications of organic food and organic agriculture: a comprehensive review. *Environmental Health, 16*(111), 1-22. doi:10.1186/s12940-017-0315-4

Mishra, P., & Passos, D. (2021). Realizing transfer learning for updating deep learning models of spectral data to be used in new scenarios. *Chemometrics and Intelligent Laboratory Systems, 212*, 104283. doi:10.1016/j.chemolab.2021.104283

Mishra, P., Rutledge, D. N., Roger, J.-M., Wali, K., & Khan, H. A. (2021). Chemometric pre-processing can negatively affect the performance of near-infrared spectroscopy models for fruit quality prediction. *Talanta, 229*, 122303. doi:10.1016/j.talanta.2021.122303

Mishra, P., Woltering, E., & El Harchioui, N. (2020). Improved prediction of 'Kent'mango firmness during ripening by near-infrared spectroscopy supported by interval partial least square regression. *Infrared Physics and Technology, 110*, 103459. doi:10.1016/j.infrared.2020.103459

Moghimi, A., Aghkhani, M. H., Sazgarnia, A., & Sarmad, M. (2010). Vis/NIR spectroscopy and chemometrics for the prediction of soluble solids content and acidity (pH) of kiwifruit. *Biosystems Engineering, 106*(3), 295-302. doi:10.1016/j.biosystemseng.2010.04.002

Mogollón, R., Contreras, C., da Silva Neta, M. L., Marques, E. J. N., Zoffoli, J. P., & de Freitas, S. T. (2020). Non-destructive prediction and detection of internal physiological disorders in'Keitt'mango using a hand-held Vis-NIR spectrometer. *Postharvest biology and technology, 167*, 111251. doi:10.1016/j.postharvbio.2020.111251

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2010). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *International Journal of Surgery, 8*(5), 336-341. doi:10.1016/j.ijsu.2010.02.007

Molognoni, L., de Sá Ploêncio, L. A., Deolindo, C. T. P., de Oliveira, L. V. A., Hoff, R. B., & Daguer, H. (2020). FT-NIR combined with chemometrics versus classic chemical methods as accredited analytical support for decision-making: Application to chemical compositional compliance of feedingstuffs. *Microchemical Journal, 158*, 105126. doi:10.1016/j.microc.2020.105126

Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics, 106*, 213-228. doi:10.1007/s11192-015-1765-5

Mubarok, S., Sutari, W., & Hadiwijaya, Y. (2021). *Application of spectra pre-treatments on firmness assessment of intact sapodilla using vis-nir spectroscopy.* Paper presented at the IOP Conference Series: Earth and Environmental Science.

Munawar, A., Hayati, R., & Fachruddin, F. (2021). *Rapid determination of inner quality parameters of intact mango fruits using portable near infrared spectroscopy.* Paper presented at the IOP Conference Series: Earth and Environmental Science.

Munawar, A. A., Hayati, R., & Wahyuni, D. (2019). The application of near infrared technology as a rapid and non-destructive method to determine vitamin C content of intact mango fruit. *INMATEH-Agricultural Engineering, 58*(2). doi:10.35633/INMATEH-58-31

Munawar, A. A., Hörsten, D. v., Mörlein, D., Pawelzik, E., & Wegener, J. K. (2013). Rapid and non-destructive prediction of mango sweetness and acidity using near infrared spectroscopy.

Munawar, A. A., Meilina, H., & Pawelzik, E. (2022). Near infrared spectroscopy as a fast and non-destructive technique for total acidity prediction of intact mango: Comparison among regression approaches. *Computers and Electronics in Agriculture, 193*, 106657. doi:10.1016/j.compag.2021.106657

Munawar, A. A., von Hörsten, D., Wegener, J. K., Pawelzik, E., & Mörlein, D. (2016). Rapid and non-destructive prediction of mango quality attributes using Fourier transform near infrared spectroscopy and chemometrics. *Engineering in Agriculture, Environment and Food, 9*(3), 208-215. doi:10.1016/j.eaef.2015.12.004

Nawar, S., & Mouazen, A. (2019). On-line vis-NIR spectroscopy prediction of soil organic carbon using machine learning. *Soil and Tillage Research, 190*, 120-127. doi:10.1016/j.still.2019.03.006

Ncama, K., Magwaza, L. S., Mditshwa, A., & Tesfay, S. Z. (2018). Application of visible to near-infrared spectroscopy for non-destructive assessment of quality parameters of fruit. *Infrared Spectroscopy-Principles, Advances, and Applications*.

Ncama, K., Opara, U. L., Tesfay, S. Z., Fawole, O. A., & Magwaza, L. S. (2017). Application of Vis/NIR spectroscopy for predicting sweetness and flavour parameters of 'Valencia'orange (Citrus sinensis) and 'Star Ruby'grapefruit (Citrus x paradisi Macfad). *Journal of Food Engineering, 193*, 86-94. doi:10.1016/j.jfoodeng.2016.08.015

Ni, P., Niu, H., Tang, Y., Zhang, Y., Zhang, W., Liu, Y., & Lan, H. (2023). Bibliometrics and Visual Analysis of Non-Destructive Testing Technology for Fruit Quality. *Horticulturae, 9*(10), 1091. doi:10.3390/horticulturae9101091

Nicolaï, B. M., Defraeye, T., De Ketelaere, B., Herremans, E., Hertog, M. L., Saeys, W., . . . Verboven, P. (2014). Nondestructive measurement of fruit and vegetable quality. *Annual review of food science and technology, 5*(1), 285-312. doi:10.1146/annurev-food-030713-092410

Ninh, D. K., Phan, K. D., Ninh, C. K., & Le Thanh, N. (2022). Determination of fruit freshness using near-infrared spectroscopy and machine learning techniques. In *Intelligent Systems and Networks: Selected Articles from ICISN 2022, Vietnam* (pp. 455-464): Springer.

