Machine Learning Driven Non-Invasive Approach for the Detection of Anomalies in Living Plant Leaves and Water at Cellular Level Using Terahertz Sensing

Adnan Zahid

Submitted in fulfilment of the requirements for the
Degree of Doctor of Philosophy

James Watt School of Engineering
College of Science and Engineering
University of Glasgow

October 2020

Abstract

In recent times, an increasing global aridification due to climate transformations and unceasing expansion of population have posed enormous challenges on the environment and its agricultural provision. Researchers and scientists are faced with significant challenges to enhance yield while facing shortage of fertile land due to environmental changes. In this regard, many technologies have been employed to monitor and enhance the crops production. However, certain limitations such as low resolution, destructive nature, cost, sensitivity and reactive nature of technology have markedly reduce their application in modern agriculture. The mounting pressure of more yield with limited fertile land due to environmental changes demands for proactive, cost-effective, real-time, feasible and non-destructive technique in perpetual plants' health monitoring in order to maintain a healthy physiological status of plants leaves, and to drive the crops productivity and achieve economic benefits. With this motivation in mind, we potentially highlight the evolving application of terahertz (THz) technology (due to its non-ionising and less pervasive radiation properties) with machine learning (ML) for the proactive vegetation monitoring.

In this thesis, we proposed a novel, non-invasive, and cost-effective technique to characterise and estimate the real-time information of water contents (WC) in plants leaves and fruits at cellular level in terms of electromagnetic parameters at THz frequency range from 0.75 to 1.1 THz. . It is was noticed that loss observed in WC on day 1 was in the range of 5% to 22%, and increased from 83.12% to 99.33% on day 4. Furthermore, we observed an exponential decaying trend in the peaks of the real part of the permittivity from day 1 to 4, which was reminiscent of the trend observed in the weight of all leaves.

The study also highlights the proactive approach by integrating THz with ML for the accurate and precise estimation of WC in plants and fruits slices including apple and mango, respectively. The results obtained from the amalgamation of ML with THz for the estimation of WC in plants leaves demonstrated that support vector machine (SVM) outperformed other classifiers using tenfold and leave-one-observations out cross-validation for different days classification with an overall accuracy of 98.8%, 97.15%, and 96.82% for Coffee, pea shoot, and baby spinach leaves respectively. In addition, using sequential forward selection (SFS) technique, coffee leaf showed a significant improvement of 15%, 11.9%, 6.5% in computational time for SVM, K-nearest neighbour (KNN) and Decision-tree (D-Tree). For pea-shoot, 21.28%, 10.01%, and 8.53% of improvement was noticed in operating time for SVM, KNN and D-Tree classifiers, respectively. Lastly, baby spinach leaf exhibited a further improvement of 21.28% in SVM, 10.01% in KNN, and 8.53% in D-tree in overall operating time for classifiers. The results illustrated that the performance of SVM exceeded other classifiers results using 10-fold validation and leave-one-observation-out-cross-validation techniques. Moreover, all three classifiers exhibited 100% accuracy for day 1 and 4 with 80% Moisture content (MC) value (freshness) and 2% MC value (staleness) of both fruits' slices, respectively. Similarly, for day 2 and 3, an accuracy of 95% was achieved with intermediate MC values in both fruits' slices.

In addition, in this work, the preservation of clean water without any harmful impurities is also addressed for the health, environmental protection, and economic development. For this purpose, a realistic technological solution method and application of Fourier transform Infrared Spectroscopy (FTIR) operates at THz waves enabled by ML is also discussed in detail. The suggested technique can provide the approximate prediction and detection of even the smallest of contaminants in distilled water due to high sensitivity and non-destructive nature and also produce high optical throughput. Moreover, it was found that random forest (RF) with 97.98%, outperformed other classifiers for estimation of salts concentration added in aqueous solutions. However, for sugar and glucose concentrations, SVM exhibited a higher accuracy of 93.11% and 96.88%, respectively, compared to other classifiers

The proposed novel study using THz wave and incorporating ML are beneficial and provide prolific recommendations, and insights for cultivators, and horticulturists to take proactive actions in relations to both vegetation and water health monitoring, which in turn, can help in reducing the health and purification expenses by providing early alerts to protect the public health, increase yield with limited land, which will ultimately optimise economic benefits.

Statement of Originality
University of Glasgow
James Watt School of Engineering

Name: Adnan Zahid
Registration Number:

I certify that the thesis presented here for examination for a PhD degree of the University of Glasgow is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it) and that the thesis has not been edited by a third party beyond what is permitted by the University's PGR Code of Practice.

I declare that the thesis does not include work forming part of a thesis presented successfully for another degree.

I declare that this thesis has been produced in accordance with the University of Glasgow's Code of Good Practice in Research.

I acknowledge that if any issues are raised regarding good research practice based on review of the thesis, the examination may be postponed pending the outcome of any investigation of the issues.

Signature: ....
Date: ....

List of Publications

Journals

1. **Zahid, A.**; T. Abbas, H.; Imran, M.A.; Qaraqe, K.A.; Alomainy, A.; Cumming, D.R.S.; Abbasi, Q.H., "Characterization and Water Content Estimation Method of Living Plant Leaves Using Terahertz Waves ", *Appl. Sci. 2019, 9*, 2781 **(Published)**.

2. Ren, A., **Zahid, A.**, Dou, F., Imran, M.A., Alomainy, A., Abbasi, Q. H, "State-of-the-art in terahertz sensing for food and water security – A comprehensive review", *Trends in Food Science & Technology, Vol. 85*, pp. 241-251, March 2019 **(Published)**.

3. **Zahid, A.**, Abbas, H.T., Ren, A. et al, "Machine learning driven non-invasive approach of water content estimation in living plant leaves using terahertz waves", *Plant Methods 15*, 138 (2019) **(Published)**.

4. A. Ren, **Zahid, A.** et al., "Machine Learning Driven Approach Towards the Quality Assessment of Fresh Fruits Using Non-Invasive Sensing", *IEEE Sensors Journal, vol. 20, no. 4*, pp. 2075-2083, 15 Feb.15, 2020 **(Published)**.

5. **Zahid, A**, Kia Dashtipour, Muath Al-Hasan, Ismail Ben Mabrouk, Muhammad A. Imran, Akram Alomainy, Qammer H. Abbasi, "Machine Learning Enabled Identification and Real-time Prediction of Living Plants' Stress Using Terahertz Waves", *IEEE Sensors Journal*, **(Under Review)**.

6. M. K. Sott, L. B. Furstenau, L. M. Kipper, Ê. L. Machado, J. R. López-Robles, M. S. Dohan, M. J Cobo, **A. Zahid**, Q. H. Abbasi, M. A. Imran, "Precision Techniques and Agriculture 4.0 Technologies to Promote Sustainability in the Coffee Sector: State of the Art, Challenges and Future Trends," *in IEEE Access, vol. 8, pp. 149854-149867*, 2020 **(Published)**.

7. **A. Zahid**, Kia Dashtipour, Ivonne E. Carranza1, Hasan T. Abbas, Muhammad A. Imran, David R. S. Cumming, James P. Grant, Qammer H. Abbasi, "Machine Learning Framework for the Detection of Anomalies in Aqueous Solutions Using Terahertz Waves", *Sci Reports*, nature, **(To be Submitted)**.

Conference Proceedings

1. **Zahid, A.**, Abbas H.T.,Carranza Ivonne E., James P. Grant James P., Imran, Muhammad A., Cumming, David R. S., and Abbasi.Qammer H., "Assessing the Salt Constituents Characteristics in Aqueous Solutions Using Terahertz Waves", *IEEE International Symposium on Antennas and Propagation and North American Radio Science Meeting in Montréal, Québec*, Canada on 5-10 July 2020.(**Accepted for Publication)**

2. **Zahid, A.**, and Abbasi, Q. H.,(2018), "Electromagnetic Properties of Plant Leaves at Terahertz Frequencies for Health Status Monitoring", *IEEE MTT-S 2019 International Microwave Biomedical Conference (IMBioC-2019), Nanjing, China, 6-8* May 2019,**(Published)**.

3. **Zahid, A.**, Abbas, H. T., Sheikh, F, Kaiser, T., Zoha, A., Imran, M.A., Abbasi, Q.H., (2018) "Monitoring Health Status and Quality Assessment of Leaves Using Terahertz Frequency", *IEEE 2019 Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting, Atlanta, GA, USA*, 07-12 Jul 2019, **(Published)**.

4. **Zahid, A.**, Abbas, H.T., Alomainy, A., Imran, M.A., Abbasi, Q.H., (2018), "Monitoring the Variability of Water Dynamics in Plant Leaves at Cellular Level Using Terahertz Sensing", *Second International Workshop on Mobile Terahertz Systems (IWMTS), 1-3 July 2019*, Bad Neuenahr, Germany. **(Published)**.

5. **Zahid, A.**,Yang, K., Heidari, H. , Li, C., Imran, M. , Alomainy, A. and Abbasi, Q.H., (2018), "Terahertz Characterization of Living Plant Leaves for Quality of Life Assessment Applications." *Microwave and Radar Week (MIKON-2018 and Baltic URSI Symposium-2018, Poznan, Poland*, 14-17 May 2018. (doi:10.23919/URSI.2018.8406770). **(Published)**.

6. Ren, A., **Zahid, A.**, Imran, M. A., Alomainy, A., and Abbasi, Q.H., (2019), "Introducing a Novel Technique of Detecting Fruits Contaminations Using Terahertz Sensing," accepted by *IWAT 2019, Miami*, Florida.**(Published)**.

7. Ren, A., **Zahid, A.**, Imran, M. A., Alomainy, A., and Abbasi, Q.H., (2019), "Establishing A Novel Technique for the Detection of Water Contamination Using Terahertz Waves", *China Microwave Week International Wireless Symposium (IEEE IWS-2019), May 19-22, 2019*, Guangzhou, China. **(Published)**.

8. Ren, A., **Zahid, A.**, Imran, M. A., Alomainy, A., Dou, F., and Abbasi, Q.H., "Establishing a Novel Method for Monitoring Quality Control of Fruits Using Terahertz Sensing", Submitted to the 2019, *IEEE AP-S Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting, July 7-12*, 2019, Atlanta, GA. **(Published)**.

9. Ren, A., **Zahid, A.**, Imran, M. A., Alomainy, A., and Abbasi, Q.H.,"Terahertz Sensing for Fruit Spoilage Monitoring", *Submitted to the Second International Workshop on Mobile Terahertz System (IWMTS), 1-3 July 2019*, Neuenahr, Germany. **(Published)**.

10. Y. Wei, **A. Zahid**, H. Heidari, M. Imran and Q. H. Abbasi, "A compact Non-Invasive WearableVital Signal Monitoring System," 2018 *IEEE Asia Pacific Conference on Post-graduate Research in Microelectronics and Electronics (PrimeAsia), Chengdu, 2018*, pp. 55-59, **(Published)**.

Book Chapters

1. **Zahid, A.**, Abbas, H.T., Ren, A., Alomainy, A., Imran, M.A. and Abbasi, Q.H. (2020). Application of Terahertz Sensing at Nano-Scale for Precision Agriculture. *In Wireless Automation as an Enabler for the Next Industrial Revolution (eds M.A. Imran, S. Hussain and Q.H. Abbasi*. https:// doi:10.1002/9781119552635.ch11 **(Published)**.

2. Aifeng Ren, **Zahid, A.**, Xiaodong Yang, Alomainy, Akram, Imran, Muhammad Ali, Abbasi, Qammer H. 'Terahertz (THz) application in food contamination detection' (Electromagnetic Waves, 2019), *'Nano-Electromagnetic Communication at Terahertz and Optical Frequencies: Principles and Applications', Chap. 5*, pp. 77-100, doI: 10.1049/SBEW 542E, ch5 IET Digital Library.

Other Publications - Journal

1. Taylor, W.; Shah, S.A.; Dashtipour, K.; **Zahid, A.**; Abbasi, Q.H.; Imran, M.A. An Intelligent Non-Invasive Real-Time Human Activity Recognition System for Next-Generation Healthcare. *Sensors 2020, 20*, 2653. **(Published)**.

2. Z. Yu, A. M. Abdulghani, **Zahid, A.**, H. Heidari, M. A. Imran and Q. H. Abbasi, "An Overview of Neuromorphic Computing for Artificial Intelligence Enabled Hardware-Based Hopfield Neural Network," *in IEEE Access, vol. 8, pp. 67085-67099*, 2020. **(Published)**.

3. S. A. Shah, A. Tahir, J. Ahmad, **A. Zahid**, H. Pervaiz, S. Y. Shah, A. M. A. Ashleibta, A. Hasanali, S. Khattak, and Q. H. Abbasi,"Sensor Fusion for Identification of Freezing of Gait Episodes using Wi-Fi and Radar Imaging", *in IEEE Sensors Journal*, 2000, doi: 10.1109/JSEN.2020.3004767. **(Published)**.

4. L. B. Furstenau, M. K. Sott, L. M. Kipper, Ê. L. Machado, J. R. López-Robles, M. S. Dohan, M. J Cobo, **A. Zahid**, Q. H. Abbasi, M. A. Imran, "Link Between Sustainability and Industry 4.0: Trends, Challenges and New Perspectives," *in IEEE Access, vol. 8*, pp. 140079-140096, 2020, doi: 10.1109/ACCESS.2020.3012812.**(Published)**.

5. Ashleibta, A.M.; **Zahid, A.**; Shah, S.A.; Abbasi, Q.H.; Imran, M.A. Flexible and Scalable Software Defined Radio Based Testbed for Large Scale Body Movement. *Electronics 2020, 9*, 1354. **(Published)**.

6. Zheqi Y., Pedro M., **Zahid A.**, Abdulghani A. M., Dashtipou K., Heidari H.,Imran M. A., and Abbasi Q. H. Energy and Performance Trade-off Optimization in Heterogeneous Computing via Reinforcement Learning. *Appl. Sci. 2020*, pp 1-11. **(Under Review)**.

7. Z Yu, **A Zahid**, S Ansari, H. Heidari, M. A. Imran and Q. H. Abbasi. IMU Sensing-based Hopfield Neuromorphic Computing for Human Activity Recognition. *IEEE Sensors. 2020*, pp 1-10. **(Under Review)**.

Contents

List of Tables

List of Figures

List of Acronyms

WC	Water Content
THz	Terahertz
ML	Machine Learning
DT	Decision Tree
NB	Naïve Bayes
KNN	K-nearest neighbors
SVM	Support Vector Machines
MCK	Material Characterisation Kit
EM	Electromagnetic
FCC	Federal Communications Commission
CW	Continuous-wave
EOR	Electro-optical rectification
BAN	Body-Area Networks
THz-TDS	Terahertz Time-Domain Spectroscopy
THz-PEAS	Terahertz Penetration-Enhancing Agents
QCL	Quantum Cascade Laser
QD	Quantum-Dot
THz-FEL	Terahertz low-gain free electron lase
PCG	Planar comb Generator
PC	Photonic Crystals
TPX	Polymethylpentene
ATR	Attenuated Total Reflection
DNA	deoxyribonucleic acid
NDE	Non-Destructive Evaluation
EICIC	Enhanced Inter-Cell Interference Coordination
SS-OCT	Swept Source Optical Coherence Tomography
SMF-PLD	Single Mode Fabry-Perot Laser Diodes
NIR	near-infrared spectroscopy
MRI	Magnetic Resonance Imaging
NA	Network Analyser

NRW	Nicholson-Ross-Weir
VNA	Virginia Diodes Analyzer
PTFE	Polytetrafluoroethylene
MUT	Material Under Test
SFR	Sensitive Frequency Region
TRR	Target Response Region
IFFT	Inverse Fast Fourier Transform
STFT	Short-Time Fourier Transform
MAV	Mean of Absolute Value
STD	Standard Deviation
PCC	Pearson correlation coefficien
IQR	Interquartile Range
WT	Wavelet Transform
RBF	Radial Basis Function
SFS	Sequential Forward Selection
SBS	Sequential Backward selection
SOLT	Short-Open-Load-Thru
CDF	cumulative dis-tribution function
PSD	Power Spectral Density
CPSD	Cross Power Spectral Density
IR	Infrared
PMP	Polymethylpentene
RF	Random Forest

Acknowledgements

First and Foremost, My most sincere and deepest gratitude would like to go to my primary supervisor Dr. Qammer H. Abbasi and secondary supervsior's Prof. Muhammad Ali Imran, Dr. Hadi Heidari, for their endless support, guidance, motivation and encouragement throughout my entire journey of PhD I am highly indebted and extremely grateful to Dr. Qammer H. Abbasi for always being there, keeping me calm during tough times and showed lot of patience, and has given me lot of strength and confidence necessary for the development of the work. I strongly believe that he truly understand my potential and always motivate me to achieve the best. I would always be grateful to him for always being there and for his valuable feedback for all the publications, sharing his worthy knowledge, developing an unbounded energy and passion and have set an excellent models for me to achieve my goals. I am also grateful to Dr Chong Li and Prof David Cumming for giving me an access to the Terahertz Laboratory that enabled me to perform my experiments freely and with full confidence.

A special thanks would also like to go to Dr. Hasan T. Abbas, his knowledge and willingness to meet all the time for discussions, his valuable feedback and rigorous comments in all the publications were extremely helpful. I truly admire and admit that without his support, sincere guidance and discussions on methodology of experiments , this journey would be much harder and arduous, I wouldn't be able to execute all measurements successfully. Also, would also like to thank Dr Ahmed Zoha for sharing his knowledge and providing guidance in technical matters particularly in machine learning approach.

Additionally, I must acknowledge and would like to express my deepest gratitude to Dr Aifeng Ren for always being a great mentor, colleague and wonderful human being. The time I spend with him has given a huge opportunity to learn so much from detailed discussions relating to experiments, valuable comments and feedback on papers, presentations and always admire his in-depth knowledge. I strongly believe that he truly inspire me to go out of my comfort zone and tap my true potential and broaden my horizon a lot. Our in-lab measurements and laughs whilst performing the experiments would always be the pearls in my memory and wish him a great success from my deep heart.

I am thankful Dr Ivonne Escorcia Carranza for sparing time out from her busy schedule and helped me alot. She always ensured her support and availability for discussions, valuable input and feedback especially when I needed the most whilst dealing with expensive equipment.I would also like to thank my friends and colleagues especially Dr Paulo Valente Klaine who helped me keep going during my PhD with their moral support and valuable company.

I must acknowledge and appreciate the support of my wife Farah Adnan who has been great support to me and took care of many of my responsibilities to help me focus on my PhD, without her ubiquitous support this journey could not be possible.

Above all, I would like to thank all my family. Regardless of physical distance,their support and affection have always been with me. My last thought goes to my parents who are always standing by me through thick and thin and also my brothers and sister. I dedicate this thesis to my family, to honour their unconditional love, endless encouragement and support, and most importantly, firm confidence in me throughout my life.

Declaration

With the exception of chapters 1, which contain introductory material, chapter 2,3,4,5 and 6 are taken from my published work. I can confirm that all work in this thesis was carried out by the author unless otherwise explicitly stated.

Chapter 1

Introduction

Over the past decade, climate transformations and increasing growing population have predicted an increment in the occurrence of scarcity of water resources in many parts of world [1]. This growing deficiency of water and startling environment has posed enormous challenges in various fields of plant science sector and received strong impetus to researchers and scientists [1]. Since, agriculture is considered as the spine in the overall development of countries especially in developing countries due to its significant role in enhancing economic development of country [1,6]. Therefore, it is imperative to develop a feasible strategies and cost-effective techniques for the timely detection of drought symptoms in leaves at an early stage, while effectively utilizing water distribution in the field of plant to take proactive action in relation to plants health monitoring closely, and for precision agriculture applications, which is of high importance to improve the overall crops productivity and obtain economic benefits [2–4]. Moreover, the quantitative determination of water content WC in living plant leaves provides significant and valuable information to growers, cultivators to circumvent any unforeseen situation of plant drought stress [4, 5].

In recent years, the emerging applications of terahertz (THz) technology has captivated numerous researchers and scientists with a different background working on a wide range of subjects due to its distinctive characteristics and viable opportunities offered in various fields [7, 8]. Some of the applications that have been derived from these unique features are medical imaging for non-invasive diagnostics of dental care and skin cancer, security imaging of invisible items, atmospheric studies, processing the quality control of food,high- frequency communication and non-contact imaging for protection of paintings and manuscripts [7, 9, 10].

Although there has been significant progress of THz applications in various fields such as biomedical imaging, diagnostic applications. However, its potentials to propagate satisfactorily through plants is still one of the least examined research areas until now. In this respect, some of the elementary and significant aspects of THz transmission through various sections of vegetations, e.g. fruits, vegetables, leaves still need to be accomplished [11, 12]. The estimation of leaf water content is of immense interest to the cultivator, horticulturists, researchers,as well as scientists in various fields of applied plant biology due to growing aridification [13]. It contributes immensely to avoiding plant drought stress as well as provides valuable information in irrigation management. The water detection in plants has mainly relied upon the degree of water existence in plants and is considered as a strong absorber of THz frequency.

In recent times, significant amount of efforts have been made in the field of agriculture such as monitoring and controlling of environmental systems, crops productivity enhancement, protection of crops from any pathogen attacks, and especially, monitoring the appropriate amount of water content in leaves by using various methods [7, 16, 17]. In this context, the focus is mainly deployed on THz technology for the investigation of WC in leaves, which have been vastly researched in numerous publications aiming for the detection of dehydration response in plants and remote sensing of vegetation conditions [8, 18].

Presently, a large part of current literature about THz communications sector is mainly focused on the channel characterization and modelling for body-centric communications and also presented on-body path loss models for different body scenarios [8]. However, those models are inadequate to anticipate any transmission loss directly in a medium that contains plants leafage. To overcome this corresponding drawback, authors in [8, 19] have proposed the numerical method of elementary THz path-loss models that envisage the absorption and scattering reaction of plant leaves for estimating the total signal loss.

In spite of the fact, that these models are based on simulation results, described in the literature [19] for distinct leaves, can be approved by comparing them with some practical measurements of plants leaves. The results from measurements process will pave the way to develop a feasible, cost-effective and practical solution for the timely and proactive determination of water-stress leaves at an early stage to further explore and comprehensive understandings of plants physiological process of growth at cellular level. According to [1], the food demand will be increased by 70% to match current trends in population growth. Therefore, there is a vital need for an effective, realistic and reliable solution not only to satisfy agricultural production demand but also to revolutionize the agricultural sustainability. In this regard, there is unanimity in recognizing the role of THz technology in agriculture, especially with regard to its significant contributions in various disciplines due to its high sensitivity and strong penetration features [4]. Hence, it is strongly envisioned that aforementioned technique is a promising tool to design a smart and plant-specific irrigation system that monitors the leaf WC in a non-invasive manner in order to maintain the structural integrity of leaves at cellular level and is, therefore, critical in the current circumstances governed by global climate change that demand water conservation [3].

1.1 Research Motivation

In recent years, the growing scarcity of water resources has created grand challenges in various fields of plants sector [4]. In this respect, many techniques have been proposed by horticulturists, researchers at various levels in the field of plant biology to develop feasible strategies for the timely detection of drought symptoms in plant leaves as shown in Figure 1.1. In addition, standard sensors and systems have been employed to meet the huge requirement of crops productivity, appropriate usage of fertilizers, capable of detecting small amounts of impurities in soil and pathogens in plants, nutrients deficiencies in plants have not obtained prolific results in agriculture sectors and clearly appears to be unfeasible and unachievable.

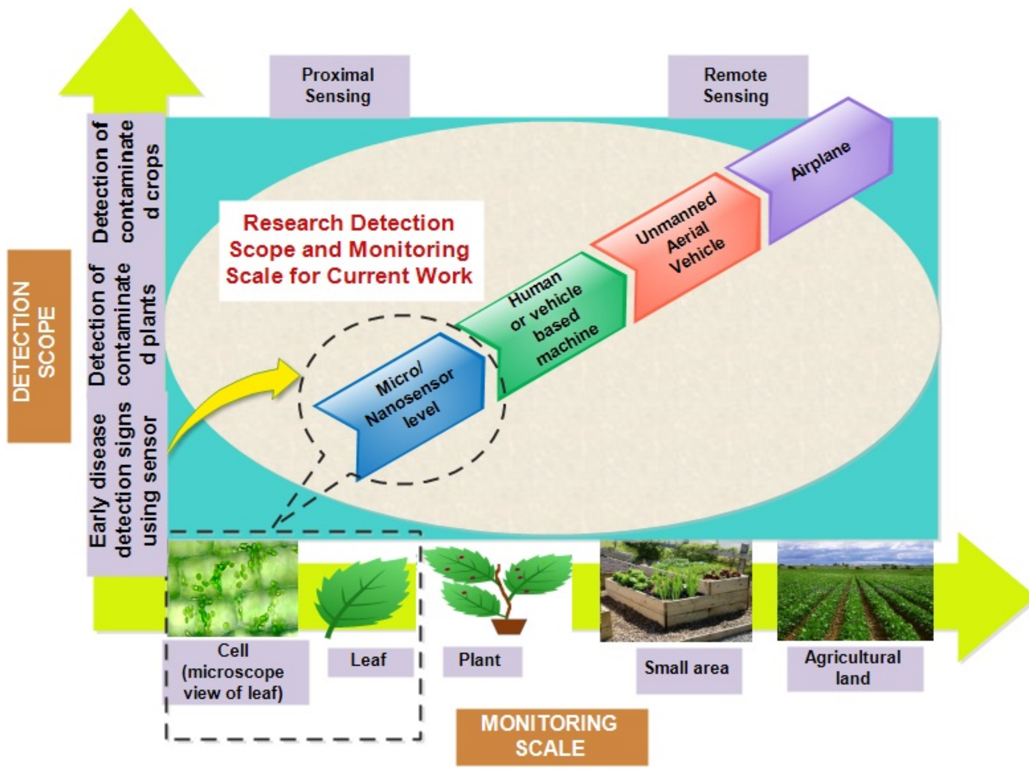


Figure 1.1: Overview of various sensing technologies used for the automated detection and identification of host-plant interactions

On the other hand, some non-destructive methods such as spectroscopy and imaging techniques can be utilized to identify plant diseases caused by pathogens as well as drought from a very early stage. In spectroscopy, the transmission or the reflection spectrum of plants (for example, from their leaves) are measured at different frequencies. The changes in the spectrum can be mapped to different diseases. In the case of imaging, such spectral information is collected with multiple pixels, enabling the generation of images, which also capture the shape of the elements being measured. However, these techniques are limited by the resolution and thus, cannot provide any cellular-scale information about the plants [242]. In addition, these techniques do not consider any environmental influence due to its low photon energy [242].

These facts and prevailing challenges in agriculture sector motivated us to develop a novel, pragmatic and non-invasive approach by utilizing the THz technology to timely detect the water-stressed at cellular level to maintain a healthy physiology and lessening the early inception of disease that may damage the crops. It is envisioned that it can deepen the understandings of biological traits and metabolic process of leaf tissues. In addition, it will help cultivators, growers to take effective and appropriate measures to monitoring the WC in plants leaves, and to maintain healthy biological traits of leaves to sustain crops productivity.

1.1.1 Why Terahertz sensing

The current state of the art technology, specifically in bio-engineering and bio-electromagnetism dictates in the microwave (50 - 900 MHz) and millimetre wave (30 - 300 GHz) regions and gradually expanding in the THz domain [21]. With an increasing awareness of hazardous diseases specially in plants, the diagnostic research is more emphasised on extracting intrinsic details of the samples and plants, precision of measurement and efficient data transfer. Over the past decade, the interest in THz technology has drastically increased due to the opportunities both in communication and sensing that this frequency band has to offer [24].