Nogales-Bueno, J., Hernández-Hierro, J. M., Rodríguez-Pulido, F. J., & Heredia, F. J. (2014). Determination of technological maturity of grapes and total phenolic compounds of grape skins in red and white cultivars during ripening by near infrared hyperspectral image: A preliminary approach. *Food chemistry, 152*, 586-591. doi:10.1016/ j.foodchem.2013.12.030

Nordey, T., Davrieux, F., & Léchaudel, M. (2019). Predictions of fruit shelf life and quality after ripening: Are quality traits measured at harvest reliable indicators? *Postharvest biology and technology, 153*, 52-60. doi:10.1016/j.postharvbio.2019.03.011

Nowak, D., Gośliński, M., Wojtowicz, E., & Przygoński, K. (2018). Antioxidant properties and phenolic compounds of vitamin C-rich juices. *Journal of Food Science, 83*(8), 2237-2246. doi:10.1111/1750-3841.14284

Ntsoane, M. L., Zude-Sasse, M., Mahajan, P., & Sivakumar, D. (2019). Quality assesment and postharvest technology of mango: A review of its current status and future perspectives. *Scientia Horticulturae, 249*, 77-85.

Nuanmeesri, S., & Poomhiran, L. (2022). Improved Classification of Intact Ripe Mango Sweetness using Fusion Deep Learning and Enhanced Near-Infrared Spectroscopy. *International Journal of Engineering Trends and Technology, 70*(7), 60-67. doi:10.14445/22315381/IJETT-V70I7P207

Nyangena, I., Owino, W., Ambuko, J., & Imathiu, S. (2019). Effect of selected pretreatments prior to drying on physical quality attributes of dried mango chips. *Journal of food science and technology, 56*, 3854-3863. doi:10.1007/s13197-019-03857-9

Odoi-Yorke, F. (2024). A systematic review and bibliometric analysis of electric cooking: evolution, emerging trends, and future research directions for sustainable development. *Sustainable Energy Research, 11*(1), 24. doi:10.1186/s40807-024-00119-x

Oliveira, M., Cruz-Tirado, J., Roque, J., Teófilo, R., & Barbin, D. (2020). Portable near-infrared spectroscopy for rapid authentication of adulterated paprika powder. *Journal of Food Composition and Analysis, 87*, 103403. doi:10.1016/j.jfca.2019.103403

Omar, A. F., Atan, H., & MatJafri, M. Z. (2012). Peak response identification through near-infrared spectroscopy analysis on aqueous sucrose, glucose, and fructose solution. *Spectroscopy letters, 45*(3), 190-201. doi:10.1080/00387010.2011.604065

Omolola, A. O., Jideani, A. I., & Kapila, P. F. (2017). Quality properties of fruits as affected by drying operation. *Critical reviews in food science and nutrition, 57*(1), 95-108. doi:10.1080/10408398.2013.859563

Opara, U. L., & Pathare, P. B. (2014). Bruise damage measurement and analysis of fresh horticultural produce—A review. *Postharvest biology and technology, 91*, 9-24. doi:10.1016/j.postharvbio.2013.12.009

Othman, O. C., & Mbogo, G. (2009). Physico-chemical characteristics of storage-ripened mango (Mangifera indica L.) fruits varieties of eastern Tanzania. *Tanzania Journal of Science, 35*.

Ouyang, Q., Chen, Q., Zhao, J., & Lin, H. (2013). Determination of amino acid nitrogen in soy sauce using near infrared spectroscopy combined with characteristic variables selection and extreme learning machine. *Food and Bioprocess Technology, 6*(9), 2486-2493.

Owino, W. O., & Ambuko, J. L. (2021). Mango fruit processing: Options for small-scale processors in developing countries. *Agriculture, 11*(11), 1105. doi:10.3390/agriculture11111105

Ozaki, Y., Christy, A. A., & McClure, W. F. (2006). *Near-infrared spectroscopy in food science and technology*: John Wiley & Sons.

Padda, M. S., do Amarante, C. V., Garcia, R. M., Slaughter, D. C., & Mitcham, E. (2011). Methods to analyze physico-chemical changes during mango ripening: A multivariate approach. *Postharvest biology and technology, 62*(3), 267-274. doi:10.1016/j.postharvbio.2011.06.002

Padhi, S. R., John, R., Tripathi, K., Wankhede, D. P., Joshi, T., Rana, J. C., . . . Bhardwaj, R. (2024). A Comparison of Spectral Preprocessing Methods and Their Effects on Nutritional Traits in Cowpea Germplasm. *Legume Science, 6*(2), e2977. doi:10.1002/leg3.229

Panigrahi, N., Bhol, C. S., & Das, B. S. (2016). Rapid assessment of black tea quality using diffuse reflectance spectroscopy. *Journal of Food Engineering, 190*, 101-108. doi:10.1016/j.jfoodeng.2016.06.020