The THz technology has attracted significant attention for its unprecedented sensing ability and its non-invasive and non-ionizing properties making it a remarkable option for in-vivo characterisation and imaging of biological systems [21, 22]. The THz spectrum hosts a large number of interesting microscopic phenomena such as inter/intra-molecular motions and Debye relaxation processes. This translates into the ability to use THz radiation to sense the presence of or characterize a vast array of materials which are inaccessible to other frequency bands. For communication perspective, the THz Band channel is highly frequency selective and exhibits a unique distance-dependent bandwidth behaviour due to the absorption from mainly the content of the biological medium, mainly water and related constituents [23].

A distinguishing feature of the THz waves is their absorption by water molecules, a phenomenon that is found at the basis of innovative bio-sensing applications, ranging from skin cancer early detection [205] to plant health monitoring [3, 26]. THz technology can provide detailed insight into the health of a plant specimen in terms of the water content (WC) in the leaves [27]. Plant leaves consist of bio-molecules, such as cellulose, and synthesis compounds, including proteins, carbohydrates and many other molecular weight compounds as shown in Figure 1.2. Water is not only an essential component but an important nutrient to the process of photosynthesis and transpiration [28]. Due to the high sensitivity and strong penetration feature of THz, it has a strong potential to disseminate through plant leaves at cellular level and can yield significant information of WC in leaves. Hence, designing a smart and plant-specific irrigation system that monitors the leaf WC in a non-invasive manner is, therefore, critical in the current circumstances governed by global climate change that demand water conservation.

It is strongly believed that the terahertz sensing has the potential and is deemed to have more fast, reliable response for the overall monitoring and maintaining the health of the leaves. In this work, it is aimed to introduce a novel, simple and non-invasive method that overcome the limitations of earlier proposed methods [29–33].

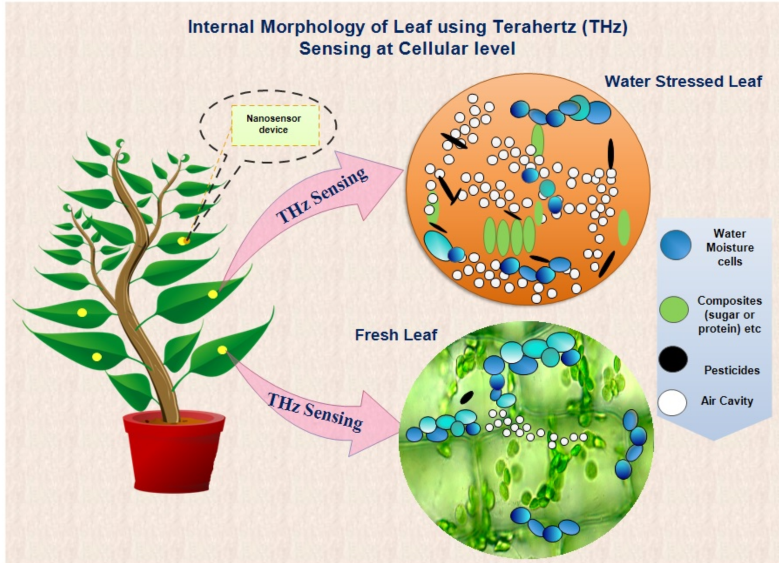


Figure 1.2: Internal morphological structure of fresh and water stressed leaf using THz sensing.

1.1.2 Why Machine Learning (ML) is needed

The ongoing aridification and continuing growth of human population has put immense pressure on the agricultural system. Researchers and scientists have made substantial effort and proposed various techniques to tackle this mounting pressure [35]. In this regard, THz technology has been immensely considered by scientists and researchers in plants scientific community to minutely observe the behaviour and physiological changes that occur in plants leaves at cellular level i.e. WC and other nutrients level in plants leaves. It is believed that though THz technology has obtained notable contributions in agriculture sector due to its non-ionising and strong penetration feature. However, an effective and feasible strategy can be significant for the timely detection of WC in plants leaves and deliver the appropriate amount of WC to plants leaves, and to maintain healthy physiological traits of leaves to sustain crops productivity. Hence, precise estimation of WC in plant leaves is of crucial to all plants physiologists and biochemists as it provides valuable information in terms of facilitating a suitable irrigation management [3].

In this context, ML has emerged as an innovative and modern agricultural approach to improve the agricultural system production while reducing its effect on the environment [35]. ML applications create an innovative opportunity to unravel, quantify, and understand data-intensive processes in agricultural operational environments. The data collected in this regard is provided by various sensors that facilitate a broader knowledge and understanding of the operational environment (an interaction of dynamic leaves, grain, seed, soil, and weather conditions) and the process itself (machinery data), leading to more precise and accurate action to sustain and meets the crops productivity in an effective manner [35].

Evidence from multi-disciplinary agri-technology studies show that reliable and early detection of WC in plants leaves at a cellular level can drive agricultural productivity and optimize the economic benefits [30–32]. In recent times, the applications of ML have been immensely used in various scientific fields [22] such as healthcare sector, food security, meteorology, medicine, meteorology, economic sciences [35]. Furthermore, researchers are very keen to discover its possibilities, specifically in modern digital agriculture systems to develop intelligent management of plants by applying the water distribution effectively [35]. In this study, a state-of-the-art method is proposed by integrating ML with THz to closely monitor the diminutive variations and physiological status in plants leaves health at cellular level. In addition, the underlying aim of amalgamation of ML with THz waves is to carefully predict the future trends of WC in plants' leaves and closely observe the internal morphological characteristics of leaves in an automated fashion, which, in turn, can help to reduce superfluous costs [3]. Furthermore, forecasting the future trends of WC in leaves using ML, will not only help to maintain the healthy status of leaves but it can save crops from stresses by taking timely action. This will ultimately help to maximize yield production and obtain financial gains.

Considering the sensory characteristics of plants leaves, water is essential to the overall growth, transpiration, and nutritional process of plants leaves [30]. Therefore, timely delivery of the appropriate amount of resource inputs such as water and its precise quantification can be very beneficial to drive and sustain overall crops productivity in an advanced agricultural system [30]. In an era, where most of the farmlands around the globe are water-stressed, the outcomes of this study will have a positively impact the agriculture sector, and can help in the design and implementation of smart, sustainable digital agricultural technologies, which is of high importance to boost the overall crops productivity.

1.2 Research Objectives

As previously mentioned, this thesis focuses on the utility of THz in the field of agriculture due to its non-ionizing and less pervasive radiation properties. In ongoing climate transformations and unceasing expansion of population that have posed grand challenges recently, researchers and scientist strongly believe that application of THz technology at cellular level can revolutionize modern agriculture discipline. In this realm, numerous researchers, physiologists, and scientists have brought considerable attentions towards this significant issue at various level and achieved advancements to develop feasible techniques for the precise estimation of WC in plant leaves to circumvent any unforeseen circumstances, foreseeing in major decline in crops production and economic loss in agriculture sector. Although, some techniques have been markedly suffered from some major limitation; time consuming methods, limited accessibility, and less sensitivity at cellular changes in plants and lower resolution issues. Moreover, some of them are considered as inappropriate for enduring studies due to their destructive nature [3].

With these limitations in mind, the initial objectives of this research are listed below:

- To study and perform the thorough literature review and recognize the recent research performed various techniques of THz technology and understand the working principle of them and their applications in precision agriculture.

- To comprehend and critically analyse the results and identify the gap that requires further investigation and research to be carried out for the timely detection of WC in living plants leaves at cellular level.

- To emphasize on novel, cost-effective, feasible and non-invasive technique that can determine the internal morphological structure, characterization and mainly estimate the precise amount of WC in various living plants leaves at cellular level using THz sensing to maintain the healthy physiological growth and lessening the early inception of age dependant disease.

- To explore the utilization and suitability of various ML techniques with an integration of THz to estimate and predict the future trends of WC in plants' leaves in an automated fashion and can ravel, quantify, and understand data-intensive processes in agricultural operational environments.

- To study, investigate and propose a unique and non-destructive THz technique enabled by ML for the precise detection of impurities in water quality at cellular level. Since water is an important and integral component for plants, therefore, the fresh and pure water delivery to the plants leaves in a timely manner can be vital to maintain the overall growth and physiological traits of leaves at cellular level to sustain the crops productivity and significantly help in mitigating yield and economic losses.

- Lastly, to provide thorough conclusion with future work as well as open challenges are also addressed to emphasize the utility of THz in the field of agriculture discipline.

1.3 Key Contributions

The growing demand and effective use of various types of sensors to determines the plants' responses to changes in environmental conditions, pathogen attacks, and other stresses have become the most fascinating areas to be explored over the past few years. Concerns have arisen when the existing methods experience some major drawbacks, such as; time-consuming, low-resolution, sensitivity issues, and unable to provide information of WC in the leaves at cellular level [8]. Moreover, some of them are considered as inappropriate for enduring studies due to their destructive nature [8]. Consequently, this has led researchers to quest for more viable, practical and focus on non-destructive and non-invasive techniques and have considered the THz technology to determine internal morphology of leaves, relating to WC in leaves due to its strong absorption feature [8]. Based on THz technology and objectives in mind, the contributions of thesis are summarised as follows:

1. Provides literature review covering the advances in the THz sensing for microbiological contamination of food and water, in addition to state-of-the-art in network architectures, applications and recent industrial developments. Furthermore, various studies have been analyzed, examined and discussed covering the various aspects of THz systems, components, THz spectroscopy and imaging technologies. In addition, it also highlight the THz detecting technologies and different potential applications that can be effectively used for food and water contamination detection and linked those to future directions. Towards the end, it also present some of the open challenges such as distraction and absorption of water, low penetration depth, scattering effect, system compactness and equipment cost that need to be addressed.

2. This thesis presents a novel electromagnetic method to monitor the WC and characterisation in plant leaves using the absorption spectra of water molecules in the THz frequency for four consecutive days. It was markedly noticed that after four days, WC in all leaves were fully evaporated and this was also validated by observing no substantial changes in transmission response of leaves. We extracted the material properties of leaves of eight types of leaves namely, coffea- arabica, baby-leaf, parsley, pea-shoot, baby-spinach, basil, coriander and lamb's lettuce from the scattering parameters, measured using a material characterisation kit (MCK) in the frequency range of 0.75 to 1.1 THz. From the computed permittivity, it is deduced that the leaf specimens increasingly become transparent to the THz waves as they dry out with the passage of days and showed a exponential decay response. Moreover, the loss in weight and thickness of leaves were observed due to the natural evaporation of leaf moisture cells and change occurred in the morphology of fresh and water-stressed leaves.

3. The demand for effective use of water resources has increased because of ongoing global climate transformations in the agriculture science sector. Cost-effective and timely distributions of the appropriate amount of water are vital not only to maintain a healthy status of plants leaves but to drive the productivity of the crops and achieve economic benefits. In this regard, employing a terahertz (THz) technology can be more reliable and progressive technique due to its distinctive features. This thesis presents a novel, and non-invasive machine learning (ML) driven approach using terahertz waves with a swissto12 MCK in the frequency range of 0.75 to 1.1 THz in real-life digital agriculture interventions, aiming to develop a feasible and viable technique for the precise estimation of WC in plants leaves. For this purpose, using measurements observations data, multi-domain features are extracted from frequency, time, time–frequency domains to incorporate three different machine learning algorithms such as support vector machine (SVM), K-nearest neighbour (KNN) and decision-tree (D-Tree).

4. In agriculture science, accurate information of WC in fruits and vegetables in an auto-

mated fashion can be vital for astute quality and grading evaluation. This demands for a viable, feasible and cost-effective technique for the defect recognition using timely detection of MC in fruits and vegetables to maintain a healthy sensory characteristic of fruits. Here we propose a non-invasive ML driven technique to monitor variations of MC in fruits using the THz waves with Swissto12 MCK in the frequency range of 0.75 THz to 1.1 THz. In this regard, multi-domain features are extracted from time, frequency, and time-frequency domains, and applied three ML algorithms such as SVM, KNN, and D-Tree for the precise assessment of MC in both apple and mango slices in non-invasive manner.

5. Considering the health of plants leaves, providing a fresh and pure water in a safe and reliable way to crops with limited resources is a huge challenge as the demand increases with a rising population. Also, fresh and unpolluted water is worsened by climate transformations, more regular droughts in many parts of the word, and by water pollution, making it more demanding and costly to handle. To safeguard the public health and environment protection, it is essential to develop a fast, real-time mechanism through which any hazardous impurities in the water supply can be readily detected. In thesis, the prospects of integrating non-invasive terahertz THz waves with ML enabled technique is studied. The research explores a method of using Fourier transform Infrared Spectroscopy (FTIR) system to observe the absorption spectra and characteristics of three solvents solution, including salt, sugar and glucose with various quantity in aqueous solutions in the frequency range of 1 THz to 20 THz. In this study, due to the different molecular configuration and vibration modes of substances, distinct absorption spectra peaks were achieved for different concentrations of solvent solutions at certain sensitive THz region.

1.4 Thesis Outline

The rest of the thesis is organised as follows:

- Chapter 2 presents a thorough and first-time analysis of the available literature covering advancements in THz sensing for food and water microbiological infection, as well as state-of-the-art network design, implementations and recent industrial innovations. It also describes the THz sensing techniques for food and water contamination detection. The open challenges for food and water contamination detection are also addressed in this chapter.

- Chapter 3 gives detailed introduction and highlight some of the notable contributions and advancements of THz obtained in various disciplines. It also provides insight of various techniques that have been employed by researchers, horticulturists for the quantification of the WC in living plant leaves. In addition, it presents a novel, non-invasive approach for characterising and monitoring the WC of plant leaves using the scattering parameters of a THz waves. The proposed scheme can be used to design efficient irrigation systems on-site without any need to remove the leaves from plants.The measurement results and different parameters are discussed such as permittivity, the effect of weight and thickness, followed by a comparison of transmission response of all eight leaves between day 1 and 4.

- Chapter 4 describes ML driven non-invasive approach for the estimation of WC in living plants leaves using THz waves. It includes description of exhausting techniques followed by their limitations. In addition, the proposed method incorporating ML using THz is discussed including data collection pre-processing methodology, feature extraction and optimal feature selection and analysis of three classifiers results are discussed along with an initial. From the results, it is proved that proposed technique can be beneficial for precise estimation of WC in leaves and can provide prolific recommendations and insights for growers to take proactive actions in relations to plants health monitoring.

- Chapter 5 presents the utilization of ML driven non-invasive technique for the quality assessment of fresh fruits using THz waves. In this chapter, the proposed method is described starting from the experimental set followed by the data-collection and pre-processing methodology, feature extraction and selection technique and performance of three classifiers are discussed in detail. This study will pave a new direction for the real-time quality evaluation of fruits in a non-invasive manner by incorporating ML with THz sensing at a cellular level.

- Chapter 6 contributes to the use of non-invasive THz waves and ML enabled optimized technological solution to detect various substances and their distinct concentrations in aqueous solutions. In this chapter, introduction is presented and briefly highlight the techniques that have been employed by researchers for the detection of water contaminants in aqueous solutions. After that, it discusses why the current techniques are not feasible and propose a realistic method and application of FTIR enabled by ML that can provide the approximate detection of smallest of contaminants in distilled water due to high sensitivity and non-destructive nature. In this regard, results are showcased and analysis of results are explained.

- Chapter 7 provides a summary and findings of the study and concludes the accomplished work. It also discusses future trends and research activities in relation to terahertz sensing for future agriculture applications are also described.

Chapter 2

Background and Literature Review

2.1 Introduction

The electromagnetic EM radiation in the range of 0.3 THz to 3 THz has unique properties that makes it particularly attractive for various applications including biomedical imaging, packaged goods inspection, food inspection, water contamination detection and so on. The THz radiation with much lower photon energies (4meV for 1THz) is a non-ionizing electromagnetic wave [36]. Ionizing radiation is defined with a photon energy greater than 10eV by US Federal Communications Commission (FCC), which corresponds to the ionization energy of oxygen and hydrogen about 14 eV [37, 38]. Unlike X-ray radiation and microwaves with their ionizing properties, THz waves with a non-ionizing form do not cause any negative effect to damage tissues and biomolecules on food or other live cells. Not only that, the THz waves can be used to determine the differences between molecules due to the vibrational and rotational modes of many molecules in the THz region. However, generating radiation at terahertz level presents many practical challenges. In recent years, several techniques for generating both continuous-wave (CW) and pulsed terahertz radiation have been developed [39].

In turn, these are spawning the early development of terahertz applications, particularly in microbial pollution in water and food. New trends, discoveries and applications have been observed in diverse fields, especially in the field of photonics and nanotechnology. The first demonstration of THz wave time-domain spectroscopy exploiting femtosecond laser sources to produce and identify freely transmitting THz pulses were developed by Auston at Bell Labs and Grischkowsky at IBM in 1980s [40]. The development of optimized techniques for THz pulse generations, particularly the optimized non-collinear beam geometry [41], the production of THz pulses using the energy of few microjoules or even higher values which has the potential to access the 2nd or 3rd order non-linearities [42, 43] is a standard followed in almost every laboratory settings.

In [44], it generated the highest-energy ultrashort THz pulses by electro-optical rectification

[0]This chapter is from paper no 2* in publication list

(EOR) of femtosecond pulses in LiNbO3 leveraging the tilted-pulse-front pumping technique, which is intrinsically expansible to boost the available THz pulse energy and corresponding field strength. The THz radiation has many advantages over other imaging modalities in many aspects including non-ionizing [45], classifying species of tissue [46], and scattering and water absorption for porosimetry [47]. Consequently, the THz radiation has shown widespread potential applications in diverse areas to address the real-world problems including genomics [48], medical diagnostics [49], pharmacology [50], healthcare applications [51, 52], Body-Area Networks (BAN) [53, 54], wireless communication [55, 56], environmental monitoring [57], agriculture [58] and food analysis [36, 59], defense and security [60, 122].

The recent research work on THz radiation generation, detection, and the spectroscopic and imaging tools, such as terahertz time-domain spectroscopy (THz-TDS) [61–63] that has the potential to be operated in the time and spatial domain to process the data obtained from the transmitted or reflected THz beam, opens up new ventures by applying the particular technology for food or water contamination detection [64]. The core idea is food and liquid have unique physical features and presents a unique spectral imprint when exposed in the THz frequency domain [65–68]. The food quality and safety monitoring comprising microbiological contamination detection including toxic metals [70], pesticides [71], veterinary drug residues, organic pollutants, radionuclides and mycotoxins, is a pressing issue for public health and well-being. Several studies have focused on different types of food additives, such as flour and talc mixtures [72], melamine in milk powders [74], which indicate that the absorption spectra of the pure and mixed products were by and large distinct.

The strong absorbing nature of THz radiation by water, the vibrational modes between molecules of hydrogen bonds positioned in the THz region presents limitations on the application of THz technology considering the detection of concealed cases [125, 129]. However, the absorbing nature does not pose challenge to the applications of THz imaging in identifying humidity in different substances. There are several methods that can resolve limited penetration depth issue, for example the paraffin-embedding technique, freezing technique and terahertz penetration-enhancing agents (THz-PEAs) [49]. There are several existing sensing techniques that can be used for contamination detection at THz level [73]. The existing studies on the THz sources and processing methods, such as the spectral imprints of commodity of food in the THz region, enable the synthesis of THz sensing systems into real-world.

This chapter analyses the recent trends observed in THz technology, THz detection techniques and reports different issues that need to be addressed in order to apply THz sensing for food and water contamination detection. In this regard, the contribution of this paper is discussed as:

In section 2.2, the application, architecture and techniques for implementing THz systems are presented in comprehensive manner. Section 2.3 describes the THz detection techniques and the associated technologies. Section 2.4 illustrates the THz sensing for food and water contamination detection techniques. Section 2.5 presents the summary drawn from existing literature.

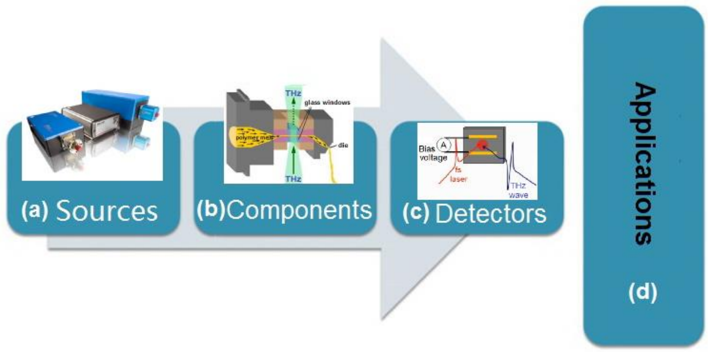


Figure 2.1: Typical terahertz configuration. (a) Sources, (b) Components, (c) Detectors, (d) Applications, such as spectroscopy and imaging [56]

.

2.2 Terahertz Components and Application Systems

The terahertz sources, components, detectors, and application-oriented goals are the main focus of the development of THz technologies. Figure 2.1 indicates the typical THz configuration that consists of multiple sources, components and detectors. The introduction of solid-state mode-locked and quantum cascade lasers (QCL), which can provide a significant continuous-wave THz source, also the technological breakthrough in laser-based THz-TDS and microelectronic fabrication have led to the actualization of several significant THz equipments and systems as in [76–78]

2.2.1 The THz Sources

A terahertz source produces a broad range of radiation [79, 80] presents six different types of THz sources which include thermal [81], vacuum electronic [82], solid-state electronic [83], lasers [84], sources pumped using lasers (including continuous [85], pulsed [86] and mechanical excitation sources [135]. Past few years have witnessed several new THz source techniques that have been brought forth. For examples, terahertz source techniques were summarized by [87]. An InAs/GaAs Quantum-Dot (QD) Laser Source [88] can be ultracompact, tunable in room temperature and its experimental setup involves a QD based photomixer resonantly pumped by a compact, broadly-tunable dual-wavelength QD laser in the double-grating quasi-Littrow configuration. A high peak power THz source for ultrafast electron diffraction was constructed operating at the wavelength ranging between 50-200µm [89]. A device model consisting of electron source and terahertz low-gain free electron laser (THz-FEL) oscillator was proposed in [90]. The THz radiation source based on electron linear accelerator (linac) was introduction in [91]. A frequency reference source based on the optical combs was proposed to simultaneously stabilize a series of THz source devices in [92]. The optical combs, which deliver uniformly-spaced CW lights having significantly high accuracy, can be produced using combination of a planar comb generator (MZ-FCG) based on a Mach-Zehnder modulator and a highly nonlinear dispersion-shifted fiber.

2.2.2 The THz Components

All THz components such as mirrors, lenses, and polarizers manipulate the specific radiation. Conventionally, the THz mirrors were made of metals such as aluminum, silver, copper and gold. In order to achieve maximum reflectivity characteristics from these metals, the width of the metal coating should be at least two skin depths at the frequency of the incident beam [93]. Recent innovations in THz mirror technology include semiconductor [94], hybrid mirrors [95], and the tunable mirrors are based on photonic crystals (PCs) [96]. THz lenses are typically made of plastics such as polyethylene or Polymethylpentene (TPX). Recently, the significant innovations allow the rapid production of a large number of lenses (Wichmann et al., 2011) that include Fresnel zone plates [97], plasmonic resonances [98], variable focal length lenses [99], 3D printed diffractive lenses [100], and even THz lenses made of paper [101]. The polarizers are one of the significant components in THz imaging, data transmission, and spectroscopy. Wire-grid polarizers are simpler to realize in the THz than in the visible region [102]. Recently, reconfigurable polarizers in [103] and carbon nanotube fiber polarizers in [104] have been reported.

2.2.3 The THz Detectors

The THz detectors that can measure the THz radiation upon reception, play a significant role in many areas including astrophysics, biological, chemistry and explosions detection, imaging, astronomy applications and so on [106] discussed various types THz radiation detectors, including basic direct [107] and heterodyne detectors [108] Schottky barrier diodes [109], pair braking photon detectors [110], thermal detectors [113], and field-effect transistor detectors [111]. [134] introduced the design and testing of broadband quasi-light THz Schottky diode detectors, which

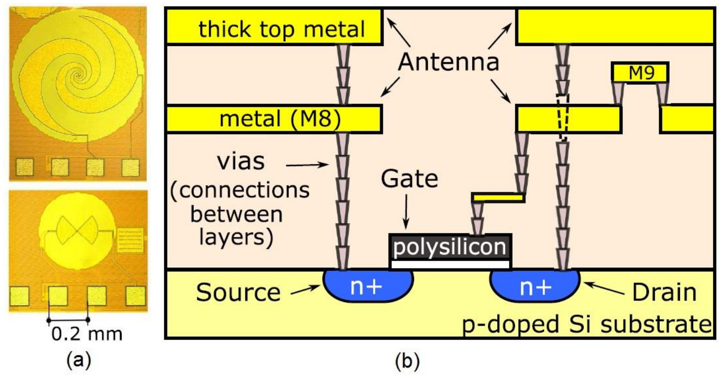


Figure 2.2: . (a) Top is micrographs of the log-spiral TeraFETs and bottom is bow-tie antenna. (b) A simplified cross-sectional view of the transistor region of the TeraFET [105].

.

can be used as a set of waveguide detectors to provide a reference point at the same location of a source beam. [105] presented a wideband bow-tie antenna-coupled THz detector on a 90nm silicon CMOS FET (TeraFET) for the identification of free-space THz radiation, which was enhanced employing an internally designed circuit based on physics model and the THz detection well beyond 2THz with an optical responsivity of 45mA/W (or 220V/W) [139]. Figure 2.2 illustrates the micrographs and the layout of this detector.

2.2.4 The THz Systems and Applications

The broad applications of the THz radiation can be categorized into four main groups: sensing, imaging, spectroscopy and communication [126, 179], and can also be applied in preventive health care to quality control, surgery, and non- destructive evaluation [136] as shown in Figure 2.3

The diverse areas of application of THz technology have been summarized in [59] and in [36], especially the applications in safety and quality control of commodity of food are shown in Table 2.1. The table indicates that the potential of THz technologies in the food and water detection fields and presents various other facets including the identification of foreign bodies, detection of pesticide and antibiotic residues, characterization of edible oil and discrimination of transgenic crops, and so on. For example, the identification of undesired and potentially unhealthy organisms in food is excessively important in the food industry [77]. So many attractive applications of THz radiation totally attribute to its enormous intrinsic properties and unique fingerprints characteristics of biological tissues at some specific frequencies in the THz electro-

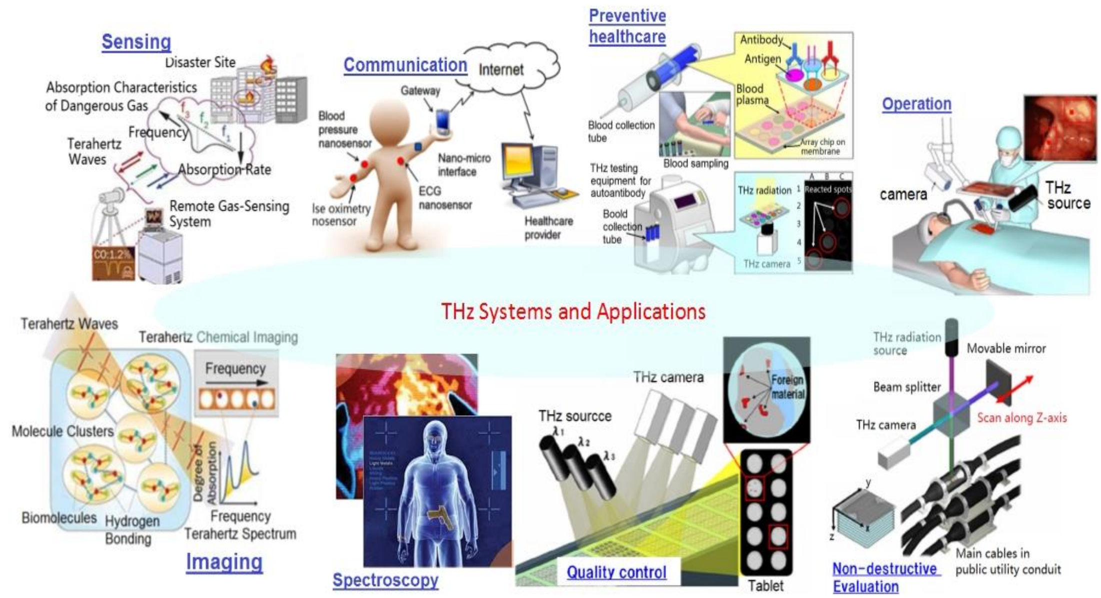


Figure 2.3: Envisioned applications for THz radiation
[136], [179], and [12] .

magnetic region.