Passos, D., & Mishra, P. (2023). Deep Tutti Frutti: Exploring CNN architectures for dry matter prediction in fruit from multi-fruit near-infrared spectra. *Chemometrics and Intelligent Laboratory Systems, 243*, 105023. doi:10.1016/j.chemolab.2023.105023

Paul, V., Pandey, R., & Srivastava, G. C. (2012). The fading distinctions between classical patterns of ripening in climacteric and non-climacteric fruit and the ubiquity of ethylene—An overview. *Journal* *of food science and technology, 49*, 1-21. doi:10.1007/s13197-011-0293-4

Pauziah, M., Malip, M., Norhayati, M., Tham, S., & Ibrahim, M. (2012). *Physical properties and chemical compositions of'Maspine'pineapple at different stages of maturity.* Paper presented at the VII International Postharvest Symposium 1012.

Payasi, A., & Sanwal, G. (2010). Ripening of climacteric fruits and their control. *Journal of food Biochemistry, 34*(4), 679-710. doi:10.1111/j.1745-4514.2009.00307.x

Peng, X., Shi, T., Song, A., Chen, Y., & Gao, W. (2014). Estimating soil organic carbon using VIS/NIR spectroscopy with SVMR and SPA methods. *Remote Sensing, 6*(4), 2699-2717. doi:10.3390/rs6042699

Peng, Y., & Lu, R. (2008). Analysis of spatially resolved hyperspectral scattering images for assessing apple fruit firmness and soluble solids content. *48*(1), 52-62.

Pérez-Marín, D., Paz, P., Guerrero, J.-E., Garrido-Varo, A., & Sánchez, M.-T. (2010). Miniature handheld NIR sensor for the on-site non-destructive assessment of post-harvest quality and refrigerated storage behavior in plums. *Journal of Food Engineering, 99*(3), 294-302. doi:10.1016/j.jfoodeng.2010.03.002

Phuangsombut, K., Phuangsombut, A., & Terdwongworakul, A. (2020). Empirical approach to improve the prediction of soluble solids content in mango using near-infrared spectroscopy. *International Food Research Journal, 27*(2).

Pissard, A., Fernández Pierna, J. A., Baeten, V., Sinnaeve, G., Lognay, G., Mouteau, A., . . . Lateur, M. (2013). Non-destructive measurement of vitamin C, total polyphenol and sugar content in apples using near-infrared spectroscopy. *Journal of the Science of Food and Agriculture, 93*(2), 238-244. doi:10.1002/jsfa.5779

Pivato, S., Misani, N., & Tencati, A. (2008). The impact of corporate social responsibility on consumer trust: the case of organic food. *Business ethics: A European review, 17*(1), 3-12.

Polinar, Y. Q., Yaptenco, K. F., Peralta, E. K., & Agravante, J. U. (2019). Near-infrared spectroscopy for non-destructive prediction of maturity and eating quality of 'Carabao'mango (Mangifera indica L.) fruit. *Agricultural Engineering International: CIGR Journal, 21*(1), 209-219.

Pourdarbani, R., Sabzi, S., Kalantari, D., Karimzadeh, R., Ilbeygi, E., & Arribas, J. I. (2020). Automatic non-destructive video estimation of maturation levels in Fuji apple (Malus Malus pumila) fruit in orchard based on colour (Vis) and spectral (NIR) data. *Biosystems Engineering, 195*, 136-151. doi:10.1016/j.biosystemseng.2020.04.015

Praiphui, A., & Kielar, F. (2023). Comparing the performance of miniaturized near-infrared spectrometers in the evaluation of mango quality. *Journal of Food Measurement and Characterization, 17*(6), 5886-5902. doi:10.1007/s11694-023-02097-y

Purwanto, Y. A., Sari, H. P., & Budiastra, I. W. (2015). Effects of preprocessing techniques in developing a calibration model for soluble solid and acidity in 'Gedong Gincu'mango using NIR spectroscopy.

International Journal of Engineering and Technology, 7(5), 1921-1927.

Puska, P., Kurki, S., Lähdesmäki, M., Siltaoja, M., & Luomala, H. (2018). Sweet taste of prosocial status signaling: When eating organic foods makes you happy and hopeful. *Appetite, 121*, 348-359. doi:10.1016/ j.appet. 2017.11.102

Qi, X., Chen, G., Li, Y., Cheng, X., & Li, C. (2019). Applying neural-network-based machine learning to additive manufacturing: current applications, challenges, and future perspectives. *Engineering, 5*(4), 721-729. doi:10.1016/j.eng.2019.04.012

Qi, Z., Wu, X., Yang, Y., Wu, B., & Fu, H. (2022). Discrimination of the red jujube varieties using a portable NIR spectrometer and fuzzy improved linear discriminant analysis. *Foods, 11*(5), 763.

Qu, J.-H., Liu, D., Cheng, J.-H., Sun, D.-W., Ma, J., Pu, H., & Zeng, X.-A. (2015). Applications of near-infrared spectroscopy in food safety evaluation and control: A review of recent research advances. *Critical reviews in food science and nutrition, 55*(13), 1939-1954. doi:10.1080/10408398.2013.871693

Raghavendra, A., Guru, D., & Rao, M. K. (2021). Mango internal defect detection based on optimal wavelength selection method using NIR spectroscopy. *Artificial Intelligence in Agriculture, 5*, 43-51. doi:10.1016/j.aiia.2021.01.005

Ramírez-Morales, I., Rivero, D., Fernández-Blanco, E., & Pazos, A. (2016). Optimization of NIR calibration models for multiple processes in the sugar industry. *Chemometrics and Intelligent Laboratory Systems, 159*, 45-57. doi:10.1016/j.chemolab.2016.10.003

Ramos, B., Miller, F., Brandão, T., Teixeira, P., & Silva, C. (2013). Fresh fruits and vegetables—An overview on applied methodologies to improve its quality and safety. *Innovative Food Science and Emerging Technologies, 20*, 1-15. doi:10.1016/j.ifset.2013.07.002

Rocha, W. F. d. C., Prado, C. B. d., & Blonder, N. (2020). Comparison of chemometric problems in food analysis using non-linear methods. *Molecules, 25*(13), 3025. doi:10.3390/molecules25133025

Rong, D., Wang, H., Ying, Y., Zhang, Z., & Zhang, Y. (2020). Peach variety detection using VIS-NIR spectroscopy and deep learning. *Computers and Electronics in Agriculture, 175*, 105553. doi:10.1016/j.compag.2020.105553

Rungpichayapichet, P., Chaiyarattanachote, N., Khuwijitjaru, P., Nakagawa, K., Nagle, M., Müller, J., & Mahayothee, B. (2023). Comparison of near-infrared spectroscopy and hyperspectral imaging for internal quality determination of 'Nam Dok Mai'mango during ripening. *Journal of Food Measurement and Characterization, 17*(2), 1501-1514. doi:10.1007/s11694-022-01715-5

Rungpichayapichet, P., Mahayothee, B., Khuwijitjaru, P., Nagle, M., Müller, J., & Analysis. (2015). Non-destructive determination of β-carotene content in mango by near-infrared spectroscopy compared with colorimetric measurements. *Journal of Food Composition, 38*, 32-41.

Rungpichayapichet, P., Mahayothee, B., Nagle, M., Khuwijitjaru, P., & Müller, J. (2016). Robust NIRS models for non-destructive prediction of postharvest fruit ripeness and quality in mango. *Postharvest biology and technology, 111*, 31-40. doi:10.1016/j.postharvbio.2015.07.006

Rungpichayapichet, P., Nagle, M., Yuwanbun, P., Khuwijitjaru, P., Mahayothee, B., & Müller, J. (2017). Prediction mapping of physicochemical properties in mango by hyperspectral imaging. *Biosystems Engineering, 159*, 109-120. doi:10.1016/j.biosystemseng. 2017.04.006

Ruwali, A., Thakuri, M. S., Pandey, S., Mahat, J., & Shrestha, S. (2022). Effect of different ripening agents in storage life of banana (Musa paradisiaca) at Deukhuri, Dang, Nepal. *Journal of Agriculture and Food Research, 10*, 100416. doi:10.1016/j.jafr.2022.100416

Sadler, G. D., & Murphy, P. A. (2010). pH and titratable acidity. *Food analysis, 4*, 219-238.

Saloni, S., Chauhan, K., & Tiwari, S. (2017). Pineapple production and processing in north eastern India. *Journal of pharmacognosy and Phytochemistry, 6*(6S), 665-672.

Sánchez, M.-T., Garrido-Varo, A., Guerrero, J.-E., & Pérez-Marín, D. (2013). NIRS technology for fast authentication of green asparagus grown under organic and conventional production systems. *Postharvest biology and technology, 85*, 116-123. doi:10.1016/j.postharvbio. 2013.05.008

Sandrey, R., & Edinger, H. (2009). The relevance of Chinese agricultural technologies for African smallholder farmers: Agricultural technology research in China. *Stellenbosch: Centre for Chinese Studies*.

Sankaran, S., Mishra, A., Maja, J. M., & Ehsani, R. (2011). Visible-near infrared spectroscopy for detection of Huanglongbing in citrus orchards. *Computers and Electronics in Agriculture, 77*(2), 127-134. doi:10.1016/j.compag.2011.03.004

Santos, C. S., Cruz, R., Goncalves, D. B., Queiros, R., Bloore, M., Kovacs, Z., . . . Casal, S. (2021). Non-destructive measurement of the internal quality of citrus fruits using a portable NIR device. *Journal of AOAC International, 104*(1), 61-67. doi:10.1093/jaoacint/qsaa115

Santos, J. R., Sarraguça, M. C., Rangel, A. O., & Lopes, J. A. (2012). Evaluation of green coffee beans quality using near infrared spectroscopy: A quantitative approach. *Food chemistry, 135*(3), 1828-1835.

Saranraj, P., & Geetha, M. (2012). Microbial spoilage of bakery products and its control by preservatives. *International Journal of Pharmaceutical biological archives, 3*(1), 38-48.

Saranwong, I., Sornsrivichai, J., & Kawano, S. (2003). On-tree evaluation of harvesting quality of mango fruit using a hand-held NIR instrument. *Journal of Near Infrared Spectroscopy, 11*(4), 283-293. doi:10.1255/jnirs.374

Saranwong, S., Sornsrivichai, J., & Kawano, S. (2004). Prediction of ripe-stage eating quality of mango fruit from its harvest quality measured nondestructively by near infrared spectroscopy. *Postharvest biology and technology, 31*(2), 137-145. doi:10.1016/j.postharvbio.2003.08.007

Schmilovitch, Z. e., Mizrach, A., Hoffman, A., Egozi, H., & Fuchs, Y. (2000). Determination of mango physiological indices by near-infrared spectrometry. *Postharvest biology and technology, 19*(3), 245-252.