The antibiotics and harmful residues in food or the other agricultural products were identified due to the different absorption characteristics of the chemical materials in the THz region. Regarding the penetration capability of THz radiation to the fabric, foam and plastic, it offered effective technology for foreign body detection in quality control of food products [36]. With regard to the advantages of non-destructive, non-invasive, and non-ionizing nature, the THz technologies is termed as the alternative methods for sensing food or water safety control and contamination detection.

2.3 Terahertz Detection Methods and Their Enabling Technologies

The field of THz technology has changed rapidly in the past few years. The research work on terahertz technology has greatly promoted the generation and development of many new technologies such as THz-TDS, terahertz radiation, and THz imaging. These technologies have been

Table 2.1: Overview of the applications of THz technology for food and water detection

Types	Methods	Detected materials
Food products	THz-TDS	Melamine detection in foods [116].
		Detection of a carbamate insecticide in food matrices [115].
		Moisture content in wheat grain [174].
		Pesticide detection in food powders [137]
		Antibiotic detection in food and feed [180]. matrices.
		Discrimination of transgenic crops [155, 190]
	THz-FDS(a)	Recognition of CCH(b) and TCH(c) in soil, chicken, and rice [188]
	Imaging	Metallic and nonmetallic foreign bodies in chocolate [220].
Water (including beverages)	THz-TDS	Characterization of optical properties of vegetable oil.
		Characterization of the dielectric properties of water solutions [130, 149]
		Prediction of sugar and alcoholic content in beverages and liquors
		Identification of adulterated dairy product [60].
		Detection of harmful chemical residues in honey [158].
	TDS-ATR(d)	Dielectric constants determination of distilled water and sucrose solution [165].
	Imaging	Sugar detection in beverages [153].
Note: a THz-FDS: Terahertz Frequency-Domain Spectroscopy		
b CCH: Chlortetracycline hydrochloride		
c TCH: Tetracycline hydrochloride		
d TDS-ATR: Terahertz Time-DomainAttenuated Total Reflection spectroscopy		

widely studied and applied in various fields.

2.3.1 Time-domain spectroscopy

The THz-TDS can be configured in three modes including transmission, reception and attenuated total reflection (ATR) modes. A typical setup of THz-TDS system and three modes are shown Figure in 2.4. Compared with transmission and reflection modes, the ATR THz-TDS is highly sensitive and suitable for measuring high moisture samples [36, 140]. The TDS system has the potential to aquire direct measurement of the field amplitude information and phase information as the function of time by sweeping out the transient field of the THz pulse. The terahertz frequency range can be accessed by applying Fourier transformation on time-domain data.

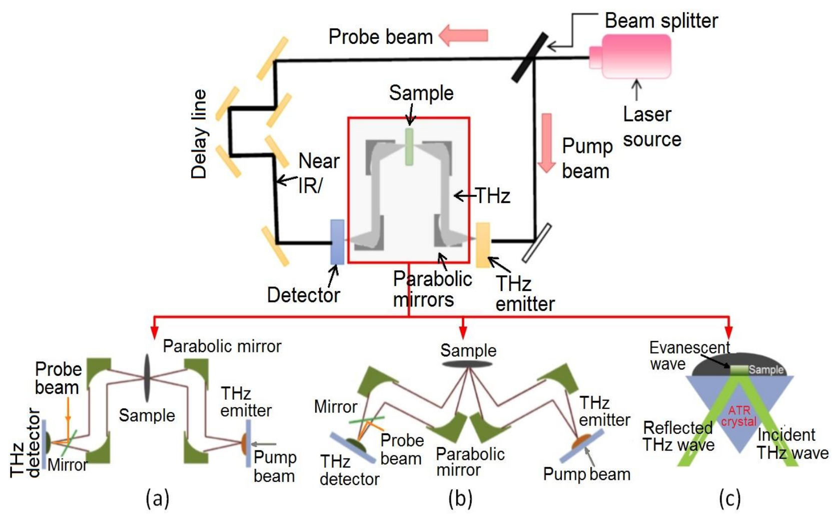


Figure 2.4: Schematic of a typical THz time-domain spectroscopy system.(a) transmission mode, (b) reflection mode, (c) attenuated total reflection (ATR) mode. (reproduced from [36]
.

As the effective method of identifying the material spectra information in the THz region using ultrafast laser beam technology, THz-TDS has comparatively high signal-to-noise (SNR) ratio up to 120 (@peak maximum) in time-domain data due to the broader frequency range, and can be used to determine the nature of material (or composition) and profound structure changes [118, 121, 163]. [177] detected the molecular classification and the associate imprint peak of three plant growth regulators including 2, 4-Dichlorophenoxyacetic acid, forchlorfenuron and indole-3-acetic acid. [116] demonstrated the feasibility of classifying the melamine in food products and found the characteristic absorption peaks of melamine at 2 THz, 2.26 THz, and 2.6 THz. In [196], it quantified the percentage contributions of the inter-molecular and intramolecular vibrations to the normal mode leveraging THz vibrational spectroscopy to molecular characterization of saccharide molecules. In [119], it reported that the THz molecular resonance imprints of deoxyribonucleic acid (DNA) methylation in cancer DNA, which can be characterized to identify various types of cancer cells. In [161], it detected the different pollutants such as oil, gypsum and water with the glass fibre reinforced composites materials and showed that the wave propagation velocity is slower in contamination region.

2.3.2 Terahertz Radiation

Terahertz radiation refers to the electromagnetic radiation region between 0.1 THz and 10 THz in frequency range and between millimeter wave and infrared wave. The resulting T-ray has a broad application prospect in object imaging, medical diagnosis, environmental detection and communication. However, in the frequency range of terahertz radiation, the transmission of THz wave will suffer slight fading and group velocity dispersion due to the water molecule in the atmosphere [69]. In [127], the field of communication is discussed and the possibility of the THz communications due to the larger bandwidth and higher transmission rate, even though there exits the strong atmospheric absorption and low source efficiency of the THz radiation.

2.3.3 Terahertz Imaging

The terahertz imaging is an emerging and significant non-destructive evaluation (NDE) technique used for safety checks, quality control [75, 162, 169, 171–173], and even fanciful applications, such as using terahertz image detection technology to count almond chocolates in bars. One of the great advantages of terahertz in the field of image processing is that terahertz radiation has a unique ability to penetrate ordinary packaging materials, so it can provide spectral information of internal materials. In [117], it introduced an active THz imaging automatic biometric security access control and tracking gate system consisting of the THz scanning/ imaging, autonomous transition and X-Ray modules, which can detect the concealed object on the human body. In [162], it constructed a swept source optical coherence tomography (SS-OCT) system with operating frequency of 600-665GHz, which can be used for quality control. In [181], it presented a compact and portable THz imaging system using Single Mode Fabry-Perot Laser Diodes (SMFPLD), which has the simple structure and low cost. In [183], it proposed a THz tomography system with the 3D scanner, which allows for the reflection THz tomography image of the sample with arbitrary surface shape by placing the transmitter and receiver on the robotic arm and can be used to study variety of non-planar objects, including biological specimens, mummies, cultural heritage and industrially made parts.

2.4 Food and Water Contamination Detection Methods

2.4.1 Terahertz for food contamination detection

Even though many policies and efforts have been made to prevent food contamination, it is more important and challenging to inspect the quality of food products and evaluate the contents of food-material. Especially, the safe methods for the bio-materials and the agricultural products are of utmost importance, which are effective to identify contaminated food, to detect toxic substances, such as pesticide residues, additives, antibiotics, and pathogens in foodstuffs. Although many technologies have been used for qualitative and quantitative analysis [156] in wide range of food products, such as near-infrared spectroscopy (NIR) imaging, bioassays, and molecular imprint-based sensors [123] the non-ionizing and nondestructive THz radiation is considered the novel, accurate and economical way for analyzing the food compositions.

Detection by analyzing moisture content

Moisture content, or water content, is the quantity of water contained in a material, such as seed and grain, food and beverage, agricultural and soil, pulp and paper products, petroleum products. The THz radiation has the potential advantages for sensing and imaging of the moisture map, non-destructively assessing hydration levels in materials [124].

Detection of water content in living plant leaves: The first application of THz imaging and sensing for detecting the WC is to monitor and evaluate the moisture level of plant leaves, which can provide the valuable information to farmers and scientists regarding plant drought stress and irrigation management. In [178], it reviewed the theoretical methods and the applications of THz spectroscopy and imaging for detecting leaf water content. A theoretical basis for detecting the water content in plants was proposed in [167], which can be conducted to establish prediction models of MC base on the processed THz transmission and absorption spectra. The meaningful information attained from the water detection of the living plant leaves can be useful to analyze the existence of any pesticides in leaves [195]. A complete analysis of the measured THz absorbance for detecting water content in the plant leaves should consider the effect of interface roughness of the leaves [220], as shown in equation 2.1 and equation 2.2, by defining an effective attenuation coefficient of the leaf, a_{total},as a sum of absorptive and scattering losses, a_{abs} and a_{scat}.

$$a_{total} = a_{abs} + a_{scat} \tag{2.1}$$

where a_{scat} is defined as

$$a_{scat}(\lambda) = \frac{1}{D}\left[(\sqrt{\varepsilon_L(\lambda)} - 1)\left(\frac{4\pi\Gamma\cos\theta}{\lambda}\right)\right]^2 \tag{2.2}$$

where D is the leaf thickness, (Γ) is the standard deviation of the height profile from the measure of surface roughness, (θ) is the incidence angle, (λ) is the THz free space wavelength, and ε_L is the wavelength dependent permittivity of the leaf.

Quality control of damaged fruit: Depending on the different absorption coefficients and refraction index of moisture on the thin surface layers of fruits, the THz wave can be used to evaluate the internal quality on fruit by non-destructive detection [67]. The damage on the tomatoes caused by pressure on the outer surface was detected and assessed in [170]. The THz reflectivity from the pressed regions revealed the decreased tendency due to the damaged area losing the moisture. Through the detection of the refractive index and the absorption coefficient in regression model, the soluble solids content, which is an important index for fruit quality, was determined in apple products by THz-TDS technique [132]. In [148], the potential of dielectric spectroscopy related to temperature, soluble solid content, moisture content with the fruit, which can be gained by THz technology, was applied to characterize the fruit quality factors in 13 fruits or vegetables at various levels, including apple, avocado, carrot, grape, guava, mango, orange, peach, melon, coconut, eggplant, potato, and tomato.

Detection of moisture in dried food: Dried foods, or low-moisture foods, are more easily preserved because the lower moisture content inhibits the activities of food-spoilage and food-poisoning microorganisms. It is significant to detect and monitor the moisture content of dried food for avoiding the mold growth when the water content exceeds the normal level [151]. A quantitative demonstration of moisture measurements in dried food was performed using THz radiation in [174] and [175]. The advantages of THz technology are that the THz absorbance of water is higher than the other constituents and no significant heating is produced in the sample due to the minimal THz power [124].

Detection by inspecting food

One of the widely used applications of the THz technology is for inspecting the foreign object such as glass shards, stone, rubber, insects and plastic, or adulteration in food products. [171], [173] performed the Sub-Terahertz transmitted Quasi-Bessel Beam imaging to determine foreign substance in food and gained a deep focus and the high spatial resolution up to 0.75 wavelength at 140 GHz frequency. In [172], they developed a reflection-mode CW THz imaging system employing a beam-steering tool for use in food quality inspection. A THz real-time imaging system with a coherent source was developed in [194], which has a 1mm diffraction limited resolution. In order to reduce the limitations on the thickness of foods inspected due to the strong absorption of THz waves by moisture, [185] newly developed a dual-polarization imaging device exploiting a THz noise source that enhanced the classification capability of foreign substances in foods, even existing similar dielectric constants. As shown in Figure 2.5, the inspected foods can be transmitted by a conveyor belt, and the 1.2 mW low coherence THz noise signals, the frequency between 0.075 THz and 0.11 THz, are generated by the source. This THz imaging system can reach up to 4 mm maximum resolution with dual-polarization imaging.

2.4.2 Detection by chemical recognition with THz pulses

Due to the excessive use of direct or indirect chemicals in the growth of plants and animals, chemical contamination of food and water has become one of the main reason of diseases. The non-contact and non-destructive detection on the non-edible additives, biological and chemical

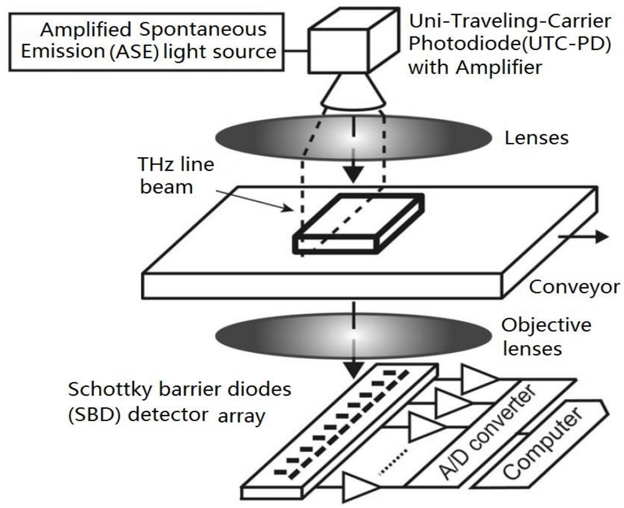


Figure 2.5: Schematics diagram of the 0.1 THz imaging system [185].

items has become a major objective in food security [67]. Due to rotation or vibration level transitions within a molecule, various chemicals and bio-molecules will appear intrinsic resonance features in the terahertz wave region, that are correlated to the physical nature and chemical composition of the sample [180]. In [116], it implemented THz spectra and images of melamine mixtures to detect melamine in foodstuffs, which showed the potential of THz-TDS as an identification and quantification technique in the food security.

In [128], they applied the simple and fast recognition algorithm in THz imaging combined with the THz-TDS data analysis to detect the drugs, such as cocaine and morphine, and other chemicals, such as lactose, sucrose, aspirin and tartaric acid, hidden in containers. In [158], they applied THz-TDS technique to discriminate the harmful chemical residues in honey and found that each chemical substance exists the distinct absorption peak in the THz frequency regime range from 0.5 THz to 6.0 THz. In [153], obtained several carbohydrate samples THz spectra in the frequency range of 0.5-2.5 THz, which included D-glucose, fructose, cellulose, and cellulose, shown as in Figure 2.6.

The results showed clearly that the strong absorption peaks appeared at monosaccharide group, such as two peaks for D-glucose at 1.4 THz and 2.1 THz in Figure 2.6(a), two peaks for fructose at 1.8 THz and 2.1 THz in Figure 2.6(b), two peaks for sucrose at 1.4 THz and 1.8 THz in Figure 2.6(c). However, the polysaccharide group appeared no recognizable absorption features in the THz spectrum, such as in Figure 2.6(d). Meanwhile, a highly sensitive and selective THz meta-material sensing tool for discriminating different kinds of carbohydrate molecules even at low concentrations of real market beverages was proposed in [153].

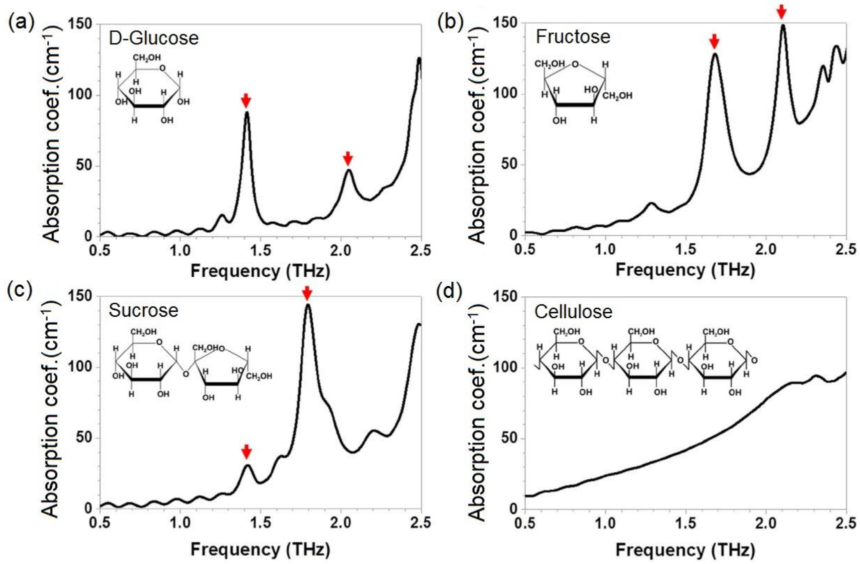


Figure 2.6: Absorption coefficients measured from THz-TDS for various sugars (a) D-glucose($C_6H_{12}O_6$), (b) fructose($C_6H_{12}O_6$), (c) sucrose($C_{12}H_{12}O_{11}$), (d) cellulose($C_6H_{10}O_5$) pellets. The structural formulas for each saccharide are shown in (a-d) as insets [153].

2.4.3 Terahertz for water contamination detection

Contrast mechanism for detection

The permittivity, the thickness, and the shape of the sample lead to the different reflection and transmission coefficients of each sample. The permittivity of the sample is a significant parameter to recognize its behaviour in theory. Although water is a strong-absorbing liquids in the THz radiation, the special THz-TDS system is available for sensitive liquid transmission spectroscopy [160]. Consequently, the contrast mechanism with the high permittivity of liquid water can be considered for investigating the molecular nature of the aqueous solutions [124, 168]. For example, in [141], it determined the alcohol content in water-alcohol mixtures based on the measurement of the THz dielectric functions of the mixture. In [187], it showed that the absorption coefficient and refraction index in the aqueous mixtures with other dipolar liquids are substantially less than that of water. The reason of these results is the slight difference between the permittivities of pure water and water molecules bounded to other materials weakly.

THz response of water – Debye model

The data received from THz TDS system is often described by a double-Debye model [166], which was first applied the complex permittivity of the water to represent the reaction to an external electric field [192]. When water is irradiated by THz radiation, the complex permittivity function is termed as:

$$\varepsilon_{\omega} = \varepsilon'(v) - j\varepsilon''(v) = \varepsilon_{\infty} + \frac{\Delta\varepsilon_1}{1+j\omega\tau_1} + \frac{\Delta\varepsilon_2}{1+j\omega\tau_2} \tag{2.3}$$

where $\varepsilon'(v)$ and $\varepsilon''(v)$ are the real part and imaginary or loss part of the complex permittivity, respectively. ε_{∞} is the permittivity at high-frequency limit. $\Delta\varepsilon_1$ denotes the dispersion in amplitude information of the slow relaxation procedure, also the bulk bonding among hydrogen molecules is liberated to equilibrium state beneath the impact of external electric field. $\Delta\varepsilon_2$ illustrates the fast relaxation procedure, where hydrogen-bond generation and destruction happen. τ_1 represents the overall relaxation time of the slow process and τ_2 is the relaxation time of the fast procedure [146]. ω is the angular frequency. The Euclidean distance between the input raw data and the output of equation 2.3 was applied to obtain the parameters of Double-Debye model (Lazebnik et al., 2007), shown in equation 2.4:

$$e = \frac{1}{N}\sum_{i=1}^{N}\left[\left(\frac{\varepsilon'(v)-\widehat{\varepsilon}'(v)}{\mathrm{median}\left[\varepsilon'(v)\right]}\right)^2 + \left(\frac{\varepsilon''(v)-\widehat{\varepsilon}''(v)}{me\,\mathrm{dian}\left[\varepsilon'(v)\right]}\right)^2\right] \tag{2.4}$$

where $\widehat{\varepsilon'}(v)$ and $\widehat{\varepsilon''}(v)$ denote the total output of the in equation (4), representing the real and imaginary part of the dielectric properties, respectively. Furhtermore, the value of N is the points number collected from the frequency of 0.1THz-1.5THz. The optimisation problem algorithm can be applied to solve equation (4) which subject to $\varepsilon_{\infty} \geq 1, \Delta\varepsilon_1 \geq 0, \Delta\varepsilon_2 \geq 0$, and $\tau_1 \geq 0, \tau_2 \geq 0$. [193] applied particle swarm optimisation (PSO) with 500 particles to solve this problem and received all Double-Debye parameters for the five groups of variables when the Euclidean distance is less than the threshold of 0.0012 or 1000 iterations reached. The result showed that the dielectric model accurately represent he dielectric properties of collagen not only at the higher values of frequency but also at the lower frequency range.

Application of water contamination detection

Although there are many studies on the detection of food using terahertz technology, limited research has been done on water contamination detection. Recent developments in the integration of biosensor and the vibrational spectroscopic techniques are particularly suitable for water contamination analysis [197]. In [176], the THz-TDS with metamaterial sensors was applied to detect viable and live microorganisms such as yeast cells, molds and bacteria in both ambient and aqueous environments. [184] applied the reflective pulsed terahertz imaging to monitor the water contaminated by sesame oil. In [159], it discriminated the water contamination levels (0%, 0.1%, and 0.2%) in the diesel engine oil samples (SAE 15W-40) by the THz-TDS techniques. In [131], the oil-water mixtures with water content from 1.8%-90.6% were quantitatively recognised by the combination of THz spectroscopy and 3D-printing technology.

2.4.4 Conventional plant disease identification methods

Traditional diagnostic strategies can be mostly classified in two group tools. On the one hand, plant diseases can be detected by means of molecular techniques, such as quantitative polymerase chain reaction (Q-PCR) [195]. However, such studies require a relatively large amount of target tissue and rely on multiple assays to accurately identify distinct plant pathogens. In addition, the high price and short shelf half-life of some molecular biology reagents, such as enzymes and primers, limit the application of traditional molecular methods. Moreover, most of these procedures cannot be applied in the field (on-site detection). On the other hand, some non-destructive methods such as spectroscopy and imaging techniques can be utilized to identify plant diseases caused by pathogens as well as drought from a very early stage. In spectroscopy, the transmission or the reflection spectrum of plants (for example, from their leaves) are measured at different frequencies. The changes in the spectrum can be mapped to different diseases. In the case of imaging, such spectral information is collected with multiple pixels, enabling the generation of images, which also capture the shape of the elements being measured [195].

However, these techniques [195] are limited by the resolution and thus, cannot provide any cellular-scale information about the plants [198]. In addition, these techniques do not consider any environmental influence due to its low photon energy [199]. Furthermore, many spectroscopy and imaging techniques at different frequencies have been proposed in the literature for plant disease detection [200]. Traditionally, the mid and near (close to visible) infrared (IR) spectrum are used, as the very small corresponding wavelength enables the precise detection of relevant molecular and structural information of the plants. In addition to spectroscopic techniques, hyper-spectral remote sensing has also been used to identify stress in plants such as tomatoes, which is governed by late blight disease. Late blight spreads rapidly in tomato fields. It is caused by the fungal pathogen Phytophthora infests. It is found that the near-infrared (NIR) region is the most valuable range of wavelengths than the visible range in identifying plant diseases [195, 201].

Many functional techniques have been suggested such as infrared imaging, hyper-spectral thermal imaging and magnetic spectrum imaging to determine the comprehensive analysis of spatial and spectral information of plant leaves [195, 225]. Although there has been substantial progress achieved by utilizing these techniques, however, there are some limitations related to each technique and can be enhanced further. The resolution of the imaging measurements in the microwave spectrum is limited due to the relatively hefty wavelength that this method employs [195, 225]. In fact, the imaging resolution of microwave radiation is restricted to a nominal wavelength of nearly 2.5 mm . Comparing the image resolution of a microwave with

MRI, it displays the more advanced resolution but has a drawback of limited accessibility and a high-cost. However, the THz spectrum is highly sensitive to water contents and hosts a large number of interesting microscopic phenomena such as inter/intra-molecular motions and Debye relaxation processes, which makes it ideal candidate for this application [195, 225].

2.5 Summary

The recent advancements in the THz technology has a high potential benefit to society, particularly for the public health and environmental protection. Therefore, there is a need of thorough research on the contamination detection of food products and liquid water matters. In this paper, various studies have been analyzed, examined and discussed covering the various aspects of THz systems, components, terahertz spectroscopy and imaging technologies. In addition, we have highlighted the THz detecting technologies and different potential applications that can be effectively used for food and water contamination detection and linked those to future directions. We have also presented some of the open challenges such as distraction and absorption of water, low penetration depth, scattering effect, system compactness and equipment cost that need to be addressed. In a nutshell, the findings of this review paper are expected to be useful for researchers, scientists, engineers, doctors, and policymakers working in the area THz domain for healthcare, fitness and wellness.

Chapter 3

Characterisation of Living Plants Leaves Using THz Waves

3.1 Introduction

Over the past decade, THz technology has seen an increased amount of interest in the scientific community chiefly due to its non-ionising and less pervasive radiation properties [210]. There has been significant progress in tapping the so-called THz gap 0.3 THz to 3 THz of the electromagnetic spectrum. The THz technology has found extensive use in applications such as the imaging of concealed items [125], material characterisation [210], diagnostic applications including treatment of skin and dental care [211,212], effective and quality control of food [213], and telecommunication [164,210,215]. Furthermore, a distinguishing feature of the THz waves is that the water molecules exhibit a strong absorption spectrum in the pertinent frequency range, leading to novel bio-sensing applications.

3.1.1 Related Work

Despite these substantial contributions, the utility of the THz technology in the environmental control/monitoring systems has not been explored in depth, especially for vegetation monitoring [195,216]. Unlike the microwave-based remote sensing techniques, the THz technology can provide detailed insight into the health of a plant specimen in terms of the water content (WC) in the leaves [217]. Plant leaves comprise of a composite biological structure of tissues, distinct bio-molecules like cellulose and synthesis compounds including proteins, carbohydrates and many other molecular weight compounds, as illustrated in Figure 3.1. On an individual basis, they vastly differ in terms of material properties such as relative permittivity [218]. Furthermore, water is not only an essential component but an important nutrient to the process of photosynthesis, and transpiration in the overall process of growth [219]. Due to the high sen-

[0]This chapter is from paper no 1* in publication list

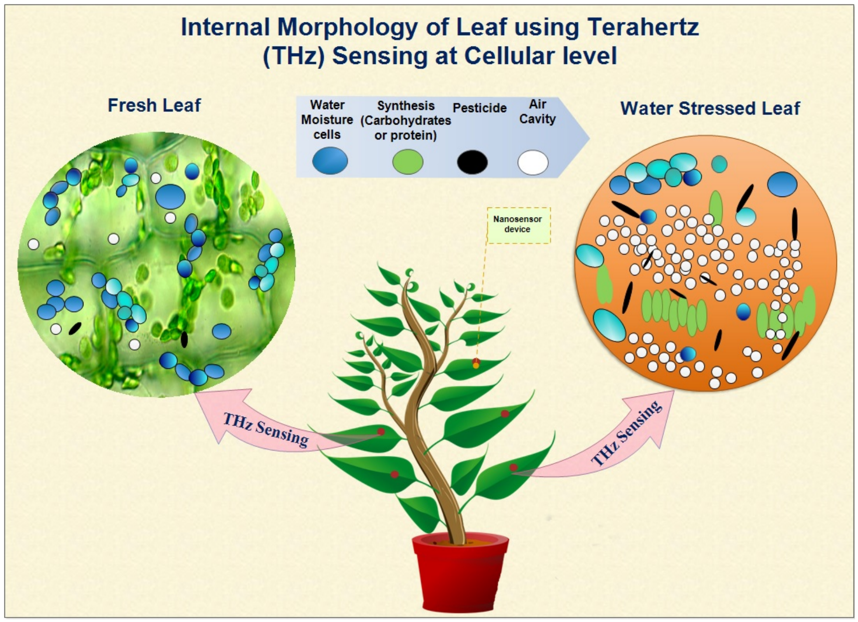


Figure 3.1: Internal morphology of fresh and water stressed leaf using THz sensing.

sitivity and strong penetration feature of THz, it has a strong potential to disseminate through plants leaves at cellular level as shown in Figure 3.1 and can yield significant information of WC in leaves. Hence, it is significant to highlight the frequency dependence of the permittivity of leaves. Designing a smart and plant-specific irrigation system that monitors the leaf WC in a non-invasive manner is, therefore, critical in the current circumstances governed by global climate change that demand water conservation. Over the years, significant contributions have been made [216–220], that address estimating the leaf's WC. There are techniques that offer high reliability, yet they are unsuitable for long-term studies of plant leaves and the validity of measurements cannot be guaranteed due to their disparaging nature. [167, 218, 221–223].

Thermogravimetric analysis has been used for the quantification of WC in plants leaves. However, due to its destructive nature and problems owing to its harmfulness to specimens have markedly reduced the usefulness, and thus is not suitable for estimation of WC in plants leaves. On the other hand, some non-destructive methods have previously been used to determine the water status in plant leaves, which include thermal and hyper-spectral imaging [224, 225], infrared [226] and magnetic resonance imaging (MRI) [227]. However, these techniques are limited by the resolution and thus, cannot provide any cellular-scale information about the plants [225].In addition, these technique do not consider any environmental influence due to its low photon energy [221, 225].

Lately, there has been a growing trend in the field of plant physiology and characterisation of liquid to use THz-TDS [218, 228], which is considered a non-invasive technique, and has been deployed in the field of plant physiology to detect anomalies proactively. Moreover, it has an enormous potential to measure the leaf water status under certain conditions, such as drought stress [218, 220, 222, 229, 230] and dehydration kinetics [225]. In addition, THz-TDS technique can also investigate the structural behaviour and complex traits of leaves under certain environment conditions [221,225]. As compared to others, THz-TDS technique has proven to be more effective and reliable. However, the experimental setup requires a complex configuration of lasers [228].

In this chapter, we present a novel, non-invasive approach and simple setup compared to other techniques to monitor the WC of plant leaves using the scattering parameters of a THz pulse. Using a well-known material extraction algorithm, we computed permittivity (an electromagnetic parameter) from the scattering parameters for eight types of leaves including coffea arabica, aromatic coriander, basil, baby-leaf, pea-shoot, parsley, lamb's lettuce, and baby spinach, which we observed for four consecutive days. It was noticed that after four days, no changes were obtained in the transmission response of all leaves, clearly indicating that WC in leaves were all evaporated. The WC was then gauged from the decrease in the permittivity as the days passed. The rest of the chapter is structured as follows: Section 3.2 describes the experimental setup followed by the material characterisation methods of plants leaves. Section 3.3 presents the measurement results and different parameters are discussed such as permittivity, the effect of weight and thickness, followed by a comparison of transmission response of all eight leaves between day 1 and 4. Finally, summary is presented in Section 3.4

3.2 Methods

3.2.1 Experimental Setup

We used a Swissto12 material characterisation kit (MCK) operating in the THz frequency range to obtain the scattering parameters of the plant leaves. The MCK was attached to a Keysight Technologies N5224A microwave network analyzer (NA), the frequency range of which was shifted in the THz frequency range via a Virginia Diode vector NA extender module WM-250 (WR 1.0), enabling operation in the frequency range of 0.75 to 1.1 THz with a resolution of 2 GHz. The MCK comprised of two conical waveguide horn transitions with further two sections of the low-loss corrugated waveguide. A small aperture between the two low-loss corrugated waveguides allows the material samples to be inserted into the system during the measurement. Moreover, each half of the MCK comprises a waveguide which transitions from a rectangular waveguide at one end to a corrugated circular waveguide at the other. Furthermore, one half of the MCK, is fixed, while the other half is movable (to easily accommodate the insertion of the sample to be measured). To avoid any structural damage while the leaf specimen was clamped in the MCK for observation, we used two PTFE caps that enabled a uniform compression of the samples as shown in Figure 3.2. Prior to the measurement, the setup was configured using the two-port short-open-load-thru (SOLT) calibration technique.

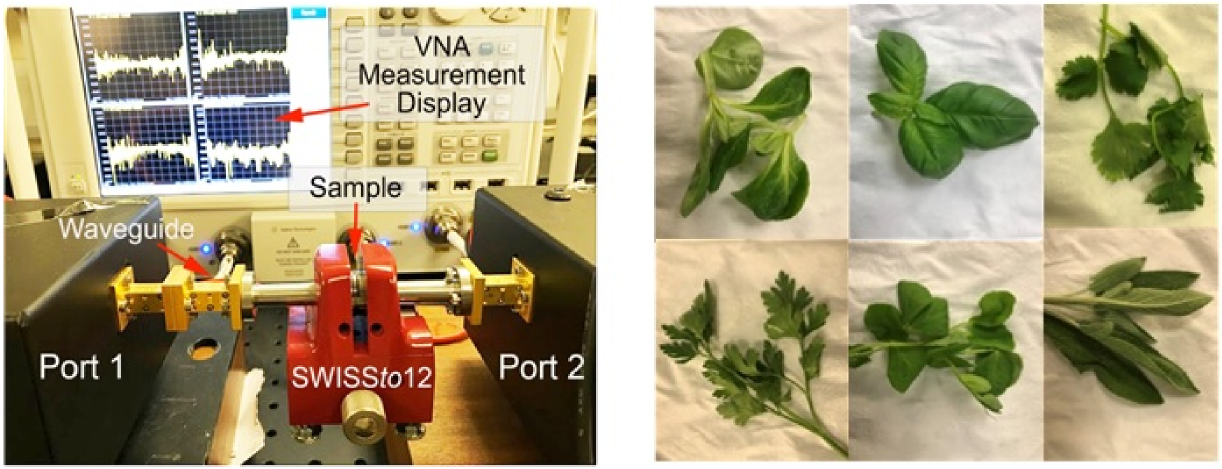


Figure 3.2: Representation of experimental setup used for measurement of leaf sample. The leaf sample is placed between the two PTFE caps fitted to waveguide.

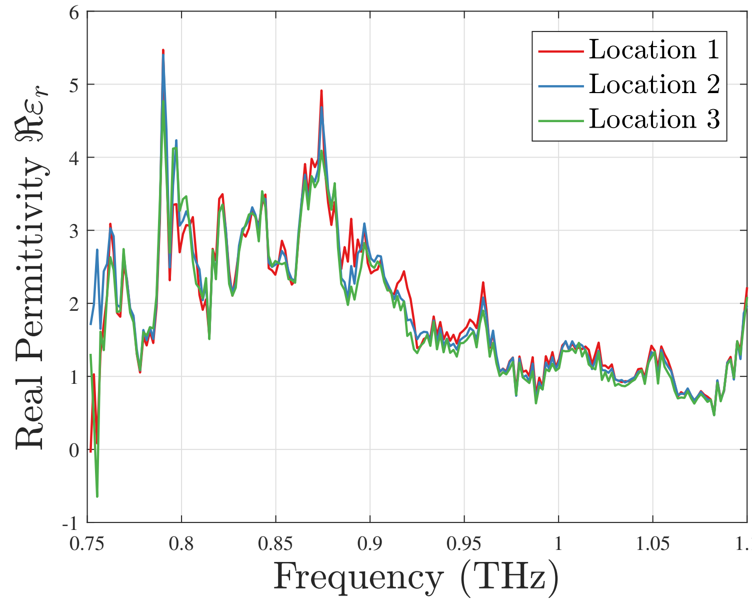
3.2.2 Sample Details

Eight different kinds of pot herbs were used, namely coffea arabica, aromatic coriander, basil, baby-leaf, pea-shoot, parsley, lamb's lettuce, and baby spinach. The fresh leaves were detached from the plants and placed in the laboratory for four consecutive days. The environment temperature for the measurements of leaves was $18.0°C \pm 0.1°C$ and the humidity was between $30\% \pm 2\%$.

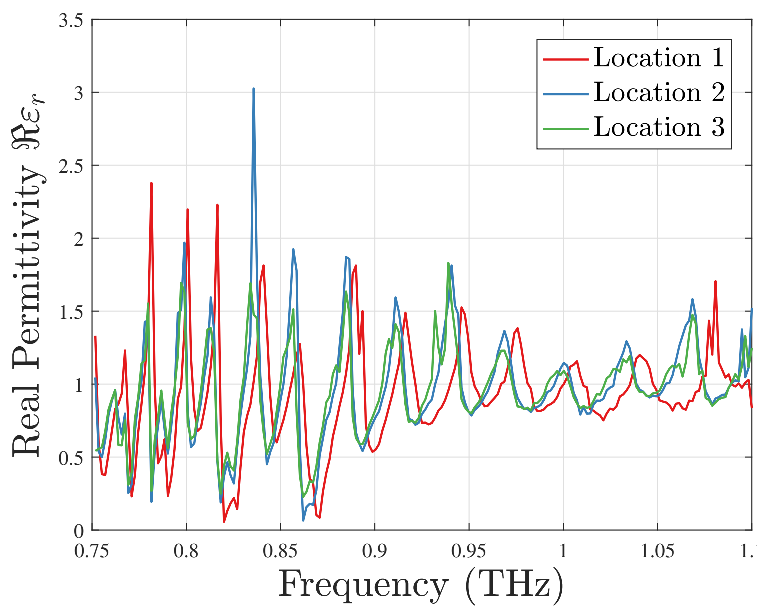
In this study, the weight and thickness of leaves were determined for four consecutive days using a precision electronic scale and Vernier caliper respectively. The leaves' thickness and weight were measured every 2 h during the natural evaporation of leaf moisture. We used a Vernier scale to measure the leaf thickness and this process was repeated to determine thicknesses at three different locations to ensure the thickness was consistent across the surface of a leaf, which was in the range of 40 µm to 4 mm. The weight of leaf was measured using a digital kitchen scale with a least count of 0.1 mg. All leaves were measured at three different locations and on every location, four various orientations were considered to investigate the behaviour of leaves. From these observations, the purpose was mainly to determine any unevenness in the surface of leaves that may result in a change in the scattering response. For further illustration, Figure 3.3 shows the response of coffea leaf at three different locations which suggests that the orientation of leaf does not affect the measurements.

3.2.3 Material Characterization of Plant Leaves

The Nicholson-Ross-Weir (NRW) method [231] is the most common technique in which the dielectric parameters ε_r and μ_r of a planar material are extracted from a two-port vector NA measurement in which the transmission and reflection coefficients are obtained through the S-parameters. This widely accepted approach is particularly attractive for THz applications and



(a) Day 1



(b) Day 4

Figure 3.3: Real part of permittivity of coffea leaves at three different locations taken on days 1 and 4.

rely on the measurement of the complex transmission where the specimen where the sample is fitted inside the wave-guide [231]. This method belongs to the category of frequency-by-frequency material extraction in which every point from the frequency sweep is used. In general, the NRW method generates both the complex permittivity $\varepsilon_r = \varepsilon_r'' - j\varepsilon_r'$ and permeability, $\mu_r = \mu_r' - j\mu_r''$ of the specimen under test. Here, we assume that the leaves are non-magnetic and compute only the permittivity. One of the intrinsic problems of the NRW method is the periodicity of the phase of the electromagnetic wave that leads to ambiguous results. This problem has been discussed at length in other works [232–234]. To rectify this, we follow the step-wise approach in which the phase ambiguity is removed by using the phase delay information from the previous frequency point [235]. In this chapter, we consider a plant leaf as a planar slab of thickness d which is positioned between two air-filled circular waveguides. With the help of an equivalent transmission line model, the reflection (Γ) and transmission (T) coefficients of a semi-infinite slab are expressed in terms of the measured s-parameters, S_{11} and S_{21} as [236],

$$\Gamma = \chi \pm \sqrt{\chi^2 - 1}, \qquad T = \frac{S_{11} + S_{21} - \Gamma}{1 - (S_{11} + S_{21})\Gamma}, \tag{3.1}$$

where the intermediate variable χ is defined as $\left(S_{11}^2 - S_{21}^2 + 1\right)/2S_{11}$. In the case of the slab with a finite thickness d, the transmission coefficient T can be described in terms of the propagation constant, γ as, $T = \exp(-\gamma d)$, which can subsequently be written in the Euler form as $|T|\exp(-j\phi)$ where ϕ denotes the phase term. The propagation constant is then determined using [234],

$$\gamma = \frac{1}{d}\{-\log(|T|) - j\phi + j2\pi n\} \text{ where } n \in \mathbb{Z} \tag{3.2}$$

which results in an infinite number of branches of the complex valued root due to the logarithmic function, demonstrated by the presence of the $2\pi n$ term. The problem of selecting the proper branch is solved by the technique proposed in [235] in which at each frequency point, the phase delay information is recovered from the previous frequency point. If the phase difference, $\phi_i - \phi_{i-1} < \pi$, the method ensures the current branch is selected. The permittivity is then calculated by [237],

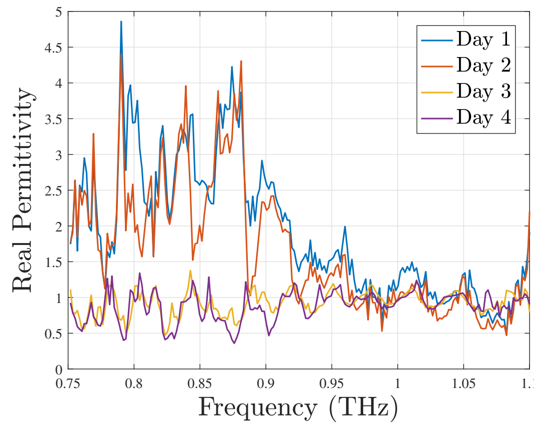
$$\varepsilon_r = \frac{\gamma}{\gamma_0}\left[\frac{1-\Gamma}{1+\Gamma}\right]. \tag{3.3}$$

3.3 Results and Discussion

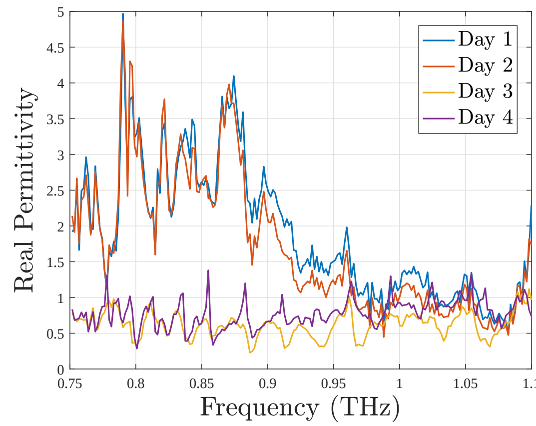
In this chapter, we aimed to determine the electromagnetic properties of leaves including permittivity, and physiological features such as weight and thickness that can affect the WC of leaves. In addition, a strong correlation between the determined properties and WC of leaves was observed. Furthermore, the transmission response of all eight leaves were investigated for four consecutive days.

3.3.1 Permittivity of Leaves

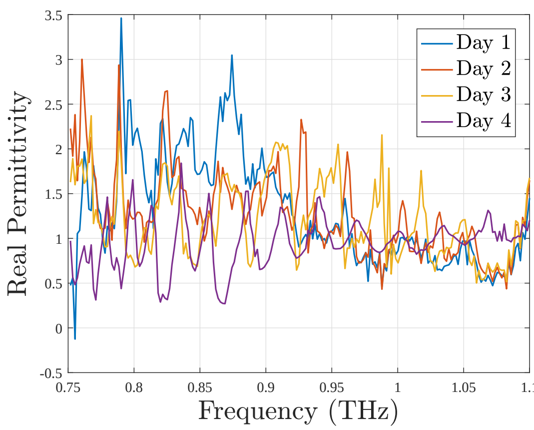
The complex-valued permittivity of eight different types leaves were extracted from measurements taken from three various locations with different percentage of WC in them. Furthermore, on every location, measurements were recorded using four different orientations of the leaves



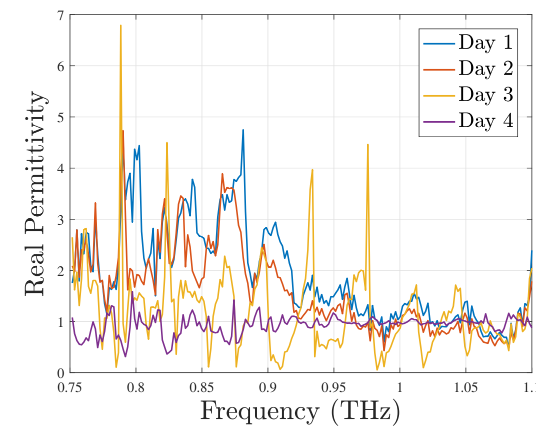
(a) Baby-leaf



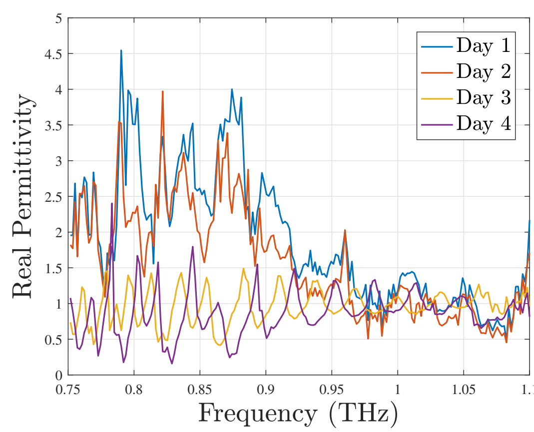
(b) Basil leaf



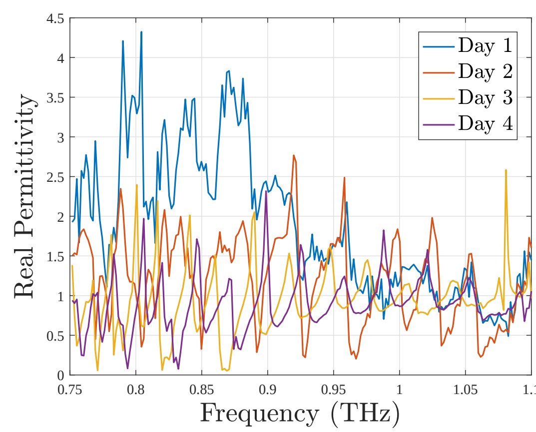
(c) Coffea Arabica



(d) Aromatic Coriander



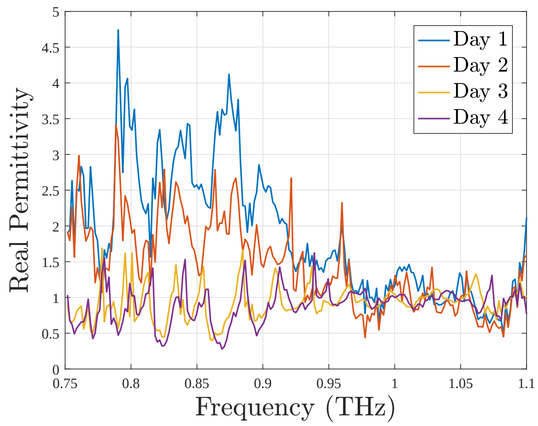
(e) Lamb's Lettuce



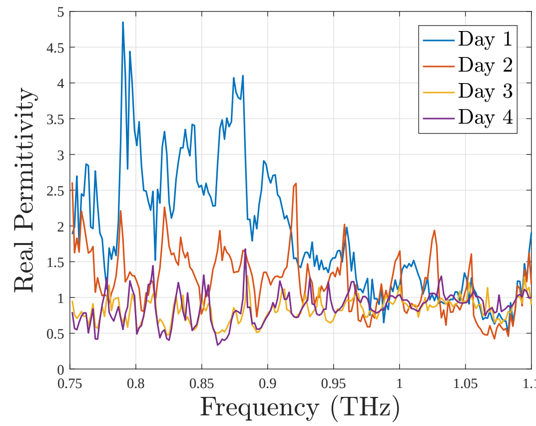
(f) Parsley

Figure 3.4: *Cont.*

to observe any anisotropic behaviour. Figure 3.3a shows that for a coffea arabica leaf, neither the location, nor the leaf orientation had any effect on the real part of permittivity on day 1. However, the effect of location was notable on the day 4 as shown in Figure 3.3b. This drastic decrease in the leaf thickness is responsible for this behaviour. Similar patterns were observed



(g) Pea-shoot



(h) Baby Spinach

Figure 3.4: Real part of permittivity of all eight leaves measured on four consecutive days. Leaves become transparent to electromagnetic waves with the passage of days as seen by the decrease in the permittivity.

for other types of leaves as well. Figure 3.4 shows the real part of the permittivity for all the leaves measured on four consecutive days. It is significant to observe that all the leaves revealed the highest permittivity on day 1 when the WC in fresh leaves was the highest, and as the days progressed, permittivity showed a decrement when leaves became water stressed. Hence, dielectric parameter measurements differed significantly on days 1 and 4 for fresh, and water-stressed leaves. From these observations, it also showed a clear correlation between the permittivity and WC of leaves, i.e., fresh leaves with a higher amount of WC would have a high permittivity and vice versa. From Figure 3.5, it is evident that the real part of permittivity shows a strong decaying correlation with WC. As observed in Figure 3.4, various leaves also showed distinct decreasing responses from each other, attributed to different lead composition and structure. In this study, we have also shown that the real part of permittivity can be used to classify leaves, that are described here with an accuracy of 98.2%.

3.3.2 Estimating Leaf Water Content

In this study, WC in leaves was observed by determining the physical parameters such as weight and thickness. Referring to the weight of leaves, the time duration between the two weight measurements were maintained from two to three hours for all four consecutive days. It was noted that there was significant decrease in the weights of some leaves as shown in Figure 3.6 on day 1, i.e., basil, baby-leaf and pea-shoot, whereas, other leaves displayed a slow decreasing trend in weight loss of leaves as days progressed. This clearly indicated that the moisture in leaves evaporated more rapidly on the day 1 and 2 compared to day 3 and 4, thereby creating more air cavities in the leaves. To assess the variation of leaf WC during the leaf's water evaporation

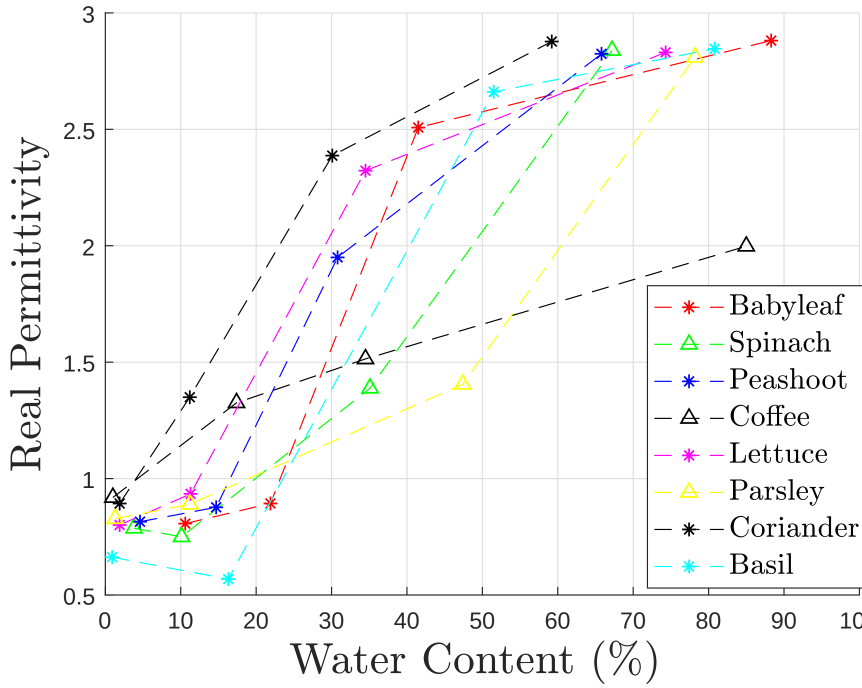


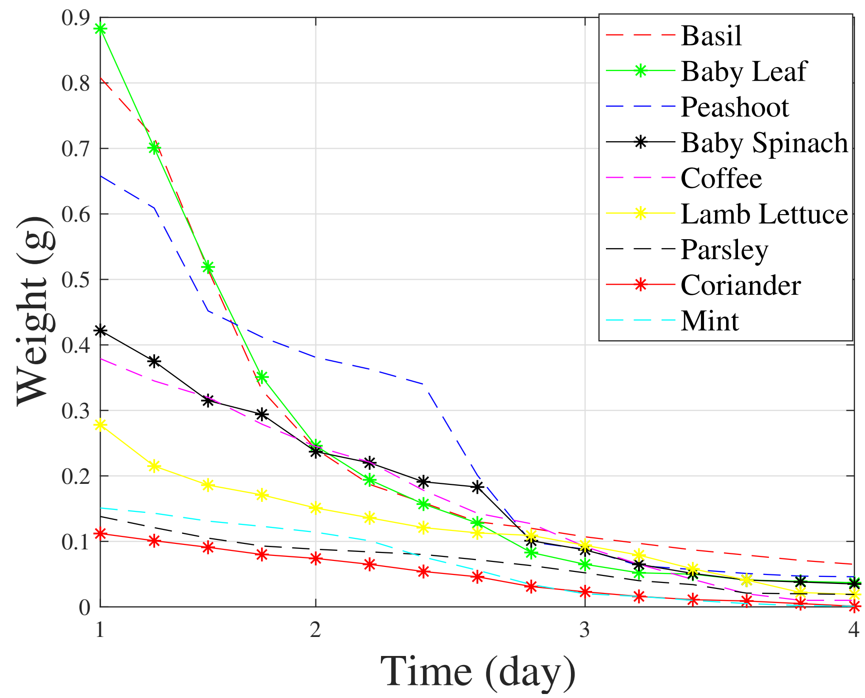
Figure 3.5: Correlation of permittivity with loss of WC in leaves.

process, the measurements were translated into WC using 3.4 [167, 238].

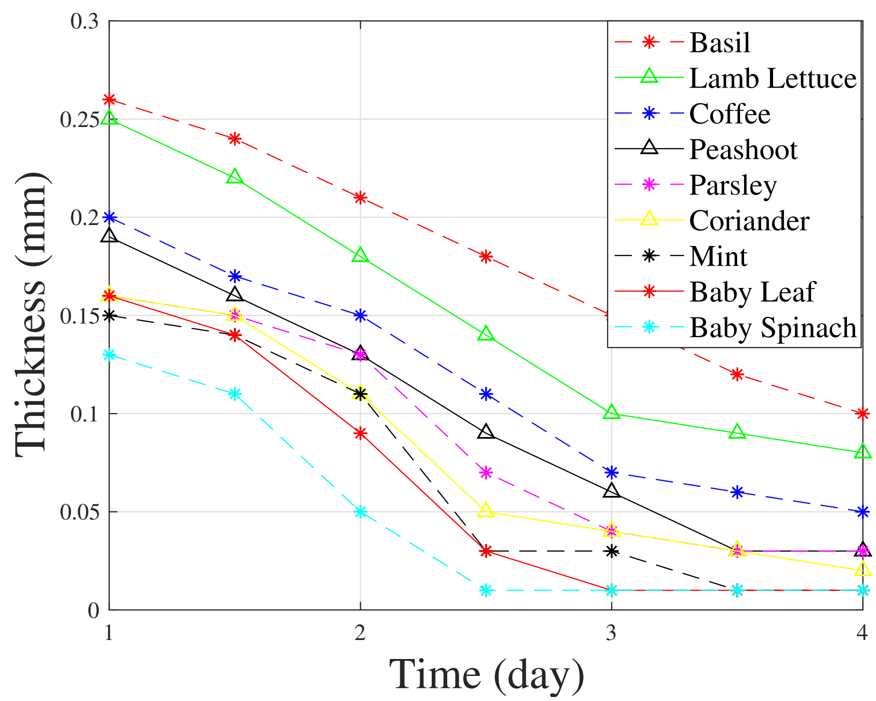
$$WC = \frac{W_{\text{time}} - W_{\text{dry}}}{W_{\text{fresh}}} \times 100 \tag{3.4}$$

where W_{fresh} is the weight of the fresh leaf, W_{time} is the weight of a leaf measured over time and W_{dry} is the weight of a dry leaf. In the beginning, the WC loss observed between the two hours on day 1 was found in the range of 5% to 22%. At the end of the investigation on day 4, this loss increased during the natural evaporation of leaf moisture and was established in the range of 83.12% to 99.33%. The obtained percentages loss of WC can be validated with Figure 3.6a. Considering this discussion, it also showed a significant correlation with the real part of the permittivity which was the highest on the first day when the weight of leaves was considerably high compared with the fourth day as shown in Figure 3.5.