Seehanam, P., Chaiya, P., Theanjumpol, P., Tiyayon, C., Ruangwong, O., Pankasemsuk, T., . . . Maniwara, P. (2022). Internal disorder evaluation of 'Namdokmai Sithong'mango by near infrared spectroscopy. *Horticulture, Environment, and Biotechnology, 63*(5), 665-675. doi:10.1007/s13580-022-00435-5

Seki, H., Murakami, H., Ma, T., Tsuchikawa, S., & Inagaki, T. (2024). Evaluating Soluble Solids in White Strawberries: A Comparative Analysis of Vis-NIR and NIR Spectroscopy. *Foods, 13*(14), 2274. doi:10.3390/foods13142274

Shah, S. S. A., Zeb, A., Qureshi, W. S., Arslan, M., Malik, A. U., Alasmary, W., & Alanazi, E. (2020). Towards fruit maturity estimation using NIR spectroscopy. *Infrared Physics and Technology, 111*, 103479. doi:10.1016/j.infrared.2020.103479

Shah, S. S. A., Zeb, A., Qureshi, W. S., Malik, A. U., Tiwana, M., Walsh, K., . . . Alanazi, E. (2021). Mango maturity classification instead of maturity index estimation: A new approach towards handheld NIR spectroscopy. *Infrared Physics and Technology, 115*, 103639. doi:10.1016/j.infrared.2021.103639

Shaik, L., & Chakraborty, S. (2022). Nonthermal pasteurization of pineapple juice: A review on the potential of achieving microbial safety and enzymatic stability. *Comprehensive Reviews in Food Science Food Safety, 21*(6), 4716-4737. doi:10.1111/1541-4337.13042

Siddiqui, M. W., & Dhua, R. (2010). Eating artificially ripened fruits is harmful. *Current science*, 1664-1668.

Siedliska, A., Baranowski, P., Zubik, M., Mazurek, W., & Sosnowska, B. (2018). Detection of fungal infections in strawberry fruit by VNIR/SWIR hyperspectral imaging. *Postharvest Biology Technology, 139*, 115-126.

Singh, S. K., & Sharma, M. (2017). Review on biochemical changes associated with storage of fruit juice. *Int. J. Curr. Microbiol. Appl. Sci, 6*(8), 236-245. doi:10.20546/ijcmas.2017.608.032

Šnurkovič, P. (2013). Quality assessment of fruit juices by NIR spectroscopy. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis, 61*(3), 803-812.

Sohrabi, C., Franchi, T., Mathew, G., Kerwan, A., Nicola, M., Griffin, M., . . . Agha, R. (2021). PRISMA 2020 statement: what's new and the importance of reporting guidelines. *International Journal of Surgery, 88*, 105918. doi:10.1016/j.ijsu.2021.105918

Solihin, M. I., Zekui, Z., Ang, C. K., Heltha, F., & Rizon, M. (2021). *Machine learning calibration for near infrared spectroscopy data: a visual programming approach.* Paper presented at the Proceedings of the 11th National Technical Seminar on Unmanned System Technology 2019: NUSYS'19.

Song, W., Wang, H., Maguire, P., & Nibouche, O. (2016). *Differentiation of organic and non-organic apples using near infrared reflectance spectroscopy—A pattern recognition approach.* Paper presented at the 2016 IEEE SENSORS.

Souza, J., Leonel, S., Modesto, J., Ferraz, R., & Gonçalves, B. (2018). Fruit physicochemical and antioxidant analysis of mango cultivars under subtropical conditions of Brazil. *Journal of Agricultural Science and Technology, 20*(2), 321-331.

Statista. (2023). Worldwide juice market. Accessed September 29, 2023. https://www.statista.com/outlook/20030000/100/juices/worldwide.

Stuart, B. H. (2004). *Infrared spectroscopy: fundamentals and applications*: John Wiley & Sons.

Subedi, P., & Walsh, K. B. (2011). Assessment of sugar and starch in intact banana and mango fruit by SWNIR spectroscopy. *Postharvest Biology Technology, 62*(3), 238-245. doi:10.1016/j.postharvbio.2011.06.014

Sun, T., Lin, H., Xu, H., & Ying, Y. (2009). Effect of fruit moving speed on predicting soluble solids content of 'Cuiguan'pears (Pomaceae pyrifolia Nakai cv. Cuiguan) using PLS and LS-SVM regression. *Postharvest biology and technology, 51*(1), 86-90. doi:10.1016/j.post harvbio.2008.06.003

Sun, Y., Wang, Y., Huang, J., Ren, G., Ning, J., Deng, W., . . . Zhang, Z. (2020). Quality assessment of instant green tea using portable NIR spectrometer. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 240*, 118576.

Taira, E., Nakamura, S., Hiyane, R., Honda, H., & Ueno, M. (2017). Development of a nondestructive measurement system for mango fruit using near infrared spectroscopy. *Engineering and Applied Science Research, 44*(3), 189-192. doi:10.14456/easr.2017.28

Tan, Y. P., & Chia, K. S. (2023). *Effects of Pre-Processing and Principal Components for Artificial Neural Network in Non-Destructive Internal Quality Prediction of Mango across Different Harvest Periods.* Paper presented at the 2023 IEEE 13th International Conference on Control System, Computing and Engineering (ICCSCE).