The thickness of all the leaves was carefully determined to avoid any excess pressure to the samples that would cause disturbances in the morphological structure of the leaves, changing the dielectric properties of the samples as a result. As seen in Figure 3.6b, the thickness of leaves was considerably higher on day 1, implying a greater WC in fresh leaves compared to day 4 when mostly, all leaves were dried out. From this significant and meaningful observation, it was concluded that dehydration of leaves with passing days affected the thicknesses to a substantial degree. On day 4, some leaves stayed invariant or slight changes occurred in the thickness of leaves i.e., coriander and spinach as shown in Figure 3.6a. These transformations in the thickness of leaves evidently showed that WC in coriander and spinach leaves had evaporated to



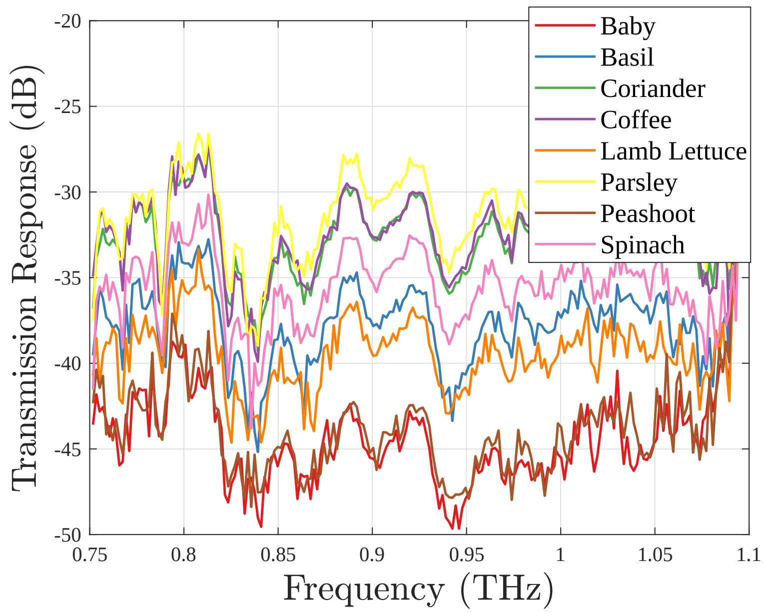
(a) Weight



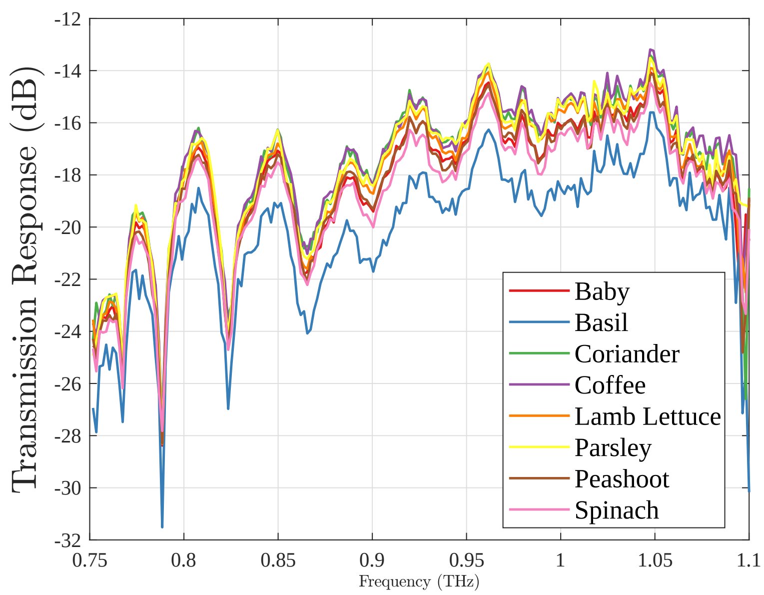
(b) Thickness

Figure 3.6: Change in the physical properties of leaves with time.

the maximum on day 4 and no further variations could be observed in thickness of leaves.



(a) Day 1



(b) Day 4

Figure 3.7: Transmission response of leaves on first and fourth days.

3.3.3 Evaluation of Leaf Transmission Response

In this section, transmission responses of all leaves were observed on day 1 and 4 as shown in Figure 3.7. The transmission response basically indicate the waves that passes through the plant leaves revealing the freshness and staleness of plants leaves. It was noticed that on day 1, attenuation's of all leaves were substantially high due to the presence of higher WC in tissues of leaves, which resulted in a higher absorption and lower transmission response. Moreover, on day 4, a substantial degree of increment in transmission response was observed as WC in leaves had evaporated to large extent, which resulted in a decrement of weight and thickness of leaves and eventually, less absorption occurred at this time. Figure 3.7 exhibited a strong correlation of transmission response with WC, weight and thickness of leaves. Baby-leaf exhibited a lower transmission response on day 1 compared to others, reflecting a higher presence of WC in leaf, which resulted in higher absorption. Contrarily, parsley displayed an increment in transmission response due to the presence of less WC in the leaf and hence, producing low absorption compared to other leaves.

3.4 Summary

In this chapter, cost-effective, novel and non-invasive technique that can determine the internal morphological structure, characterization and mainly determine the precise estimation amount of WC in various living plants leaves at cellular level using using THz sensing to main the healthy physiological growth and lessening the early inception of age dependant disease. The electromagnetic properties of eight types of leaves were determined for four consecutive days through the measured scattering parameters. The weight and thickness of the leaves were also recorded at the same time. We observed that the leaves became increasingly transparent to the terahertz (THz) waves through the course of four days experiment, as seen by the peaks in the real part of permittivity. Similar decaying trends were observed in the peak values of the real part of the extracted relative permittivity as the decreasing weight due to loss of WC. The significance of this chapter lies in the simple, cost-effective technique and other advantages such as: (a) The study proposes a unique technique to characterise and estimate WC of eight various leaves in terms of electromagnetic parameters at THz frequency range from 0.75 to 1.1 THz. (b) The electromagnetic parameters are measured in simple, fast, and non-invasive manner using a THz material characterisation kit. Moreover, The structural integrity and configuration of leaves were also considered by employing two polytetrafluoroethylene (PTFE) caps which were fitted internally to the waveguide. (c) In this chapter, we establishes a notable correlation between electromagnetic parameters with WC in leaves i.e., change in the WC of leaves is reflected in the electromagnetic parameters at certain frequencies. In the age of a climate change driven water conservation, the proposed scheme can be used to design efficient irrigation systems on-site without any need to remove the leaves from plants.

Chapter 4

Machine Learning for Precision Agriculture Using THz Waves

4.1 Introduction

The growing consciousness of fruits and vegetable quality in recent years, while utilizing natural resources such as water consumption [239], strongly demand viable and feasible techniques to detect early symptoms of plants drought stresses [239, 240]. The recent climate transformations and growing deficiency of water resources have posed enormous challenges, particularly in the applied plant biology sector [241, 242]. In this regard, much efforts have been geared by researchers, horticulturists, and plant physiologists at various levels in the plant science sector, towards developing feasible strategies for non-invasive techniques [125, 243–246] in monitoring the health status, and biological traits of leaves to sustain crops productivity. Hence, a precise estimation of WC at a cellular level in plants leaves is of high-importance to growers, and cultivators to take appropriate and efficient measures by facilitating them with appropriate amounts of resources inputs, i.e. water and nutrients to maintain healthy physiology [125, 241–246].

4.1.1 Related Work

In recent years, many conventional techniques [125, 244–250] have been suggested for accurate estimation of WC in leaves and studied the morphological structure of leaves in detail. These methods including MRI, NIRS, hyper-spectral imaging [245–250] have offered better reliability but have been suffered by some limitations and considered as time-consuming, and unsuitable for long-term studies due to disparaging nature [246–250]. Besides, some others non-destructive techniques such as thermal imaging [249–253] have been proposed, and yet they too are littered with limited resolution and sensitivity issues, and transpired as inappropriate for detecting mon-

[0]This chapter is from paper no 3* in publication list

itoring information on water dynamics and diminutive changes at the cellular level [250–253]. Consequently, the evolving applications of THz-TDS technology, which is considered as non-intrusive, has been deployed in the field of plant physiology to detect anomalies proactively and investigate the structural behaviour and complex traits of leaves under the particular environment [253–255]. This technique is proven to be more effective and reliable compared to other approaches. However, it is a costly technique, and on-site access is limited [253–255].

4.1.2 Objectives and Contributions

Meanwhile, terahertz (THz) technology has been widely used in diverse field applications such as diagnostic applications of dental and skin-care [242, 256, 257], unseen hazard items [258], material characterizations [242, 243], and telecommunications [243, 257]. However, researchers from plants science sector are of the strong view that its potential to disseminate through plants sector is still to be thoroughly revealed, considering it as a new source of vital improvements for the agricultural sector [242, 258]. The aforesaid prevailing challenges in exploring the spectral analysis of WC in leaves using THz have immensely engaged numerous scientists and captivated researchers from diverse fields. Moreover, evidence from multi-disciplinary agri-technology studies show that reliable and early detection of WC in plants leaves at a cellular level can drive agricultural productivity and optimize the economic benefits [247–249]. For this purpose, ML applications create an innovative opportunity to unravel, quantify, and understand data-intensive processes in agricultural operational environments [259] by providing precise estimation of WC in leaves. In recent times, the applications of ML have been immensely used in various scientific fields [259] such as healthcare sector, food security, meteorology, medicine, meteorology, economic sciences [259]. Furthermore, researchers are very keen to discover its possibilities, specifically in modern digital agriculture systems to develop intelligent management of plants by applying the water distribution effectively [259].

Considering the sensory characteristics of plants leaves, water is essential to the overall growth, transpiration, and nutritional process of plants leaves [247]. Therefore, timely delivery of the appropriate amount of resource inputs such as water and its precise quantification can be very beneficial to drive and sustain overall crops productivity in an advanced agricultural system [247]. This chapter presents a state-of-the-art method to closely monitoring the water dynamics in leaves using the scattering parameters of THz pulse waves through ML. In our study, we demonstrated that there is a clear relationship between the parameters of the pulse wave and the plants WC within a frequency range from 0.75 to 1.1 THz. We have performed in-lab experiments using three different plant leaves, including coffee, pea-shoot, and spinach for four consecutive days. Subsequently, the data is pre-processed for feature extraction and is fed to our proposed ML algorithms for automated classification of WC on different days. Three different machine learning algorithms such as support vector machine (SVM), K-nearest neighbour (KNN) and decision-tree (D-Tree) algorithms were used to determine the accuracy of

WC in plants leaves.

The overarching aim of this study is to estimate and predict the future trends of WC in plants' leaves in an automated fashion using THz pulse waves, which is indicative of the health status of the plants. For this purpose, we have extracted time and frequency domain-features of THz pulse wave and use it to train ML models to monitor WC in coffee, pea-shoot and spinach more precisely. By performing the leave-one-observation-out cross-validation, we strongly feel that our proposed model has the capability to monitor the WC future trend proactively. Hence, it can save crops from stresses by taking timely action, which will ultimately help to increase yield production and optimize economic benefits. The rest of the chapter is structured as follows: Section 4.2 presents methods and the implemented methodology for data collection and pre-processing 4.2.3, This is followed by the description of the feature extraction technique in section 4.2.4. Section 4.3 describes the proposed classification algorithms. In sections 4.3.1 and optimal feature selection are discussed. Section 4.4.5 shows the analysis of three classifiers results. Finally, the conclusion is drawn out in section 4.4.6.

4.2 Proposed Technique

4.2.1 Experimental Setup

In this setup, a THz Swissto12 MCK [20] was employed to obtain the scattering parameters of three plant leaves . The MCK was connected to a Virginia Diodes Analyzer (VNA) extender WM-250 (WR1.0) which operated in the frequency range of 0.75 THz to 1.1 THz. The structural integrity and configuration of leaves were also considered by employing two Polytetrafluoroethylene (PTFE) caps which were fitted internally to the waveguide and could provide a consistent compression to samples, as shown in Figure 4.1. Prior to any measurements, the setup was calibrated using the two-port short-open-load-thru (SOLT) calibration technique to confiscate any unwanted errors or noise that may have occurred while performing measurements.

4.2.2 Sample

Three various kinds of plants leaves were used for measurements are coffee-arabica, pea-shoot and baby-spinach. In this study, these fresh leaves were detached from plants, which were fully grown and nurtured in Rouken Glen Farm, East Renfrewshire, Glasgow. According to the status of these plants, these leaves grew well with no pests or disease and were kept in the laboratory under the environment temperature of 18°C ± 0.1°C, and the humidity was between 20% ± 2%. The thickness and weight of the leaves were continuously monitored for four consecutive days using the Vernier calliper and electronic scale, respectively. The thickness of leaves appeared to decrease substantially due to leaf dehydration. Hence, variations in WC of leaves was the key factor that caused spectral variation in measurements, as shown in Figure 4.1. In addition, all

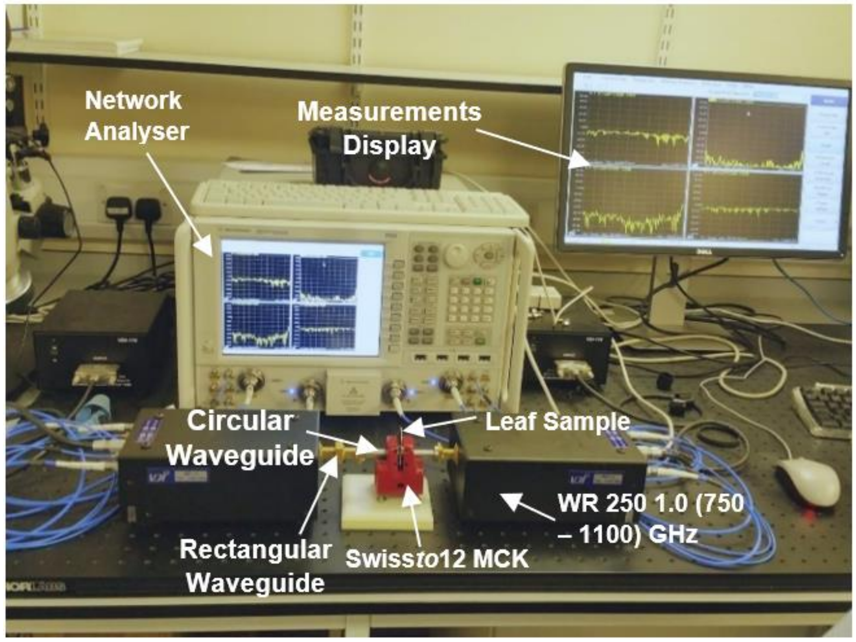


Figure 4.1: Experimental setup of Swissto12 Material Characterisation Kit (MCK) system used for measurements of leaves in the frequency range from 0.75 to 1.1 THz.

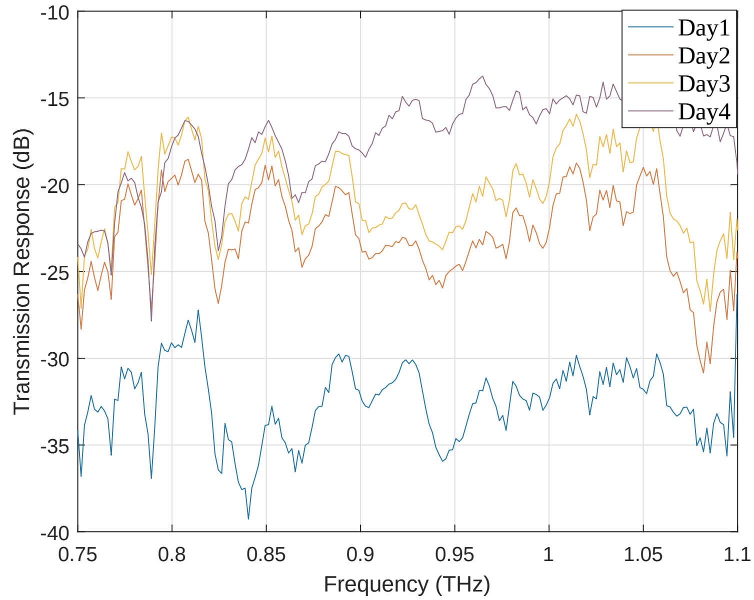
Table 4.1: The number of observations of three plants

Leaves	Number of Observations
Coffea Arabica	127
Pea shoot	76
Baby Spinach	54

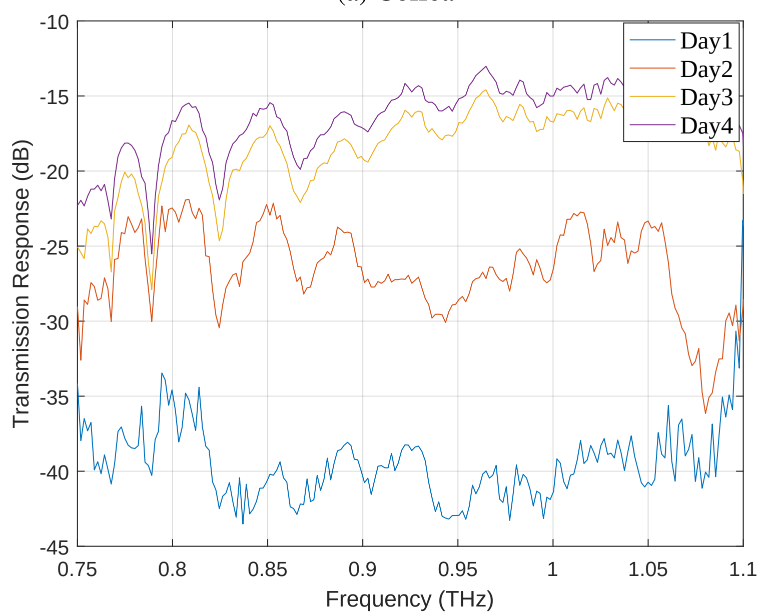
leaves' thickness and weight were measured at three various locations after every 120 minutes during the natural evaporation of WC to analyse the unevenness in the surface of leaves.

4.2.3 Procedure for Data Collection and Pre-Processing

The measurements data for all three fresh plant leaves were obtained in the Radio Frequency Laboratory at the University of Glasgow for four consecutive days. For each observation, all distinct leaves were placed between the two wave-guides, and observations were recorded. Both the transmission coefficients (S_{12},S_{21}) and reflection (S_{11},S_{22}) were determined from the measurements. The overall experimental setup for measuring the WC of all fresh plants' leaves is shown in Figure 4.1. In this work, the focus was mainly to consider the transmission response as features for all three leaves and is shown in Figure 4.2. Every day, the duration of measuring



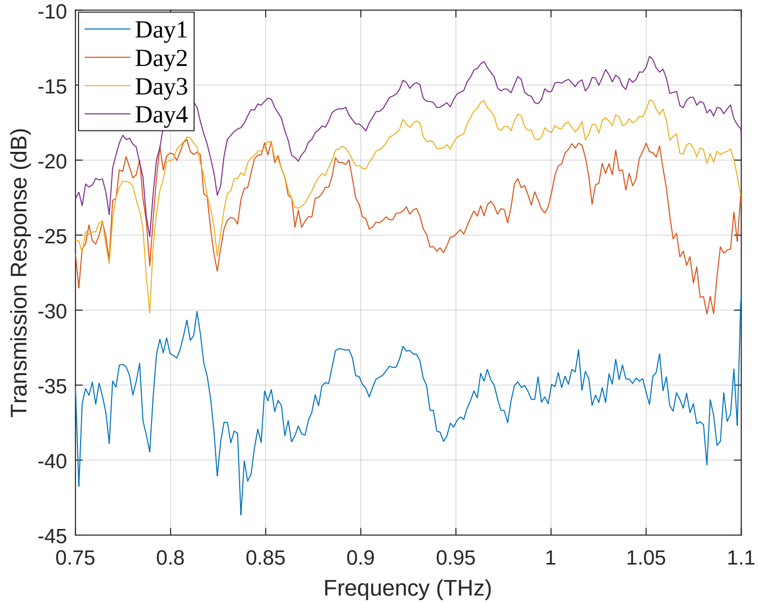
(a) Coffea



(b) Peashoot

the THz transmission response was approximately 9-10 hours to observe various degree of WC in all three leaves was, and measurements were recorded after every 120 minutes. This process was repeated for four consecutive days. Hence, the total number of observations collected for coffee, pea-shoot and baby-spinach for continuous four days are listed in table 4.1.

Table 4.1 shows the difference in the number of observations of leaves which indicates that each leaf had a variable degradation in WC during the four days of measurements. On each day, 10 rounds of weight measurements were recorded over the span of four days and converted into WC using equation 3.4 [254, 258, 260]. Upon close analysis of Figure 4.2, it was depicted



(c) Baby-Spinach

Figure 4.2: Transmission response of coffee, pea-shoot, and spinach leaves observed on four different days in the frequency range of 0.75 to 1.1 THz.

that coffee, pea-shoot and baby-spinach leaves exhibited distinct responses on all four days. On day 1, the transmission response for all leaves was significantly low due to the presence of high volumetric WC in leaves. Notably, pea-shoot revealed a response in the range of -40dB to -45dB reflecting a distinct characteristic from other leaves. The difference in transmission response also highlighted a physiological process, affecting the variability of the water dynamics in these various plants leaves.

4.2.4 Feature Extraction Methods

During the THz experimental campaign of measuring the transmission response of leaves, the observations spawned by Swissto12 MCK were erratic (exhibiting unwanted excessive variations), especially at both ends of frequency range from 0.75 to 0.80 THz and 1.05 to 1.1 THz as shown in Figure 4.2 [261]. The effect of this undesired noise could be crucial and may have produced false observations about the WC in leaves in rest of the frequency region. Inevitably, it would have produced counterfeit classification results by classifiers about the quantification of WC in leaves. Furthermore, any erroneous estimation of WC in leaves would ultimately affect their overall biological and physiological process of growth. Hence, it was significant to discover the sensitive frequency region (SFR) with the minimum effects of any unwanted errors in the overall observation data. Therefore, the target response region (TRR) was established where observations could be visibly distinguished without any overlap for leaves on all different days. The TRR for coffee leaf was selected in the range of 0.82 to 1.05 THz, as shown in Figure

4.3. Furthermore, useful observations would also have a fruitful impact on overall classification outcome.

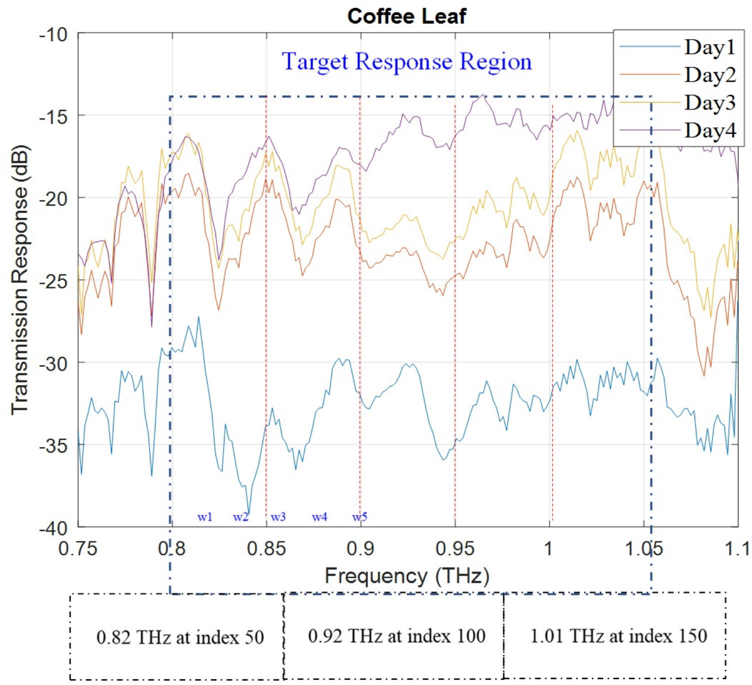


Figure 4.3: Identification of target response region (TRR) to consider only relevant and important features for the feature extraction process.

Researchers have suggested and applied many features extraction techniques to execute the classification accuracy [262]. In this work, observations recorded were in the frequency domain had to be converted into time and time-frequency domain to further minutely observe the behaviour of WC in in various leaves by analysing statistical features. Hybrid combinations of multi-dimension features domain would have a favourable response in classification accuracy by reducing overall dimensions of initial features [262]. The frequency-domain was converted into the time domain and time-frequency domain by applying Inverse Fast Fourier Transform (IFFT) and Short-Time Fourier Transform (STFT) respectively [262]. The list of different domains is summarised in Table 4.2. Hence, out of 201 features, only 25 significant features were considered which comprised of 11, 10, and 5 in the time-domain, frequency domain, and time-frequency domain respectively as indicated in Table 4.2. The block diagram of the proposed classification system for different days based on multi-domain features extraction approach is shown in Figure 4.4.

4.2.5 Evaluation of Frequency Features Extraction

Since the data obtained from VNA was in the frequency domain, it was significant to focus mainly on the region that gives the maximum and the accurate information about the existence of WC in all three leaves. For this purpose, as mentioned earlier, TRR was mainly required.

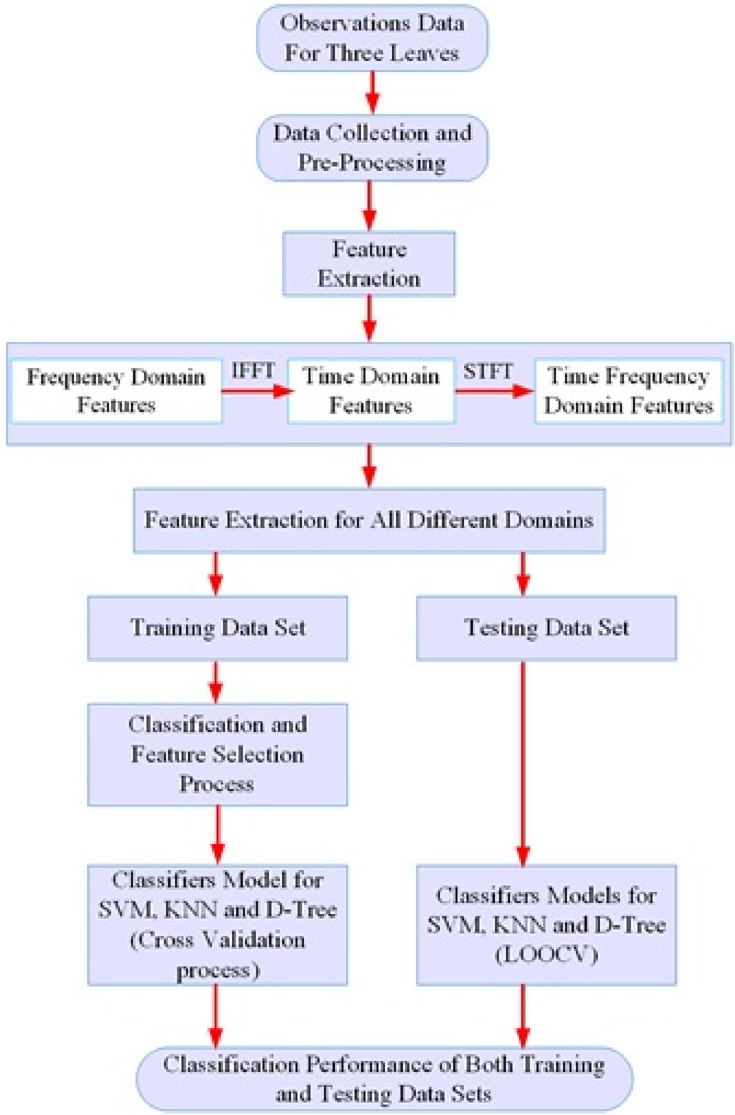


Figure 4.4: The flow chart of the proposed algorithm implementation process.

In this regard, five windows bins with a width of 20 were initiated in the middle region (0.92 THz at index=100) and symmetrically expanded to both sides of the frequency region. From Figure 4.3, the data under the observation of the selected area can be seen, and was applied to the rest of two leaves as well. In addition, the frequency domain features included a cross-power spectral density and variance of power spectral density and is given by the equation 4.1 [263] respectively. From the equation 4.1, $Y_l^n(a)$ represents the transmission response of the reference signal. In equation 4.1, $T(a)$ implies the transmission response of l -th leaf on an n$^{\text{th}}$ day. Here, ‘w’ is considered as the width of the frequency window as depicted in Figure 4.3.

$$\begin{aligned} &\text{Var}\{Y_{ll}(a)\} = \tfrac{1}{w}E\left[\{Y_l^n(a) * .Y_l^n(a)\}\right] \\ &\max\{Y_{lm}(a)\} = \max\left(\tfrac{1}{w}E\left\{\left(T(a) * .Y_l^n(a)\right)\right\}\right) \end{aligned} \tag{4.1}$$

Table 4.2: Feature Extraction Technique for all Three Leaves

Time domain (Statistical Features)	Serial No.	Frequency domain Features	Serial No.	Time-Frequency domain	Serial No.
No. of Features	11	No. of Features	10	No. of Features	(4)
Mean	1	CPSD (D = 20)	12	Subband 1	22
Variance	2	CPSD (D = 40)	13	Subband 2	23
(MAD)	3	CPSD (D = 60)	14	Subband 3	24
Skewness	4	CPSD (D = 80)	15	Subband 4	25
Kurtosis	5	CPSD(D = 100)	16		
Standard deviation	6	PSD(D = 20)	17		
MAV	7	PSD (D = 40)	18		
75^{th} (Q_3)	8	PSD (D = 60)	19		
25^{th} (Q_1)	9	PSD (D = 80)	20		
PCC	10	PSD (D = 100)	21		
IQR	11				

4.2.6 Evaluation of time features extraction

For statistical features, the transmission response of time-series of THz pulse was observed from days 1 to 4, indicating any possibilities of WC in leaves. Therefore, it was required to convert frequency domain data into the time-domain features to observe meaningful THz pulse. For this purpose, 11 time domain features were employed and they are mean, median, mean of absolute value (MAV), standard deviation (STD), mean of absolute deviation (MAD), skewness and kurtosis, Pearson correlation coefficient (PCC) [264], 25th percentile (Q1), 75th percentile (Q3) , and Interquartile Range (IQR) [265]. In which, mean and standard deviation were particularly useful to provide significant information about the distribution of data [261]. Skewness produced meaningful information about the irregularities of the examined area and its distribution around its mean [265, 266]. Moreover, kurtosis presented a measure of evenness relative to a standard distribution [265]. Q3 and Q1 showed how the observation data were dispersed in the two sides of the median. PCC was used to measure the linear relationship between the time-domain waveforms of the sample and the reference signal [265]. IQR was also used to measure the variability of the dataset and shows the difference between Q3 and Q1 while measuring the data distribution set. This information was also helpful in terms of excluding irrelevant data [265].

4.2.7 Evaluation of Time-Frequency Features Extraction

The demand for considering time-frequency technique such as Short-Time-Fourier-Transform (STFT) and Wavelet Transform (WT) was mainly to obtain the detailed information of THz pulses in this domain [267] The WT technique was more appropriate to analyse short-term THz pulse produced because of any diminutive variations occurred at the cellular level, reflecting an information of WC in leaves. After the de-noising process, the wavelet spectrum features were extracted by considering the power of various sub-bands at different levels as defined in equation 4.2 to extract the time-domain features [268, 269].

$$E(j,i) = \frac{1}{N}\sum_{k=1}^{N}[P_k(j,i)]^2 \tag{4.2}$$

In the above equation, j denotes the level of wavelet decomposition and i^{th} indicates as the sub-band and 'N' is the number of wavelet coefficients $P_k(j,i)$ is basically the wavelet coefficient vector of i^{th} sub-band in the j^{th} Hence,$E(\mathrm{j},\mathrm{i})$ denotes the average power value of ith sub-band at the jth level. Table 4.2 summarised the features extracted from time, frequency, and time-frequency domains. Each feature is assigned one serial number from 1 to 25, in which, 1-11, 12-11 and 22-25, were the serial numbers of time-domain, frequency-domain, and time-frequency domain features, respectively.

4.3 Classification Algorithm for Proactive Estimation of Water Content in Leaves and Parameters Selection

In this section, the significant of optimum parameters were determined for three classifiers including SVM, KNN, and D-Tree. In addition, on the basis of suitable parameters selection, classification algorithm was developed, and its performance was evaluated for precise estimation of WC in leaves.

4.3.1 Selection of Optimal Parameters Values

For accurate classification results, it was significant to have optimal parameters for classifiers. Here, three classifiers which include SVM, KNN and D-Tree were considered for precise estimation of WC in three leaves from day 1 to 4. For each classifier, a series of values for tuning the process with optimal parameters were determined to achieve the highest overall classification accuracy and performance of classifiers were also analysed. For SVM, two parameters i.e. the optimum parameters of cost (C) and kernel width parameter (Υ) are required to be set when applying the SVM classifier with radial basis function (RBF) kernel to achieve the optimized SVM algorithm [270]. The 'C' parameters helped to decide the actual size of mis-classification permitted for non-separable training data and adjusted the rigidity of the training data [271].

Larger values might lead to an over-fitting model and vice versa. The kernel width parameter (Υ) facilitated the shape of the class-dividing hyper-plane, and increasing or decreasing the value of (Υ) could influence the shape of the class-dividing hyper-plane, and it eventually disturbed the classification accuracy. For this purpose, a series of values were assessed and to establish the most suitable value for 'C' for available data, and finally "1" was chosen for 'C', and "0.38" was selected for (Υ).

The basic theory behind the KNN was to discover a group of 'k' samples that appeared to be nearest to the unknown samples [270]. From k-samples, the label of unknown samples could be determined by evaluating the average values for class-attributes [270, 271]. Thus, tuning this fundamental parameter of k-sample played a significant role in achieving the ultimate performance of this classifier. For this purpose, a different range of values was established, and finally, it was settled in the range from 1 to 5 to recognize the optimal 'k-value' for all training sample sets. For D-tree, again the various range of numbers for splits in D-test was analysed for the available data to identify the optimum parameter. Eventually, it was set to 5, and the rest of the settings were retained as default values for this classifier.

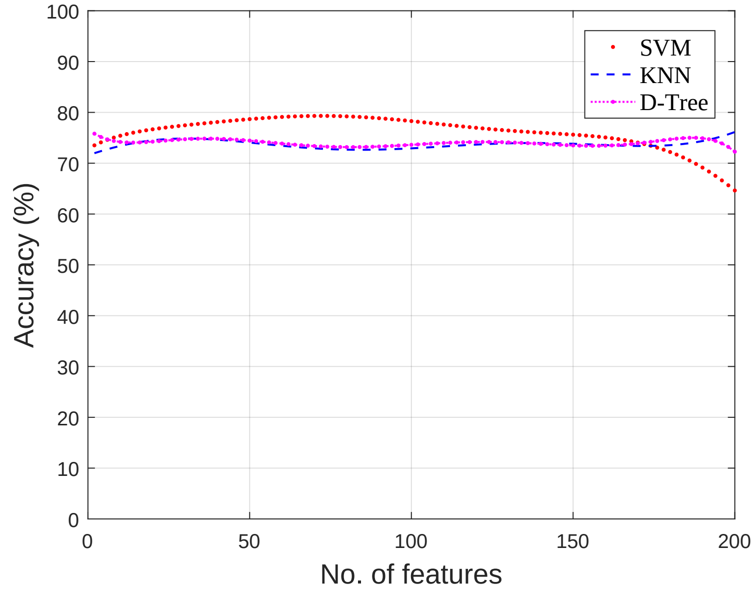
4.4 Results

4.4.1 Classification Accuracy and Features Selection

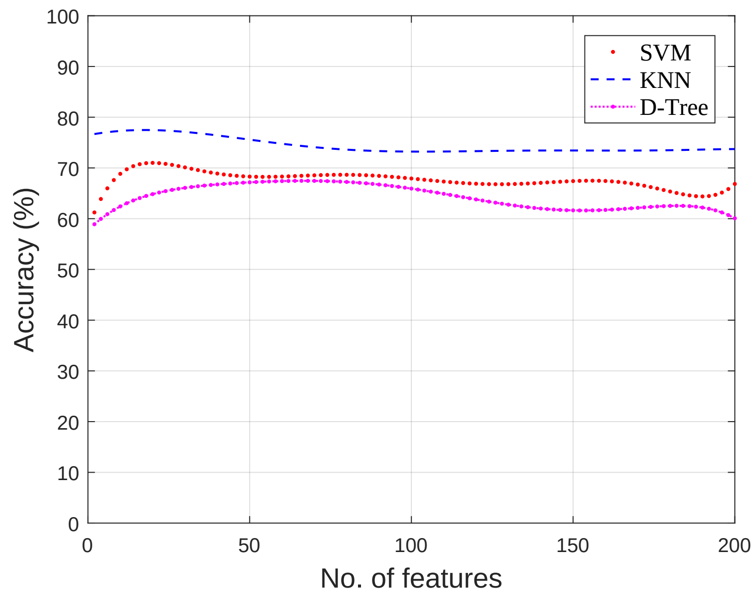
In this study, the performance of proposed classifiers including SVM, KNN, and D-Tree was assessed on raw data and on individual domain features. Furthermore, all classifiers showed distinct performances on individual domain features. Henceforward, classification accuracy for a hybrid combination of all three domains was also obtained. Towards the end, features selection was illustrated using the various state-of-the-art techniques.

4.4.2 Assessment of Classifiers on Raw Data

Before processing the classification accuracy of raw data, the frequency range of 0.75 to 1.1 THz was considered for executing classifications. Also, all observations were taken as separate features and performance of the classifiers were tested on all features. The main aim here was to evaluate the classifier response by examining all observations of three leaves at different days at every frequency point. Hence, three classifiers, including SVM, KNN, and D-tree performances were tested to estimate the WC in leaves more accurately and precisely. The classifiers were trained and validated using a k-fold and feature set was partitioned into 10 "folds" randomly. The observations data was partitioned into 70% and 30% training and testing data, respectively. Table 3 listed the average classification accuracy results of all three classifiers.

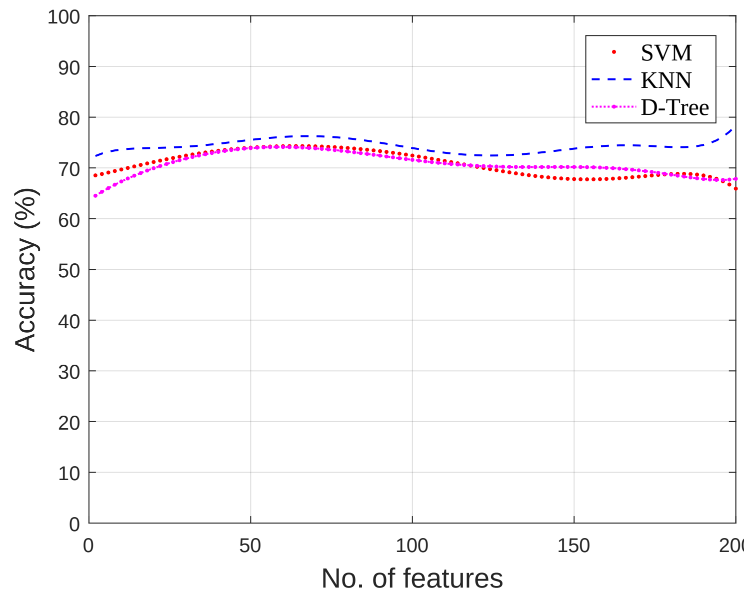


(a) Coffee



(b) Peashoot

By close investigations of results in Figure 4.5 and Table 4.3, it was depicted that classification accuracy for all leaves found in the range of 70%-75%. This low accuracy reflected some redundant or irrelevant features in the overall 201 features points, which badly affected the classification accuracy. Therefore, the performance of all three proposed classifiers could be improved by reducing undesired features and selecting more meaningful and informative features to produce an accurate estimation of WC in all three leaves. Thus, the purpose of observing the performance of the classifiers on raw data was mainly to explore the TRR, as explained in the previous section.



(c) Baby-Spinach

Figure 4.5: Classification performance of raw data for coffee, pea shoot and spinach leaves considering all features from 0.75 THz to 1.1 THz.

Table 4.3: Raw data classification Results for Three Leaves

Classification Accuracy (%)	Time domain features (11)	Frequency domain features (10)	Time-frequency domain features (4)
SVM	80.22%	76.26%	75.78%
KNN	75.1%	72.95%	74.98%
Decision Tree	76.24%	69.58%	76.93%

4.4.3 Assessment of Classifiers for Individual and Hybrid Combination Features

Once the parameters were set for all classifiers, its performance was investigated on different domain features individually and a hybrid combination of all three domain features. So, its performance accuracy was accomplished, and Table 4.4, 4.5, and 4.6 demonstrated the classification accuracy results for coffee, pea-shoot and baby spinach, respectively. The classification accuracy results were obtained for 25 extracted features which were obtained using feature selection technique such as sequential forward selection (SFS), sequential backward selection (SBS)and Relief based selection algorithm (Relief-F) [273]. These 25 features were comprised of time domain, frequency domain and time-frequency domain features. Upon close analysis, the classifiers performed relatively better for coffee leaf compared to pea shoot and baby spinach for set parameters.

Moreover, it also showed that the precise estimation of WC presence in coffee leaf from day 1

Table 4.4: Classification Results for Coffee Leaf

Classification Accuracy (%)	Time domain features (11)	Frequency domain features (10)	Time-frequency domain features (4)
SVM	92.6%	93.0%	91.6%
KNN	90.0%	91.8%	89.4%
Decision Tree	91.2%	90.7%	91.2%

to day 4 had been substantially improved compared to other leaves. Since the content of water is vital indicator for explaining the plants overall vitality and growth processes, therefore, timely detection of any deficiency in WC plays a signification role in monitoring the health status of leaves effectively. After the individual performance of three features domain, another attempt was made to assess the performance of the classifier for hybrid combinations of all three domain features collectively. Table 4.7 displayed the classification accuracy of all three classifiers for all three leaves. In this condition, classifiers were trained and cross- validated by applying k=10 folds, and the performance of all three classifiers was obtained. These classifiers, including SVM with RBF kernel, KNN with k=5 and D-Tree, were trained and cross-validated by applying k=10 folds. By comparing the results of hybrid combinations with individual classification performance, it was discovered that the combination of features produced an improvement in classification accuracy for all three leaves. Previously, individual classification only enhanced the coffee leaf, whereas the combination of all three domain collectively improved the performance for other leaves, including pea shoot and baby spinach.

Table 4.5: Classification Results for Pea shoot Leaf

Classification Accuracy (%)	Time domain features (11)	Frequency domain features (10)	Time-frequency domain features (4)
SVM	86.6%	79.2%	80.6%
KNN	79.0%	78.8%	81.4%
Decision Tree	81.2%	81.7%	82.2%

4.4.4 Optimization and Feature Selection

In this work, the aim was to remove any redundant or irrelevant features through the feature selection technique to enhance the classification performance by lessening the computational cost for deployment. The methods for feature selection contain filtering methods which were based on the evaluation of the relevance of features, and other wrapper methods were based on a strong search of a different set of features [272]. We considered three feature selection

Table 4.6: Classification Results for Baby Spinach Leaf

Classification Accuracy (%)	Time domain features (11)	Frequency domain features (10)	Time-frequency domain features (4)
SVM	82.6%	81.1%	84.6%
KNN	81.0%	78.8%	81.4%
Decision Tree	78.2%	79.7%	82.2%

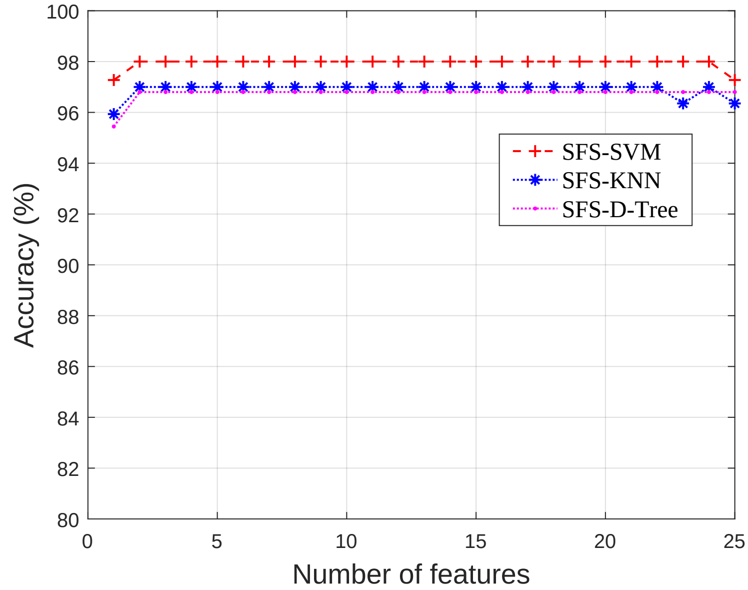
Table 4.7: Classification Results of Hybrid Combination Features for all leaves

Classification Accuracy of three Leaves	SVM	KNN	D-Tree
Coffee	94.46%	93.76%	91.15%
Pea shoot	93.42%	91.62%	90.64%
Baby spinach	91.13%	90.38%	89.01%

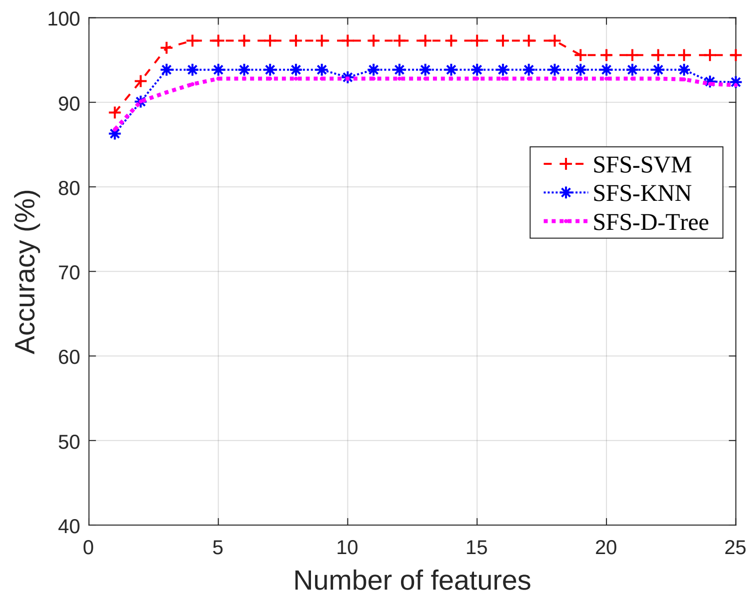
algorithms named as sequential forward selection (SFS), sequential backward selection (SBS) and Relief based selection algorithm (Relief-F) to execute the feature selection process [273]. Out of these three algorithms, SFS and SBS were considered the two most empirical selection algorithms [273]. SFS begins with an empty set and integrates the most suitable feature in every step, and exhibiting a high accuracy by employing a classifier until the pre-defined features are tallied up [273].

On the contrary, SBS operates opposite to the SFS and begins with full occupied features and disposed of unmatched features in every step by specific criterion function till the pre-defined features are [274]. Intriguingly, Relief-F can propose a more efficient technique compared to SFS and SBS and comprehend the relations of features to compute the weights of the features for accurate ranking and selection irrespective of any dependency on specific classifiers [275]. Figure 4.6 depicted the performance of SFS features selection for coffee, pea-shoot and baby spinach leaves using three classifiers.

From Figure 4.6, it was noticed that SVM performed considerably better for all leaves compared to other classifiers using different selection techniques. In addition, Table 4.8,4.9, and 4.10 displayed the classification accuracies for coffee, pea shoot and baby spinach leaves, respectively, using various features selection techniques with the required number of features. By applying a features selection algorithm to classifiers, they produced an improvement of 4%, 3% and 6% for coffee, pea-shoot, and baby spinach leaves using SVM classifiers through SFS technique. The performance of KNN for coffee, pea-shoot, and baby spinach leaves also presented progress in results by 3%, 4%, and 5% correspondingly. These tables indicated the different combinations of features including frequency, time-domain, and time-frequency domain features for classification accuracy.

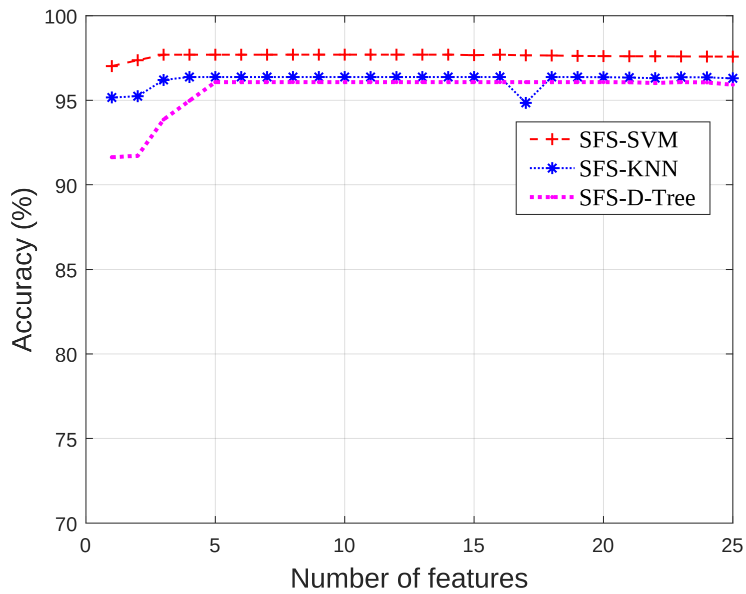


(a) Coffea



(b) Peashoot

As explained in the previous section, it was aimed at reducing the computational time using feature selection techniques. So, in this study, Table 4.11 presented the overall execution time taken by three classifiers for generating results using various feature selection techniques. It was established that execution time taken by classifiers for selected features by performing 10-fold, cross-validation showed considerable enhancement compared to extract features. For example, coffee leaf exhibited an improvement of 15%, 11.9% and 6.5% in computation time for SVM, KNN and D-Tree, respectively, using SFS technique. For pea-shoot, an upgrade of 21.28%, 10.01%, and 8.53% was noticed in operating time for SVM, KNN and D-Tree classifiers, respectively. Lastly, in baby spinach leaf, considering SFS technique, SVM showed an upgrade



(c) Baby-Spinach

Figure 4.6: Transmission response of coffee, pea-shoot, and spinach leaves observed on four different days in the frequency range of 0.75 to 1.1 THz.

of 21.28% in SVM, 10.01% in KNN, and 8.53% in D-Tree operating times. These outcomes indicated that selecting the most relevant and vital features not only enhanced the overall operation time for classifiers but also improved the classification as confirmed with Table 4.8, 4.9, and 4.10. Hence, Table 4.11 is significant for finding the performance of classifiers with less computation time for execution of classification accuracy. In this work, the core purpose was not only to achieve less computation time but also to select relevant features with maximum information using various feature selection techniques. In addition, it could utilize less time and produce maximum accuracy for estimation of WC in plants leaves to maintain a healthy physiological status.

4.4.5 Discussions

In this section, the performance of three proposed classifiers were assessed by employing two commonly quality metrics such as sensitivity or recall (also known as true-positive rate) and specificity (also called false-positive rate) [265, 276]. Here, sensitivity values indicated the possibility of correct identification of labelled class from the remaining target classes [265]. In contrast, specificity showed the probability of appropriate classification as non-target classes from the remaining un-aimed classes [276]. The purpose of utilizing these two widely accepted metrices [265, 276] was mainly to detect any misclassification that could occur, leading to inaccurate information about WC in leaves for four consecutive days.