Tasnim, F., Hossain, M. A., Nusrath, S., Hossain, M. K., Lopa, D., & Haque, K. F. (2010). Quality assessment of industrially processed fruit juices available in dhaka city, bangladesh. *Mal J Nutr, 16*(3), 431 - 438.

Teena, M., Manickavasagan, A., Ravikanth, L., & Jayas, D. (2014). Near infrared (NIR) hyperspectral imaging to classify fungal infected date fruits. *Journal of stored products research, 59*, 306-313. doi:10.1016/ j.jspr.2014.09.005

Tellspec. (2020). FAQ – Tellspec. Retrieved from https://tellspec.com/faq/

Teye, E. (2022). Mini shortwave spectroscopic techniques and multivariate statistical analysis as a tool for testing intact cocoa beans at farmgate for quality control in Ghana. *Infrared Physics Technology, 122*, 104092. doi:10.1016/j.infrared.2022.104092

Teye, E., & Amuah, C. L. (2022). Rice varietal integrity and adulteration fraud detection by chemometrical analysis of pocket-sized NIR spectra data. *Applied Food Research, 2*(2), 100218. doi:10.1016/j.afres.2022. 100218

Teye, E., Amuah, C. L., Atiah, K., Darko, R. O., Abindaw, T., Amoah, K. K., . . . Owusu, R. (2023). Quick Determination of Soil Quality Using Portable Spectroscopy and Efficient Multivariate Techniques. *Journal of Spectroscopy, 2023*. doi:10.1155/2023/2024318

Teye, E., Amuah, C. L., McGrath, T., & Elliott, C. (2019). Innovative and rapid analysis for rice authenticity using hand-held NIR spectrometry and chemometrics. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 217*, 147-154. doi:10.1016/j.saa.2019.03.085

Teye, E., Anyidoho, E., Agbemafle, R., Sam-Amoah, L. K., & Elliott, C. (2020). Cocoa bean and cocoa bean products quality evaluation by NIR spectroscopy and chemometrics: A review. *Infrared Physics and Technology, 104*, 103127.

Teye, E., Elliott, C., Sam-Amoah, L. K., & Mingle, C. (2019). Rapid and nondestructive fraud detection of palm oil adulteration with Sudan dyes using portable NIR spectroscopic techniques. *Food Additives and Contaminants: Part A, 36*(11), 1589-1596. doi:10.1080/19440049.2019.1658905

Teye, E., Huang, X., Sam-Amoah, L. K., Takrama, J., Boison, D., Botchway, F., & Kumi, F. (2015). Estimating cocoa bean parameters by FT-NIRS and chemometrics analysis. *Food Chemistry, 176*, 403-410. doi:10.1016/j.foodchem.2014.12.042

Thanh Noi, P., & Kappas, M. (2017). Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery. *Sensors, 18*(1), 18. doi:10.3390/s18010018

Tharanathan, R., Yashoda, H., & Prabha, T. (2006). Mango (Mangifera indica L.),"The king of fruits"—An overview. *Food Reviews International, 22*(2), 95-123. doi:https://doi.org/10.1080/87559120600574493

Theanjumpol, P., Noimanee, P., Pattana, S., Natwichai, J., Chang-Rue, V., & Boonyakiat, D. (2012). Tangerine and Mango Fruits Quality Detection by Near Infrared Spectroscopy. *Chiang Mai University Journal of Natural Sciences*.

Theanjumpol, P., Self, G., Rittiron, R., Pankasemsuk, T., & Sardsud, V. (2014). Quality control of mango fruit during postharvest by near infrared spectroscopy. *Chiang Mai University Journal of Natural Sciences, 13*(2), 141-157. doi:10.12982/cmujns.2014.0028

Toledo-Martín, E. M., García-García, M. d. C., Font, R., Moreno-Rojas, J. M., Salinas-Navarro, M., Gómez, P., & del Río-Celestino, M. (2018). Quantification of total phenolic and carotenoid content in blackberries (Rubus Fructicosus L.) using near infrared spectroscopy (NIRS) and multivariate analysis. *Molecules, 23*(12), 3191. doi:10.3390/molecules 23123191

Torres, I., Sánchez, M.-T., Vega-Castellote, M., & Pérez-Marín, D. (2021). Fraud detection in batches of sweet almonds by portable near-infrared spectral devices. *Foods, 10*(6), 1221. doi:10.3390/foods10061221

Tres, A., Van Der Veer, G., Perez-Marin, M., Van Ruth, S., & Garrido-Varo, A. (2012). Authentication of organic feed by near-infrared spectroscopy combined with chemometrics: a feasibility study. *Journal of agricultural and food chemistry, 60*(33), 8129-8133. doi:10.1021/ jf302309t

Tu, N. T. (2019). Genetic characterization of mango accessions through RAPD and ISSR markers in Vietnam. *SABRAO Journal of Breeding Genetics, 51*(3).

Tyl, C., & Sadler, G. D. (2017). pH and titratable acidity. *Food analysis*, 389-406. doi:10.1007/978-3-319-45776-5_22

Ulya, M., Chamidah, N., & Saifudin, T. (2021). *Predicting the sweetness level of avomango (Gadung Klonal 21) using multi-predictor local polynomial regression.* Paper presented at the IOP Conference Series: Earth and Environmental Science.