Table 4.8: Classification Performance for coffee leaf by Applying Tenfold Validation Using Proposed Algorithm with Selected Features

Feature Selection Methods	Classifiers	Serial No. of Features	Total No of Features	Accuracy (%)
SFS	SVM	24	1-19, 21-25	98.5
	KNN	22	1-6, 8-11, 13-	97.2
	D-Tree	24	21, 23-35	96.5
SBS	SVM	24	1-19, 21-25	96.5
	KNN	24	1-21, 23-35	98.6
	D-Tree	24	1-23, 25	97.6
Relief-F	SVM	10	2,4,10,11,17 -21,25	97.1
	KNN			95.9
	D-Tree			96.8

Table 4.9: Classification Performance for Pea-shoot leaf by Applying Tenfold Validation Using Proposed Algorithm with Selected Features

Feature Selection Methods	Classifiers	Serial No. of Selected Features	No of Selected features	Accuracy (%)
SFS	SVM	18	1-6, 8-14, 17, 19, 20, 22, 25	97.2
	KNN	13	1-5, 9-11, 18-20, 23, 25	94.4
	D-Tree	7	2, 4, 5, 6, 17, 18, 19	93.1
SBS	SVM	3	13, 19, 22	96.8
	KNN	5	7, 12, 17, 19, 20	94.9
	D-Tree	2	8, 20	92.3
Relief-F	SVM	12	2,4,10,11,17-21,23-25	98.6
	KNN			99.1
	D-Tree			95.5

Table 4.10: Classification Performance for Baby-Spinach leaf by Applying Tenfold Validation Using Proposed Algorithm with Selected Features

Feature Selection Methods	Classifiers	Serial Num. of Features	Total No of features	Accuracy (%)
SFS	SVM	24	1-12, 14-25	97.9
	KNN	23	1-14, 17-25	96.4
	D-Tree	5	3,5,17,20,21	96.1
SBS	SVM	23	1-11, 13, 15-25	96.8
	KNN	24	1-13, 15-25	94.5
	D-Tree	5	7,8,9,11,15	93.2
Relief-F	SVM	17	2,4-7,10,11,15-21,23-25	98.6
	KNN			99.1
	D-Tree			95.5

Table 4.11: Classification Performance of all classifiers by Applying Tenfold Validation Using Proposed Algorithms with Selected Features

Feature types and feature selection methods	Computation time (s)		
	SVM	KNN	D-Tree
Coffee leaf:			
Extracted Features	0.7282	0.5309	0.4021
Selected Features:			
SFS	0.5706	0.4123	0.3371
SBS	0.6456	0.4240	0.3202
Relief-F	0.6252	0.4842	0.3582
Baby Spinach leaf:			
Extracted Features	0.8975	0.4265	0.4053
Selected Features:			
SFS	0.6062	0.4128	0.1071
SBS	0.4259	0.3576	0.3247
Relief-F	0.4485	0.3875	0.3490
Peashoot leaf:			
Extracted Features	0.6825	0.4405	0.4196
Selected Features:			
SFS	0.4699	0.3404	0.3343
SBS	0.6504	0.1734	0.3149
Relief-F	0.5088	0.3766	0.3759

Table 4.12: Classification Performance of all classifiers by Applying Leave-One-Observation-Cross-Validation Techniques with Selected Features

Quality metrics	Water Content (%)	SVM	KNN	D-Tree
Coffee leaf Day 1	82.84			
SENS		1	1	1
SPEC			1	1
Day 2	41.22			
SENS		1	0.929	0.976
SPEC		0.988	0.965	1
Day 3	12.34			
SENS		0.963	0.889	1
SPEC		1	0.912	0.99
Day 4	0.51			
SENS		1	1	1
SPEC		1	1	1
Peashoot Day 1	76.84			
SENS		1	1	1
SPEC		1	1	1
Day 2	49.22			
SENS		1	0.892	1
SPEC		0.962	0.982	0.971
Day 3	18.91			
SENS		0.545	0.727	0.636
SPEC		0.984	0.967	0.984
Day 4	0.21			
SENS		0.919	0.85	0.833
SPEC		0.987	0.85	0.933
Spinach Day 1	71.14			
SENS		0.995	1	1
SPEC		1	1	1
Day2	34.22			
SENS		1	1	1
SPEC		0.976	1	1
Day3	10.34			
SENS		0.909	0.545	0.851
SPEC		0.923	0.949	0.897
Day4	0.10			
SENS		0.727	0.818	0.636
SPEC		0.974	0.872	0.949

These two-quality metrices depicted the performance of classifiers ranging values from 0 to 1 on days 1 to 4, indicating the presence of WC in all three leaves. Table 4.12 illustrated the performance of all classifiers using a feature selection method and showed the WC presence in all three leaves from day 1 to 4. From Table 4.12, it was also perceived that SVM outperformed

Table 4.13: The Confusion Accuracy with Leave-One-Observations-Out Cross-Validation Method of all Leaves for each day along with monitoring the water content values for each day

Samples	Classes	Classifiers Test Accuracy Performance (%)			Water Content (%)
		SVM	KNN	D-Tree	
Coffee Leaf	Day1	100	100	100	82.84
	Day2	95.2	88.1	100	41.22
	Day3	100	92.6	92.3	12.34
	Day4	100	100	100	0.71
	Variance	0.58	1.09	0.92	
Peashoot Leaf	Day1	100	100	100	76.84
	Day2	100	87.5	87.5	49.22
	Day3	93.6	78.4	74.2	18.91
	Day4	95.0	89.3	91.7	0.21
	Variance	1.55	2.27	3.60	
Baby Spinach Leaf	Day1	100	100	100	71.14
	Day2	100	100	100	34.22
	Day3	92.6	88.6	75.5	10.34
	Day4	94.7	89.7	91.3	0.10
	Variance	1.76	2.90	4.60	

other classifiers for a coffee leaf on different days. Moreover, the assessment of quality metrics for a coffee leaf on days 1 and 4 performed noticeably better, revealing the freshness and staleness of leaf. These results also discovered that the presence of WC on day 1 was high and low on day 4, which helped the classifier to execute the improved performance. Furthermore, it was worth noting that the classification accuracy for all leaves on days 2 and 3 was slightly challenging when the presence of WC in leaves was found in the range of 20% to 50% approximately.

Considering the real-life scenario, the proposed methodology can be substantial by observing the performance of the classifiers for leave-one-observation-out cross-validation method to achieve different days classifications accuracy and for accurate estimation of WC in leaves. This proposed method evaluated the actual performance of the classifier model by randomly selecting each observation from the dataset considered as a validation set, while the remaining observations were taken as the training set. This process continued until all observations from the dataset were nominated for the validation set for at least one attempt. Table 4.12 illustrated the accuracy of the classifications of all leaves for each day by applying the leave-one-observation-out cross-validation technique.

From Table 4.13, it was perceived that SVM classification accuracy outperformed other classifiers for all leaves by showing minimum variance. It also displayed that variability in WC of leaves over the course of four consecutive days. Furthermore, it was also noticed that for both days 1 and 4, classifiers produced maximum accuracy reflecting a high and low WC on days 1 and 4, respectively. Whereas on days 2 and 3, SVM performance stayed in the range from 92.6% to 100%, KNN yielded a range of 78.4% to 100%, and D-Tree produced a range of 74.2% to 100%. Hence, it was concluded that SVM achieved a better classification accuracy range on days 2 and 3 compared to other classifiers because of better variance difference. Thus, the aim of applying leave-one-observation-out cross-validation technique was to evaluate the consistency of classifiers by assessing all observations of different samples on different days as depicted in Table 4.13. It was also strongly aimed to assess the performance of the proposed ML algorithm with the incorporation of THz for real-time applications in monitoring any diminutive variations of WC in plants leaves to help in developing digital agricultural systems.

4.4.6 Summary

In this chapter, a novel ML driven approach was proposed to accurately determine the health status of plants leaves THz waves. In this process, transmission response of leaves was measured for four consecutive days, where each of the 201 frequency points were used as a feature. We performed feature selection to discard any irrelevant and spurious features that could give false observations about the WC in leaves. In this study, results showed that the performance of classifiers was drastically improved by identifying more relevant and important features that could can yield maximum information about WC in leaves. The selection of useful features also reduced the computation time for the execution of classifications by all three classifiers, which was also one of an ultimate objective. Moreover, the comprehensive cross-validation methodology demonstrated that, in most cases, support vector machine SVM yielded highest classification accuracy compared to other classifiers. It was observed that SVM achieved relatively more reliable results for predicting the accurate WC estimation in three leaves for four consecutive days.

This chapter demonstrates the potential and establishes a notable integration of machine learning (ML) using terahertz (THz) waves to assess the real-time information of WC in various plants' leaves. In an era, where most of the farmlands around the globe are water-stressed, the outcomes of this study can help in the design and implementation of smart, sustainable digital agricultural technologies, which is of high importance to boost the overall crops productivity.

Chapter 5

Quality Assessment of Fresh Fruits Using THz Sensing

5.1 Introduction

Fruits and vegetables comprise an essential part of the human diet as they are the primary source of dietary nutrients [277]. In recent years, the rising demand for fruits quality evaluation and sensory characteristics have posed significant challenges in the agriculture sector. Although, manual sorting and grading can be done for the quality assessments and freshness of fruits detection, still this method is significantly fickle, inconsistent, and tedious [277]. In this situation, the identification of any microbially contaminated fruits is quite challenging and may cause severe threats to human health by causing numerous diseases [277]. Therefore, this vital issue strongly demands an intervention of innovative, viable and feasible technological solution that can closely and accurately monitor health status and the MC of fruits to ensure its freshness and quality [278].

5.1.1 Related Work

To address this uncertainty, many significant contributions and techniques on devising non-destructive have been suggested to accurately determine the MC in fruits [278, 279]. These techniques include MRI, NIRS, hyper-spectral imaging, have been investigated and widely deployed to perform the qualitative analysis and composites level of fruits [278].

Regrettably, these techniques have mainly focused on determining the soluble solid content, acidity and physical attributes of fruits. Researchers, scientists and bio-technologists are of a strong interpretation that these aforementioned factors cannot be considered as sole quality parameters for fruits freshness [280]. Also, there are other limitations to these methods, including low resolution and low sensitivity to detect any variations at a cellular level in fruits [281].

[0]This chapter is from paper no 4* in publication list

To overcome these challenges, THz-TDS was introduced to acquire detailed information at a cellular level due to its high absorption ability, sensitivity and resolution as depicted in Figure 5.1. Due to these characteristics, it can identify the early inception of nutrients contamination in fruits [282, 283]. This technique is deemed to have a promising result to detect diminutive variations of MC in fruit slices compared to previous methods. However, it is costly and not portable [284].

5.1.2 Objectives and Contributions

Hence, aforesaid prevailing challenges in monitoring the internal morphology and biological complex traits in fruits at cellular level as shown in Figure 5.1 enthralled researchers from diverse disciplines. Therefore, consciousness of quality control for fruits is seen as a potentially important application area for THz systems provided reliable machine learning (ML) techniques can be integrated with THz sensing equipment. Researchers from multidisciplinary studies emphasize that ML has been effectively applied in various disciplines and due to its high computing performance, it creates novel opportunities to fully comprehend the intensive data processes in numerous fields [282]. Some of the significant contributions achieved by applying ML in various disciplines are food security, meteorology, medicine, economic sciences etc. However, researchers are of a strong view that its potential in the discipline of quality assessment of fruits is still one of the least explored research areas until now [282].

In this chapter, we report a novel, and non-invasive technique to closely monitor and foresee the future trends of MC in fresh fruits slices of apple and mango at molecular level by applying

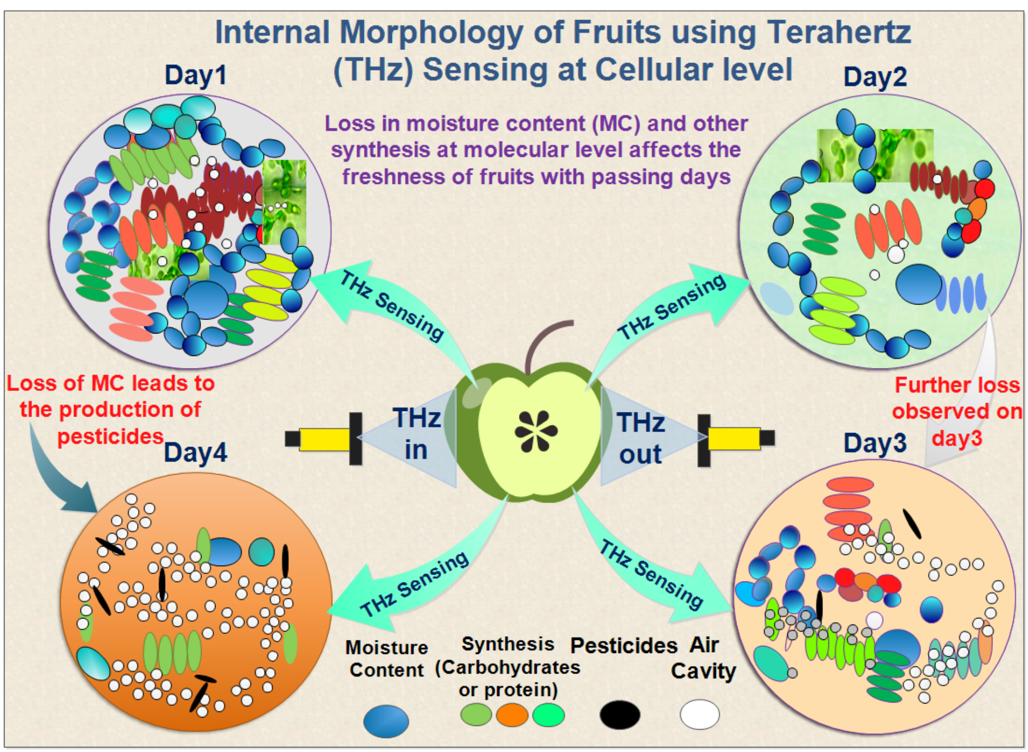


Figure 5.1: Internal Structure of the Fruits for Four Days Using the THz Sensing at Cellular Level.

the THz waves in the frequency range of 0.75 THz to 1.1 THz [285]. For this purpose, we have performed in-lab experiments utilizing the fresh fruits slices of apple and mango and carefully observed the MC in fruits by using scattering parameters of THz waves. The observed data is further pre-processed and put into proposed ML algorithm. Thus, the integration of ML and THz have demonstrated the strong potential of evaluating the freshness of fruits in an automated fashion, which in turn, can help in reducing the health and purification expenses, and optimize economic benefits by maintaining the nutrients level and MC in fruits. The chapter is organized as follows: Section 5.2 presents the data-collection and pre-processing methodology. Section 5.3 describes the feature extraction. This is followed in section 5.4 by the feature selection technique. In section 5.5, performance of three classifiers are discussed. Finally, summary and future work are highlighted in section 5.6

5.2 Data Collection and Pre-processing Methodology

5.2.1 Experimental Method

In this system, we employed a THz swissto12 MCK which was connected to Keysight Technologies VNA N5224A and with extender waveguide WM-250 (WR-1.0), enabling measurements to be performed in the terahertz frequency range of 0.75 THz to 1.1 THz as shown in Figure 5.2. The MCK comprised of two circular waveguides with further two low-loss corrugated waveguide. The movable part of the MCK enabled the material under test (MUT) to accommodate the thickness of 40μm to 4mm [286]. The scattering parameters (S-parameters) including reflection and transmission response (S_{11} and S_{21}) was obtained after performing a fully two-port Short-Open-Load-Thru (SOLT) calibration, aiming to lessen any undesired noise while performing the measurements.

5.2.2 Samples Preparation

In this study, two fresh fruits samples including mango and red apple were bought from Morrisons supermarket in Glasgow. Subsequently, two slices of each fruits were taken as samples with the average thickness between 2.3mm to 3.2mm. These slices were preserved carefully under the ambient temperature of 18 °C$\pm$0.1 °C and humidity of 25%$\pm$2%. The thickness of each slice was measured using a Vernier caliper at three different locations to ensure the evenness all over the whole slice and satisfy the threshold range of MCK.

5.2.3 Data Collection

Before collecting observations, the weight of each slice was measured using a digital scale with an accuracy of 0.1mg to monitor the variation of MC in both apple and mango slices. The weight

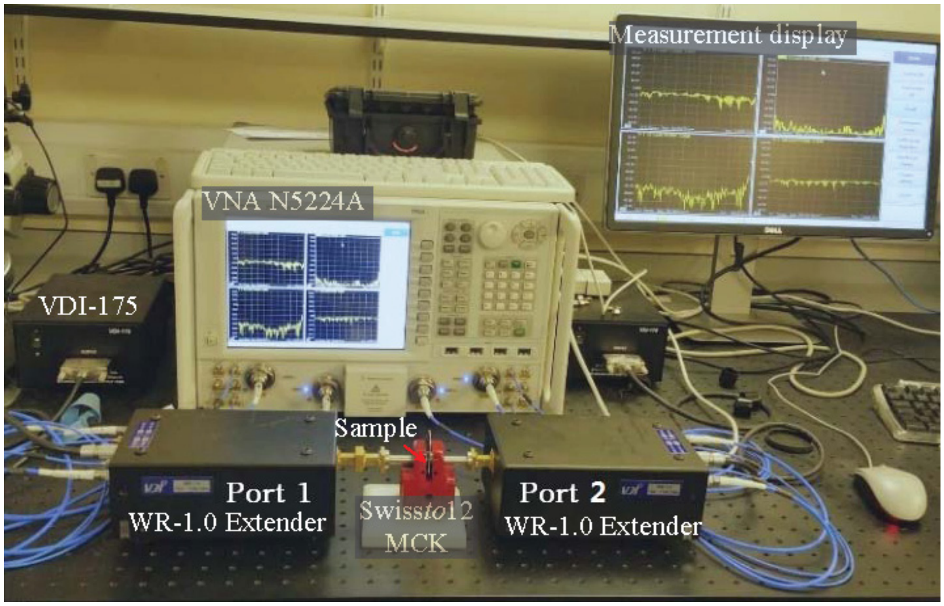


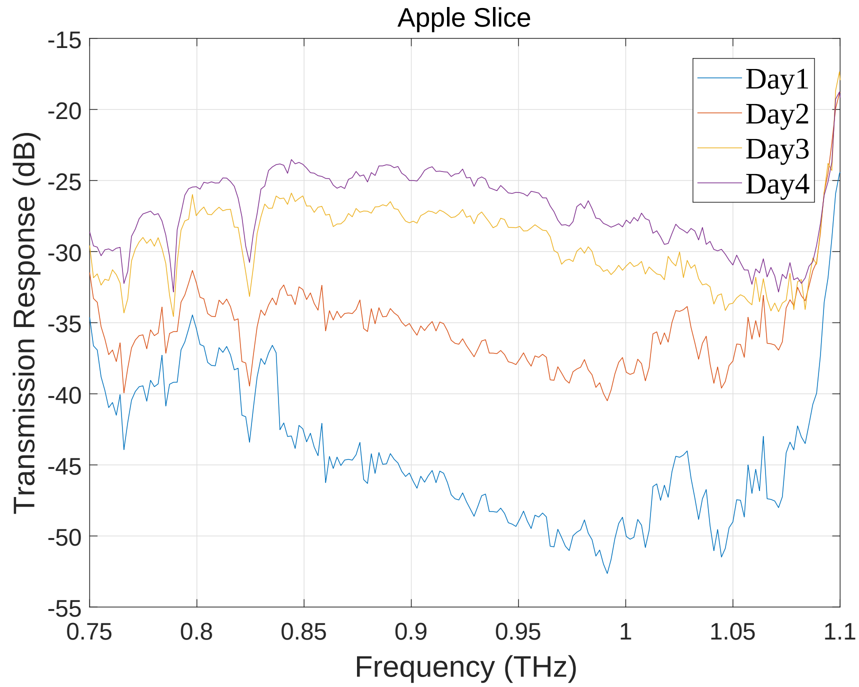
Figure 5.2: Experimental set up of system for measuring the transmission response (S_{21}) of apple and fruit slices.

of each sample was converted into MC value with the passing time as follows [287].

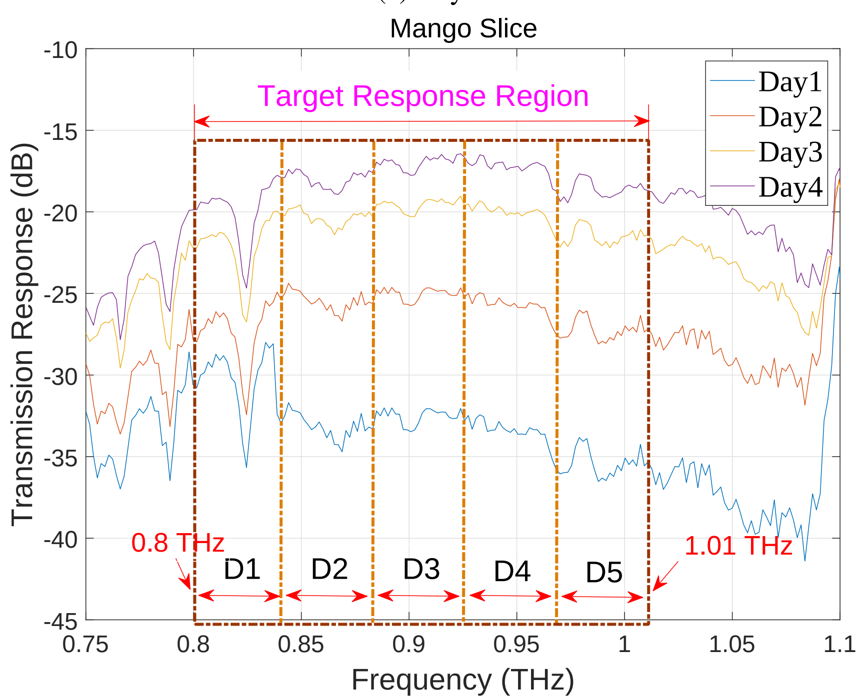
$$MC_{(k)}^{(ij)} = \frac{M_{(k)}^{(ij)} - M_{(k)}^{(dry)}}{M_{(k)}^{(ij)}} \times 100\% \tag{5.1}$$

Where $MC_{(k)}^{(ij)}$ and $M_{(k)}^{(ij)}$ indicate the MC value and the mass of the k-th sample at the j-th measurement in the i-th day, respectively. $M_{(k)}^{(dry)}$ denotes the mass of the k-th sample that dried out to the evaporation of MC over the course of four days. It was aimed to observe the various degree of MC in both samples at maximum location on slices. Hence, both samples were measured at four various locations with four different orientations. At each location, ten readings were recorded, and a total of 160 observations were collected for mango and apple slices on each day, respectively. This whole process was repeated for four consecutive days to monitor the variations of the freshness of fruit slices. At the end, there were 8 target data sets collected from VNA, $S_{(k)}^{(n)}(f_i)$, for two different slices (k=1,2 numbered for apple slice and mango slice, respectively) in four consecutive days (n=1,2,3,4) with frequency f_i (i=0:200) in the range of 0.75 TH to 1.1 THz. Figure 5.3 showed the mean transmission responses (S_{21}) of the same day observations for apple and mango slices in four consecutive days, $S_{(1)}^{(1)}(f)$, $S_{(1)}^{(2)}(f)$, $S_{(1)}^{(3)}(f)$, and $S_{(1)}^{(4)}(f)$, respectively.

In Figure 5.3(a) and 5.3(b), it can be observed that two different fruits slices including apple and mango showed distinctive transmission responses in the THz region from day 1 to 4. The difference in transmission response also revealed the presence of different degree of composites in the fruit, such as proteins, sugar and fats, causing the different absorption of the THz radiation. Upon a close analysis of Figure 5.3, it was noticed that on day 1, transmission



(a) Day 1



(b) Day 4

Figure 5.3: Shows the process of obtaining TRR area to identify useful features for performing the classification process.

response was substantially low due to the presence of the high percentage of MC in fruits.

5.2.4 Data Pre-processing

Prior to any data analysis, it was essential to pre-process the observations to reduce the effects of the any unwanted background noise caused by the THz system hardware. This undesired noise would have produced false observations of MC in samples and hence affected the overall classification accuracy [288]. The simple method was to calculate the mean observations to reduce the random noise. However, the efficient technique was to consider wavelet-based method because decomposition in the wavelet domain could provide better de-noising performance for the ultrafast pulses of THz spectrum [288]. In our work, since observations obtained from VNA were in the frequency domain, so, these observations were converted to the time domain by applying Inverse Fast Fourier Transform (IFFT).

Subsequently, a three-level wavelet decomposition was adopted to pre-process the time-domain signals of all observations with the "heuristic" and soft threshold methods. After de-noising pre-processing, the time-domain signals were arranged into a data matrix with the dimension of M×N as (2) for time domain feature extraction. And the de-noising time-domain signals were performed Fast Fourier Transform (FFT) to obtain the frequency domain response and organized into M×N matrix as (3) for frequency domain feature extraction.

$$s_{(k)}(n) = \left[s_{(k)}^{(1)}(n); s_{(k)}^{(2)}(n); s_{(k)}^{(3)}(n); s_{(k)}^{(4)}(n)\right] \tag{5.2}$$

$$S_{(k)}(f_i) = \left[S_{(k)}^{(1)}(f_i); S_{(k)}^{(2)}(f_i); S_{(k)}^{(3)}(f_i); S_{(k)}^{(4)}(f_i)\right] \tag{5.3}$$

Where k=1,2 indicates apple slice and mango slice, respectively. M is equal to 640 and indicates the number of observations of the slice k. f_i denotes the i-th frequency point. N is equal to 201 and indicates the number of discrete time or frequency points.

In THz experiment system, the observations generated by MCK are raw data, and it can be distorted due to the system's imperfections and transmission losses at both ends of the THz region. So, these redundant and misdetection observations may affect the overall classification accuracy for different classifiers and produce forged information about freshness of slices. Hence, it is vital to identify target response region (TRR) from the overall region, as shown in Figure 5.3(b), to focus on meaningful observations and help to reduce computational loads for optimum classification results. Hence, two-sample t-test with statistical significance difference is performed on all observations of all different days [290]. It is noticed from cumulative distribution function (CDF) of the probability of t-test that the observations in the frequency range from 0.8 THz to 1.05 THz exhibited a significant difference with the value of probability p near to 0 between the different days based on MC of fruit slice. It further indicates that observations are more sensitive in this TRR area.

The block diagram of the proposed classification system for different days based on MC of fruit slices, as shown in Figure 5.4. Since observations obtained from MCK were in the

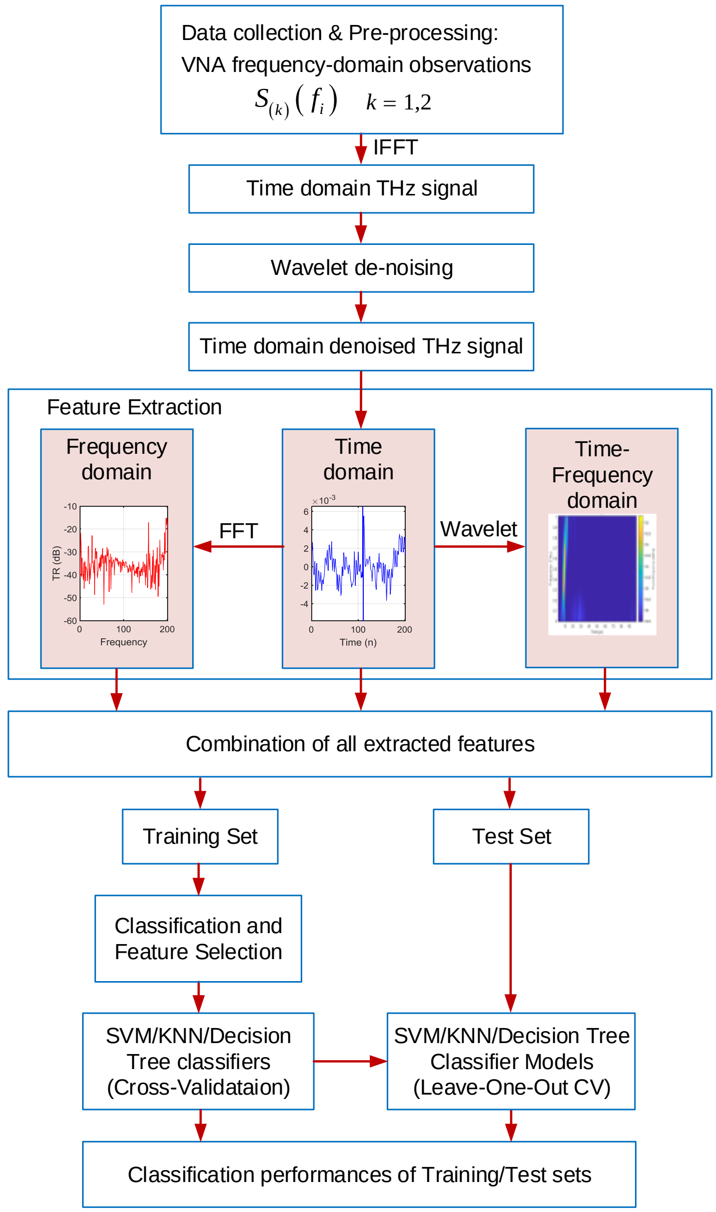


Figure 5.4: The block diagram of proposed classification process.

frequency domain, these observations were converted to the time domain and time-frequency domain features by applying IFFT and Short-Time Fourier transform (STFT), respectively, to obtain feature extraction. The description of three domain features is discussed in detail in next section individually.

5.3 Feature Extraction Technique

5.3.1 Frequency Domain Features

The observations obtained from VNA were in frequency domain form, which can be used to extract the frequency domain features in TRR range directly. For this purpose, the variance of the Power Spectral Density (PSD) and the peak value of Cross Power Spectral Density (CPSD) were considered as given in equation 5.4 and 5.5, respectively [291].

$$\mathrm{Var}\{S_{kk}(\omega)\} = \frac{1}{D} E\left\{\left(S_{(k)}^{(n)}(\omega)^{*} \cdot S_{(k)}^{(n)}(\omega)\right)\right\} \tag{5.4}$$

$$\max\{S_{rk}(\omega)\} = max\frac{1}{D} E\left\{\left(R(\omega)^{*} \cdot S_{(k)}^{(n)}(\omega)\right)\right\} \tag{5.5}$$

In equation 5.4, $S_{(k)}^{(n)}(\omega)$ is the transmission response matrix of the k-th fruit slice in n-th day. $(\cdot)^{*}$ indicates conjugate transpose. D is the width of the frequency windows. ω is the angular frequency, $\omega = 2\pi f$. In equation 5.5, $R(\omega)$ is the transmission response of the reference signal, which is measured by connecting the two halves of the corrugated waveguides of the MCK without sample between them together.