Vallone, M., Alleri, M., Bono, F., & Catania, P. (2019). Quality evaluation of grapes for mechanical harvest using vis NIR spectroscopy. *Agricultural Engineering International: CIGR Journal*, *21*(1), 140-149.

Vantarakis, A., Affifi, M., Kokkinos, P., Tsibouxi, M., & Papapetropoulou, M. (2011). Occurrence of microorganisms of public health and spoilage significance in fruit juices sold in retail markets in Greece. *Anaerobe, 17*(6), 288-291. doi:10.1016/j.anaerobe.2011.04.005

Varith, J., Hyde, G., Baritelle, A., Fellman, J., & Sattabongkot, T. (2003). Non-contact bruise detection in apples by thermal imaging. *Innovative Food Science and Emerging Technologies, 4*(2), 211-218. doi:10.1016/ S1466-8564(03)00021-3

Walsh, K., Golic, M., & Greensill, C. (2004). Sorting of fruit using near infrared spectroscopy: application to a range of fruit and vegetables for soluble solids and dry matter content. *Journal of Near Infrared Spectroscopy, 12*(3), 141-148. doi:10.1255/jnirs.419

Walsh, K., McGlone, V., & Han, D. (2020). The uses of near infra-red spectroscopy in postharvest decision support: A review. *Postharvest* *biology and technology, 163*, 111139. doi:10.1016/j.postharvbio.2020.111139

Wang, C., & Si, L. (2023). A bibliometric analysis of digital literacy research from 1990 to 2022 and research on emerging themes during the Covid-19 pandemic. *Sustainability, 15*(7), 5769. doi:10.3390/su15075769

Wang, G., Zou, X., Weng, J., Lan, G., Li, M., Wei, J., . . . Luo, H. (2024). Bibliometric analysis reveals the research hotspots and trends of nasopharyngeal carcinoma immunotherapy. *Human Vaccines and Immunotherapeutics, 20*(1), 2360341. doi:10.1080/21645515.2024.2360341

Wang, L., Sun, D.-W., Pu, H., & Cheng, J.-H. (2017). Quality analysis, classification, and authentication of liquid foods by near-infrared spectroscopy: A review of recent research developments. *Critical reviews in food science and nutrition, 57*(7), 1524-1538. doi:10.1080/10408398.2015.1115954

Wang, N.-N., Sun, D.-W., Yang, Y.-C., Pu, H., & Zhu, Z. (2016). Recent advances in the application of hyperspectral imaging for evaluating fruit quality. *Food analytical methods, 9*, 178-191. doi:10.1007/s12161-015-0153-3

Wang, Z. X., He, Q. P., & Wang, J. (2015). Comparison of variable selection methods for PLS-based soft sensor modeling. *Journal of Process Control, 26*, 56-72. doi:10.1016/j.jprocont.2015.01.003

Watanawan, C., Wasusri, T., Srilaong, V., Wongs-Aree, C., & Kanlayanarat, S. (2014). Near infrared spectroscopic evaluation of fruit maturity and quality of export Thai mango (Mangifera indica L. var. Namdokmai). *International Food Research Journal, 21*(3), 1073.

Watanawan, C., Wasusri, T., Wongs-Aree, C., Srilaong, V., & Kanlayanarat, S. (2012). *Harvest maturity determination for export mango (Mangifera indica L.'Nam Dok Mai').* Paper presented at the Southeast Asia Symposium on Quality Management in Postharvest Systems and Asia Pacific Symposium on Postharvest Quality 989.

Wedding, B. B., White, R. D., Grauf, S., Wright, C., Tilse, B., Hofman, P., & Gadek, P. A. (2011). Non-destructive prediction of 'Hass' avocado dry matter via FT-NIR spectroscopy. *Journal of the Science of Food and Agriculture, 91*(2), 233-238. doi:10.1002/jsfa.4175

Wibowo, S., Grauwet, T., Santiago, J. S., Tomic, J., Vervoort, L., Hendrickx, M., & Van Loey, A. (2015). Quality changes of pasteurised orange juice during storage: A kinetic study of specific parameters and their relation to colour instability. *Food chemistry, 187*, 140-151. doi:10.1016/j.foodchem.2015.03.131

Williams, P. (2001). Implementation of near-infrared technology. *Near-infrared technology in the agricultural food industries, eds. P. Williams, and K. Norris, 145–169. St. Paul, Minnesota: American Association of Cereal Chemists.*

Williams, P. (2014). The RPD statistic: A tutorial note. *NIR news, 25*(1), 22-26.

Włodarska, K., Szulc, J., Khmelinskii, I., & Sikorska, E. (2019). Non-destructive determination of strawberry fruit and juice quality parameters using ultraviolet, visible, and near-infrared spectroscopy.

Journal of the Science of Food Agriculture, 99(13), 5953-5961. doi:10.1002/jsfa.9870

Wohlers, M., McGlone, A., Frank, E., & Holmes, G. (2023). Augmenting NIR Spectra in deep regression to improve calibration. *Chemometrics and Intelligent Laboratory Systems, 240*, 104924. doi:10.1016/j.chemolab.2023.104924

Wojdyło, A., Teleszko, M., & Oszmiański, J. (2014). Physicochemical characterisation of quince fruits for industrial use: Yield, turbidity, viscosity and colour properties of juices. *International Journal of Food Science and Technology, 49*(8), 1818-1824. doi:10.1111/ijfs.12490

Wu, D., He, Y., & Feng, S. (2008). Short-wave near-infrared spectroscopy analysis of major compounds in milk powder and wavelength assignment. *Analytica chimica acta, 610*(2), 232-242. doi:10.1016/j.aca.2008.01.056

Wu, D., He, Y., Nie, P., Cao, F., & Bao, Y. (2010). Hybrid variable selection in visible and near-infrared spectral analysis for non-invasive quality determination of grape juice. *Analytica chimica acta, 659*(1-2), 229-237. doi:10.1016/j.aca.2009.11.045

Wu, D., Shi, H., Wang, S., He, Y., Bao, Y., & Liu, K. (2012). Rapid prediction of moisture content of dehydrated prawns using online hyperspectral imaging system. *Analytica chimica acta, 726*, 57-66. doi:10.1016/j.aca.2012.03.038

Xiao, R., Liu, L., Zhang, D., Ma, Y., & Ngadi, M. O. (2019). Discrimination of organic and conventional rice by chemometric analysis of NIR spectra: a pilot study. *Journal of Food Measurement and Characterization, 13*, 238-249. doi:10.1007/s11694-018-9937-7

Xiaobo, Z., Jiewen, Z., Povey, M. J., Holmes, M., & Hanpin, M. (2010). Variables selection methods in near-infrared spectroscopy. *Analytica chimica acta, 667*(1-2), 14-32. doi:10.1016/j.aca.2010.03.048

Xie, L., Ye, X., Liu, D., & Ying, Y. (2009). Quantification of glucose, fructose and sucrose in bayberry juice by NIR and PLS. *Food chemistry, 114*(3), 1135-1140. doi:10.1016/j.foodchem.2008.10.076

Xie, L., Ying, Y., Ying, T., Yu, H., & Fu, X. (2007). Discrimination of transgenic tomatoes based on visible/near-infrared spectra. *Analytica chimica acta, 584*(2), 379-384. doi:10.1016/j.aca.2006.11.071

Xiong, T., & Cherkassky, V. (2005). *A combined SVM and LDA approach for classification.* Paper presented at the Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.

Xu, H., Qi, B., Sun, T., Fu, X., & Ying, Y. (2012). Variable selection in visible and near-infrared spectra: Application to on-line determination of sugar content in pears. *Journal of Food Engineering, 109*(1), 142-147. doi:10.1016/j.jfoodeng.2011.09.022

Xu, X., Li, X., Qi, G., Tang, L., & Mukwereza, L. (2016). Science, technology, and the politics of knowledge: The case of China's agricultural technology demonstration centers in Africa. *World development, 81*, 82-91. doi:10.1016/j.worlddev.2016.01.003

Yahia, E., Ornelas-Paz, J. D. J., & Elansari, A. (2011). Postharvest technologies to maintain the quality of tropical and subtropical fruits.

In *Postharvest biology and technology of tropical and subtropical fruits* (pp. 142-195e): Elsevier.

Yahia, E. M., de Jesús Ornelas-Paz, J., Brecht, J. K., García-Solís, P., & Celis, M. E. M. (2023). The contribution of mango fruit (Mangifera indica L.) to human nutrition and health. *Arabian Journal of Chemistry, 16*(7), 104860. doi:10.1016/j.arabjc.2023.104860

Yang, J., Luo, X., Zhang, X., Passos, D., Xie, L., Rao, X., . . . Ying, Y. (2022). A deep learning approach to improving spectral analysis of fruit quality under interseason variation. *Food Control, 140*, 109108. doi:10.1016/j.foodcont.2022.109108

Yegon, D. K. (2023). *Evaluation of Portable Near-Infrared Spectrometer for Rapid and Non-Destructive Determination of Quality and Authenticity of Baobab (Adansonia Digitata l.) Fruit Pulp in Kenya.* JKUAT-CoANRE,

Yin, L., Zhou, J., Chen, D., Han, T., Zheng, B., Younis, A., & Shao, Q. (2019). A review of the application of near-infrared spectroscopy to rare traditional Chinese medicine. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 221*, 117208. doi:10.1016/ j.saa.2019.117208

Zakaria, N., Abdullah, A., Saad, F. A., Hassan, M. A., Bakar, M. A., Saad, A. M., . . . Ibrahim, M. (2021). *Potential of Near-Infrared (NIR) spectroscopy technique for early detection of Insidious Fruit Rot (IFR) disease in Harumanis mango.* Paper presented at the Journal of Physics: Conference Series.

Zeb, A., Qureshi, W. S., Ghafoor, A., & O'Sullivan, D. (2022). *Learning fruit class from short wave near infrared spectral features, an AI approach towards determining fruit type.* Paper presented at the 2022 8th International Conference on Mechatronics and Robotics Engineering (ICMRE).

Zhang, C., Wu, W., Zhou, L., Cheng, H., Ye, X., & He, Y. (2020). Developing deep learning based regression approaches for determination of chemical compositions in dry black goji berries (Lycium ruthenicum Murr.) using near-infrared hyperspectral imaging. *Food chemistry, 319*, 126536. doi:10.1016/j.foodchem.2020.126536

Zhao, Q., & Huang, J. (2011). *Agricultural science & technology in China: a roadmap to 2050*: Springer.

Zorn, A., Lippert, C., & Dabbert, S. (2009). Economic concepts of organic certification.