In the TRR range, five frequency windows, $D_1, \ldots, D_5$, with equal width of each frequency window were taken into consideration to extract frequency-domain features for observations of fruits slices. The total area as shown in Figure 5.3 would be from 0.8 THz to 1.01 THz. The selected region was investigated for all the observations from day 1 to day 4.

5.3.2 Time Domain Features

The main purpose of extracting time domain features was to observe the transmission properties of time series of THz pulse referring to presence of MC in fruit slices from day 1 to 4. For waveforms in time domain, statistical features can be derived with an assumption that the signal is stationary [292,293]. In this study eleven time domain features were considered including mean, median, mean of absolute value (MAV), standard deviation (STD), mean of absolute deviation (MAD), skewness and kurtosis, Pearson correlation coefficient (PCC) [294], 25th percentile (Q1), 75th percentile (Q3) , and Interquartile Range (IQR). In which, mean, median, MAV, STD and MAD are common properties used in statistics and probability theory. Skewness and kurtosis are two higher statistics terms, the former provides the symmetric information of data sets relative to the centre point, and the latter provides the flatness of data sets in a uniform distribution. PCC is used to measure the linear relationship between the time-domain waveforms of the sample and reference signal [295, 296]. Q1 and Q3 measure the value at the 25% and 75% location of the data set separately. IQR can be applied to measure the variability of a data set.

Table 5.1: Comparison of Classification Performance of Different days Using Different Datasets

Slices	Classifiers models	Classification Accuracy (%)					Execution time of Hybrid features (s)
		Raw data	Time domain features	Frequency domain features	Time Frequency domain features	Hybrid features	
Apple	SVM	65.2	88.6	97.0	93.6	98.9	1.6506
	KNN	77.1	78.0	86.4	86.4	86.7	0.8274
	D-Tree	72.4	88.2	93.2	93.2	97.8	0.6776
Mango	SVM	70.1	92.0	93.4	93.4	93.4	1.9837
	KNN	79.7	87.7	86.4	86.4	88.6	0.6579
	D-Tree	76.1	91.6	92.5	92.5	93.6	0.8760

5.3.3 Time-Frequency Domain Features

The time-frequency methods such as STFT and wavelet decomposition of the time-domain signal can provide detailed localization properties of THz pulses in time-frequency domain [296]. The wavelet-transform (WT) was more suitable for analyzing the short-duration pulse with fast and unpredictable changes to obtain interesting information [295]. After wavelet de-noising of the observations in time domain, the representation of the time-frequency domain was performed by three level wavelet decomposition using db8 wavelet, and a total of four sub-band signals are decomposed, in which one approximation coefficient, $C_3(n)$, and three detail wavelet coefficients, $D_3(n)$,$D_2(n)$, and $D_1(n)$, can be applied to extract time-frequency domain features [297, 298]. The features of the extracted wavelet coefficients can provide the energy distribution of the THz signal in time and frequency domain.

5.4 Classification Result and Feature Selection

5.4.1 Classification Accuracy for different feature datasets

In this study, three ML classifiers, namely SVM, K-nearest KNN and decision tree D-tree are considered to analyze the performances of various features datasets. These feature datasets are raw data collected from MCK, individual domain feature datasets of time, frequency, time-frequency, and hybrid combination dataset of three domain extracted features. In this regard, for each classifier, suitable parameters are selected to enhance the classification accuracy. So, for SVM classifier, Radial Basis Function (RBF) kernel is applied to set key parameters including the Gaussian kernel scale (γ), and the optimum parameters of cost (C). To establish the appropriate values, series of values are assessed and eventually, 0.35 and 1 are chosen for (γ) and (C), respectively, after the grid searching optimization of parameters $\gamma(0.1 : 0.05 : 2)$ and $\mathrm{C}(0.5 : 0.5 : 2)$ [299]. For KNN classifier, we set the number of nearest neighbours k and the distance metric to 5 after examining the range of k(1:10) and Euclidean distance, respectively [300].

Table 5.2: Classification Performance of Different Days Using selected Features

Slices	Features methods	Classifiers	Number of Selected Features				Acc. (%)	Execution time of selected features (s)
			Time-domain features	Freq. domain features	Time Freq. features	Total no. of features		
Apple	SFS	SVM	9	6	4	19	99.8	1.2561
		KNN	9	5	4	18	88.2	0.5945
		D-Tree	11	8	0	19	99.5	0.2787
	Relief-F	SVM	4	10	3	17	99.1	1.4327
		KNN					98.9	0.5788
		D-Tree					99.8	0.5510
Mango	SFS	SVM	10	8	4	22	95.4	1.3456
		KNN	9	2	4	15	94.1	0.2678
		D-Tree	3	1	0	4	95.2	0.2702
	Relief-F	SVM	6	10	3	19	98.6	1.0870
		KNN					99.1	0.6229
		D-Tree					95.5	0.5861

Lastly, for D-Tree, the maximum number of splits is set to 5. All the other parameters of three classifiers are retained as default values. Furthermore, all classifiers are trained by using 10-fold cross-validation to obtain the validation accuracy of classification of days for each fruit slice, and each type of dataset is partitioned into 80% and 20% training and testing data, respectively.

Table 5.1 depicts the comparison of all classifier's performance using various feature datasets. Upon close analysis, it is noticed that the performance of all three classifiers for raw data is unsatisfactory compared to other domain features performance. It also indicates the presence of less sensitive data or data littered with some random noise, which has severely affected the performance of classifiers and can be substantially improved by observing a sensitive region. Moreover, Table 5.1 also shows an improvement of 17% for individual domain features compared to raw data. For instance, the classification performance of frequency- and time-frequency domains exceed time-domain. Likewise, frequency- and time-frequency domain features present the same level classification performance, which is due to that fact that the THz spectrum in the frequency domain can provide more comprehensive radiation information of samples including transmittance and absorption of THz waves. Furthermore, it is also perceived that the hybrid features of three domain exhibit the best classification performance i.e. 98.9% in SVM classifier for four consecutive days of the apple slice.

By close observations of SVM, KNN and D-Tree performance, it is evidently observed that results could be further improved by feature extraction in TRR range of the raw data. Moreover, using the state-of-the-art technique of feature selection method in [301], features could be combined optimally from multi-domains, and this also helped to reduce the computational time by eliminating less-informative features.

Table 5.3: Confusion Matrix and Accuracy of Classifiers for training and Testing of Apple Slice.

Slices	Classifier Models	Sample Sets	Days	Predicted				Acc. day (%)	Overall Acc.(%)
				Day1	Day2	Day3	Day4		
Apple	SVM	Training	Day1	128	0	0	0	100	99.20
			Day2	0	62	1	1	96.88	
			Day3	0	0	64	0	100	
			Day4	0	0	2	94	97.92	
		Test	Day1	31	1	0	0	96.88	90.91
			Day2	3	12	1	0	75.00	
			Day3	0	0	15	1	93.75	
			Day4	0	0	2	22	91.67	
	KNN	Training	Day1	127	1	0	0	99.22	89.57
			Day2	16	42	5	1	65.63	
			Day3	0	6	51	7	79.69	
			Day4	0	0	7	89	92.71	
		Test	Day1	28	4	0	0	87.50	82.95
			Day2	3	11	0	0	68.75	
			Day3	0	0	13	3	81.25	
			Day4	0	1	2	21	87.50	
	D-Tree	Training	Day1	128	0	0	0	100	99.15
			Day2	0	63	1	0	98.43	
			Day3	0	1	63	0	98.43	
			Day4	0	0	1	95	98.96	
		Test	Day1	32	0	0	0	100	95.45
			Day2	0	15	1	0	93.75	
			Day3	0	0	15	1	93.75	
			Day4	0	0	2	22	91.67	

5.4.2 Features Selection

The redundant or irrelevant features extracted from the previous processing can be removed through feature selection techniques to optimize the combinations of features for improving the classification performance and minimizing the computational cost for deployment. Feature selection techniques include filter methods based on the assessments of the relevance of features and wrapper methods based on the force search of different combinations of the feature space [301]. In this section, we applied two algorithms named sequential forward selection (SFS), and the Relief-based feature selection algorithms (Relief-F) to perform the feature selection [302]. SFS is the heuristic selection algorithm, which starts with an empty set and combines one best feature in each step with high accuracy by using a classifier until the pre-defined number of features are added. Relief-F can offer a computationally efficient method to recognize the interactions of features and calculate the feature weights between 0 and 1 to rank, and select features without depending on the certain classifier [303].

Table 5.2 shows the classification results of different days for apple and mango slices with different feature selection algorithms and the number of selected features. The SFS and Relief-F algorithms yield over 2% improvement for classification of days for apple and mango slices with at least 12% features reduced out of the total of all extracted features. Interestingly, SFS with D-Tree classifier obtains the optimal feature combinations including only three time-domain and one frequency-domain features leading to about 2% improvement of accuracy for mango slice, compared to the results of extracted features in Table . In addition, the execution time obtain in Table 5.2 indicates that selected features have not only enhanced the computational time but have also improved the classification accuracy considerably. The least improvement of the execution time on classification is Relief-F KNN of a mango slice, 5.3%, and the most one is SFS D-Tree of a mango slice which goes up to 69.2%.

5.5 Results and Discussions

Table 5.3 and 5.4 represents the confusion matrix and the classification accuracy for different days with different classifier models, based on features selection using SFS method for both training and test sets of apple and mango slices, respectively. For apple slice, the results illustrate an accuracy increment of 10% for both SVM and D-Tree models compare to KNN model in training set, whereas D-Tree model exhibits highest accuracy in the test set. Furthermore, for mango slice, both SVM and KNN displays better accuracy compare to D-Tree model in training set, and D-Tree model achieves the best accuracy in the test dataset.

The performance of classifiers is assessed by applying two quality classification metrics including sensitivity (SENS) (also called true positive rate), and specificity (SPEC) [295, 304]. The value of SENS represents the probability of the target class identified correctly in the total number of one target classes. Whereas, the value of SPEC expresses the probability of classifying the sample as non-target classes correctly in the non-target classes. The values of the two-quality metrics are found in the range from 0 to 1.

Table 5.5 presents the performance of the proposed classifiers algorithms for different days using SFS method and indicates the presence of MC in both apple and mango slices. It can be observed that the D-Tree has shown substantial improvement in classification accuracy, indicating the freshness or staleness of both different slices. Furthermore, it is also depicted that the assessment parameter values of the SVM classifier achieves better precision compare to KNN for apple and mango slices. From Table 5.5, it is worth observing that the classification of the days is equivalent to that of the percentage of the MC in the slices. It is due to this reason, different MC value leads to the different absorption strength of terahertz radiation in samples, which causes variations in the features for classification. Also, the results show that classification accuracy on days 2 and 3 is slightly more challenging than for day 1 and 4, especially for KNN classifier where the MC values falls in the range between 10% to 50%. It is because the MC value of day 1 is very high (freshness), whereas on day 4 (almost dried out) is very low,

Table 5.4: Confusion Matrix and Accuracy of Classifiers for training and Testing of Mango Slice.

Slices	Classifier Model	Sample Test	Predicted					Acc days (%)	Overall Acc (%)
			Days	Day1	Day2	Day3	Day4		
Mango	SVM	Training	Day1	127	1	0	0	99.22	97.73
			Day2	0	61	3	0	95.31	
			Day3	0	4	92		95.83	
			Day4	0	0	0	64	100	
		Test	Day1	32	0	0	0	100	90.91
			Day2	0	12	4	0	75.00	
			Day3	0	3	21	0	87.50	
			Day4	0	1	0	15	93.75	
	KNN	Training	Day1	128	0	0	0	100	98.30
			Day2	0	61	3	0	95.31	
			Day3	0	2	94	0	97.92	
			Day4	0	0	1	63	98.44	
		Test	Day1	32	0	0	0	100	89.77
			Day2	0	10	6	0	62.5	
			Day3	0	3	21	0	87.5	
			Day4	0	0	0	16	100	
	D-Tree	Training	Day1	128	0	0	0	100	92.61
			Day2	0	57	7	0	89.06	
			Day3	0	18	78	0	81.25	
			Day4	0	0	1	63	98.44	
		Test	Day1	32	0	0	0	100	94.32
			Day2	0	13	3	0	81.25	
			Day3	0	2	22	0	91.67	
			Day4	0	0	0	16	100	

which leads to a clear separation of features in the feature space.

In a real-life application, it would be more convincing to obtain the classifier's performance by considering leave-one-observation-out-cross-validation technique for more accurate estimation of MC in fruits slices on different days. This propose method can be more efficient because in this process, each observation is randomly selected from the dataset as the validation set, while the remaining observations are considered as the training set to generate the classifier model. The process is repeated until all the observations in the dataset are selected as validation set once. The cumulative classification results of days are calculated from leave-one-observation-out-cross-validation approach based on MC values of each day, as shown in Table 5.5. From table 5.5, it is found that the classification accuracy for day 1 and day 4 is nearly 100% which indicates the highest MC and the best freshness in day 1 and the lowest MC and dehydration in day 4. For day 2 and day 3, the classifiers can provide the quantitative reference based on MC value for the quality assessment of the fruit slices. Thus, the aim of applying the proposed technique is to develop the consistency of classifiers by observing all observations of

Table 5.5: Quality Performance Metrices and Classification Accuracy of Test Done with Leave-one-observation-out Cross-Validation for Classification Evaluation of Days for Apple and Mango Slices Related to MC Values.

Samples	Classes	Quality Metrics				MC (%)	Test Accuracy (%)		
			SVM	KNN	D-Tree		SVM	KNN	D-Tree
Apple	Day1	SENS	1	0.96	1	82.2	100	99.8	100
		SPEC	1	0.93	1				
	Day2	SENS	0.99	0.68	1	37.8	95.8	80.0	100
		SPEC	1	0.98	1				
	Day3	SENS	1	0.86	1	12.3	100	95.0	98.8
		SPEC	1	0.96	1				
	Day4	SENS	1	0.93	1	0.50	100	96.7	100
		SPEC	1	0.98	1				
Mango	Day1	SENS	1	1	1	77.9	100	100	100
		SPEC	1	1	1				
	Day2	SENS	0.80	0.71	0.86	48.7	89.5	81.3	98.8
		SPEC	0.98	0.96	0.97				
	Day3	SENS	0.90	0.84	0.92	15.0	85.7	80.8	93.3
		SPEC	0.94	0.93	0.97				
	Day4	SENS	0.96	0.95	1	0.24	100	99.3	100

both slices on different days.

5.6 Summary

This chapter presented the novel and non-invasive sensing technique for the quality assessment of fresh of fruits by integrating ML with THz waves. For this purpose, scattering measurements of both apple and mango slices were obtained using Swissto12 MCK system. Multiple domain features such as time-, frequency-, and time-frequency domains were extracted and the classification was performed to determine the MC in fruits slices using SVM, KNN and D-tree classifiers. Results showed that by discarding unwanted features and applying comprehensive cross-validation technique, classification accuracy was substantially improved which eventually helped to reduce the computational time. Furthermore, it was perceived that, in most of the cases, the SVM classifier based on RBF kernel outperformed KNN and D-Tree classifier for both fruit slices. Thus, SVM demonstrated more precise quality assessments of fresh fruits by determining their MC more precisely for four consecutive days.

The proposed technique demonstrates the strong potential in the discipline of food and science technology by integrating ML with THz waves to assess real-time information of fruits on different days at cellular level. For future research, more fruits slices will be considered to extract additional features by employing their electromagnetic parameters for the identification of different fruits in an automated and non-invasive manner.

Chapter 6

THz Detection of Impurities in Aqueous Solutions Using ML

6.1 Introduction

In a rapidly developing and modern world, the importance and preservation of clean water without any harmful impurities for the overall global health, environmental protection, and economic development cannot be undervalued [305]. Providing sufficient and affordable water in a safe and reliable way with limited resources is a huge challenge of mounting sternness as the demand increases with a rising population [305, 306]. Also, fresh and unpolluted water is worsened by climate transformations, more regular droughts in many parts of the word, and by water pollution, making it more demanding and costly to handle [305, 306]. Mostly, the general consensus among the scientific community is that the emergence of infectious diseases such as tuberculosis, measles, and other lethal illness, often detected and caused by microbiological and microchemical contaminants in tap and drinking water sources, which cannot be detected by naked eyes, leading to jeopardize the public health and safety [305–308].

6.1.1 Related Work

Hence, the aforementioned facts underline the imperative demand for the fast, reliable, secure, and sensitive technological solutions and vitalization of resources for the precise detection and regular monitoring of water contaminants in vulnerable population to discourage any short and long-term health consequence [305–308]. In recent times, many researchers have put tremendous efforts and suggested variety of directions [306] and technologies [308] to address the substantial issues as discussed earlier. In this regard, significant achievements have been obtained through the initiation of water quality sensors, micro-fluids sensors, model-based event detection, advanced vibrational spectroscopy and miniaturized biosensors, and widely improved

[0]This chapter is from paper no 5* in publication list

the water contamination detection qualitatively [307, 308]. However, these different approaches have some advantages and limitations. For example, the deployment of numerous water quality sensors seems not very feasible due to high installation cost, time consuming, low response detection and less reliability [305–308].

In addition, results from model-based event detection have indicated certain error rate due to the low sensitivity, providing inadequate symptoms of contaminates in water [305]. Some researchers have also considered Infrared (IR) for the swift detection of impurities in solvents [309, 310]. Though it has obtained considerable advancements and yield satisfactory results [310]. However, there are some limitations and have mainly focused on the theoretical calculations to observe the characteristics of impurities added in solvents and absorption features [310]. Thus, this technique is transpired as inappropriate and feasible for precise detection of contaminants in pure water at molecular level and have markedly minimize its suitability [310].

Despite in-depth theoretical attempts and substantial significant advancements over the past years, the microscopic frameworks leading to the numerous anomalies or contaminants of water, often considered as the compact substance or a primary biological solvent, remain from being fully comprehended by the researchers in physical and biological sciences [307–310]. Consequently, the concerning effects of poor contamination technique instantly require developing a more robust, qualitative, less operating costs, and high sensitivity quantification of contaminants in solvents in a non-invasive manner [308, 310].

With this motivation and limitations found in previous techniques, this chapter proposes a realistic method and application of FTIR enabled by ML [311] that can provide the approximate prediction and detection of even the smallest of contaminants in distilled water due to high sensitivity and non-destructive nature and can also produce high optical throughput [310]. This technology includes THz, which has achieved tremendous achievements in diverse field such as diagnostic applications of dental and skincare medical imaging, invisible hazard and vulnerable items, material characterizations, and telecommunications [125, 312–314]. For this purpose, an integration of ML with THz can create a dynamic opportunity to uncover, measure, and thoroughly understand the data-intensive procedure in to minutely observe the absorbance spectra of different solutions in aqueous solutions [313]. The significant contributions of this work are as follows:

1. This chapter suggests a novel technique by employing a FTIR setup that provides a THz frequency range of interest operating from 1 THz to 20 THz to precisely determine the various solvents constituents' characteristics in aqueous solutions non-invasively.

2. The proposed methodology also suggests the ML driven approach to proactively determine the presence of any anomalies or impurities in aqueous solutions in real time to protect the environment, include early alerts to protect the public health, and reduce any superfluous costs.

3. In this study, by integrating THz with ML, we explore not only identifying the various constituents' in aqueous solutions, but also to determine the amount of impurities in each constituent solution by establishing ML algorithm technique.

4. Finally, this paper presents a notable and distinctive contributions of THz technology with ML in assessing the impurities in aqueous solutions at cellular level.

The rest of the chapter is organized as follows: Section 6.1.2 presents an experimental setup that includes a sample preparation of various solvents in aqueous solutions. Section 6.2 explores a data collection procedure and measurement results. Section 6.2.1 describes the feature extraction technique of the obtained data. Section 6.2.2 illustrates the various classification algorithm methodologies to predict the proactive detection of anomalies in water quality. This is followed by the explanation of feature selection technique 6.2.2 and 6.2.2 classifiers performance evaluations section. Section 6.3 discuss the results analysis of regression and classification algorithms, in terms of error rates and classification accuracy. Finally, section 6.4 describes the summary towards the end.

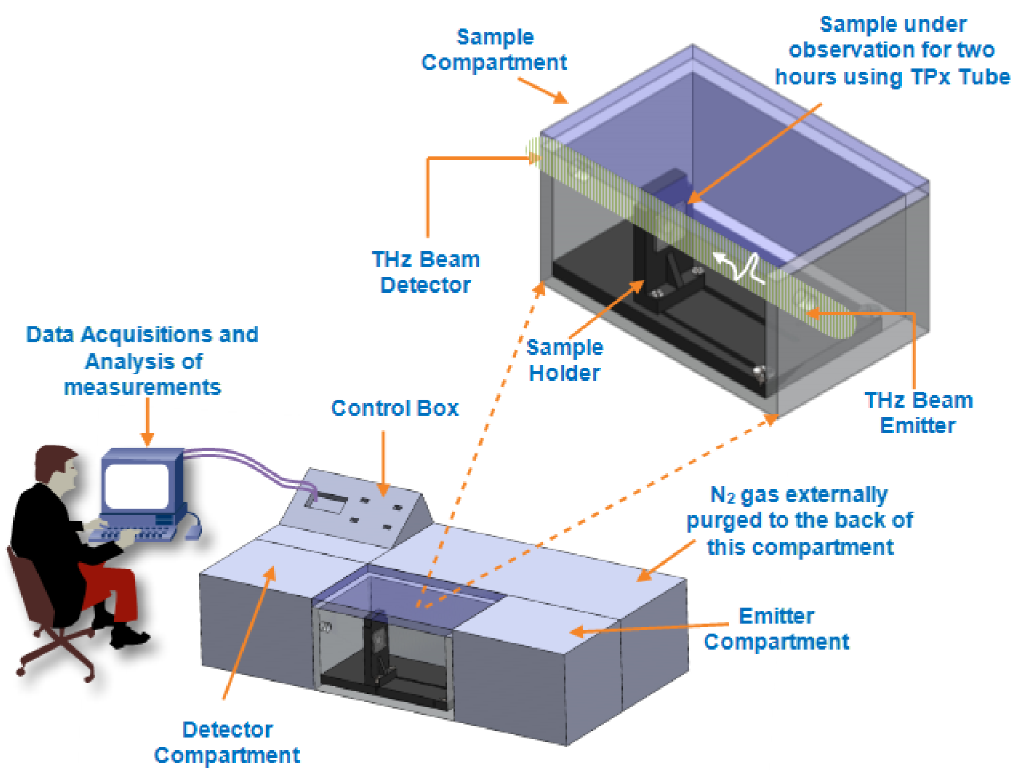


Figure 6.1: Isometric schematic of Fourier Transform Infrared Spectroscopy (FTIR) system, with the sample compartment above pointing the THz beam generated at the source, passing through the TPx tube to assess the constituents of salt, sugar and glucose in aqueous solutions. The observations for each sample ran unto 2 hours approximately.

6.1.2 Methods and Materials

Method

In this experimental setup, a Bruker 66 V/S series FTIR system was employed to accomplish the measurement of various constituent's in aqueous solution as shown in 6.1 [328, 329]. FTIR is a powerful analytical technique, providing a label-free, non-destructive method to show a rapid behaviour for any redundant impurities in solutions. The system was equipped with a DLaTGS/Polyethylene (PE) detector and a 6-micron Mylar beam splitter was employed to perform the measurements [328, 329]. The spectral distribution of the beam-splitter and detector was 1–21 THz, and 0.3–21 THz, respectively [328]. To prevent any formation of the vacuum in sample compartment, and absorption of THz power, Nitrogen (N2) gas was externally purged into the sample compartment. However, some variations in different specimens was noticed and reasons is explained in further section. The flow rate of N2 was set to 600 L/hr as specified in FTIR system guidelines [11]. Owing to high transparency and zero losses in the THz spectrum, a distinct device Polymethylpentene (PMP) commonly known as TPX tube was employed for testing the various specimens. Considerable attention and precautionary measures were taken to ensure that the sampling device is positioned in the same location every time in sample compartment to reduce the distortion in measurements [303].

Sample Details

In this study, three various specimens were considered for measurements including table salt, pure sugar, and glucose. Salt and sugar were bought from Holland Barrett glucose was ordered from Sigma-Aldrich, respectively. The list of all samples with different weight and concentrations is shown in Table 6.1. To prepare a solution of every solvent, a 50ml of distilled water was taken and mixed with different concentrations of salt, sugar and glucose such as 5%, 10%, 20% and 30% using 6.1 [330, 331].

$$Conc = \frac{gram}{gram + 50ml} \times 100 \tag{6.1}$$

These solutions were prepared at room temperature set to 23°C by adding 2.63g, 5.5g, 12.5g and 21.4g to 5%, 20%, 30% and 30%, respectively. The weights of all specimens were carefully calculated using an electronic scale with at least count of 0.1mg. Before placing mixture into the TPX tube, solutions were properly stirred for 3 to 5 minutes approximately to ensure they are being fully dissolved in distilled water. While filling the TPX tube filled with all solutions, great attention was given so that it should be filled up to 11.6ml just in lined to beam-splitter to obtain the maximum and accurate information. All the measurements were performed at an atmospheric temperature of 23 °C.

Table 6.1: Different samples with weights and concentrations.

Sample	Weight in Grams	Concentrations (%)
Table Salt (NaCl)	$2.63g \pm 0.1$	$5\% \pm 0.1$
	$5.56g \pm 0.1$	$10\% \pm 0.1$
	$12.5g \pm 0.1$	$20\% \pm 0.1$
	$21.4g \pm 0.1$	$30\% \pm 0.1$
White Granulated Sugar	$2.63g \pm 0.1$	$5\% \pm 0.1$
	$5.56g \pm 0.1$	$10\% \pm 0.1$
	$12.5g \pm 0.1$	$20\% \pm 0.1$
	$21.4g \pm 0.1$	$30\% \pm 0.1$
Glucose	$2.63g \pm 0.1$	$5\% \pm 0.1$
	$5.56g \pm 0.1$	$10\% \pm 0.1$
	$12.5g \pm 0.1$	$20\% \pm 0.1$
	$21.4g \pm 0.1$	$30\% \pm 0.1$

6.2 Data Acquisition Procedure

In this work, the focus was mainly to observe the THz absorption spectra (AS) for three various distinct solutions as explained earlier. These measurements were performed in Terahertz Laboratory, at University of Glasgow with great care. The time taken for observing AS was 2 hours in order to obtain maximum point data to minutely observe in the THz region for any impurities added in distilled water. The number of data points calculated as 338 was collected for every sample. During the measurement process which lasted for 120 minutes for every sample, and 2102 scans were obtained for each sample. The whole process was repeated for all three samples and data was pre-processed using Matlab 2019a, whereas python was used for ML classification in the form supervised learning. While performing the measurements, it was very important to monitor the (N2) gas on regular basis to ensure the continuous flowrate to the compartment, avoiding any irregular behaviour of constituents added in aqueous solutions. From Figure 6.2, we have attempted to minutely observe the absorption spectra of molecular-scale dynamics of distinguished constituents' concentrations in aqueous solutions. In particular Figure 6.2, it describes the characteristics absorption peaks of different salts concentrations falling in the range of 13 THz to 15 THz range markedly indicating a more sensitive frequency region for precise detection of distinguished concentrations pattern reaching to 0.25, 0.17, 0.09, and 0.06 approximately for 5%, 10%, 20% and 30% concentrations, respectively.

However, considering the range from 15 THz to 17 THz, it can be observed though it has distinct response for distinguished concentrations, but the peaks of absorption spectra for 5%, 10%, 20% and 30% are substantially lower than aforementioned range. This occurrence is attributed to the high sensitivity and strong penetration feature of THz that has depicted diminutive variations of salt concentrations in aqueous solutions at different region. Upon a close analysis of Figure 6.3 and 6.4, it is also depicted that both glucose and sugar have exhibited a distinct response for various concentrations in different THz region. Notably, sugar concentrations display

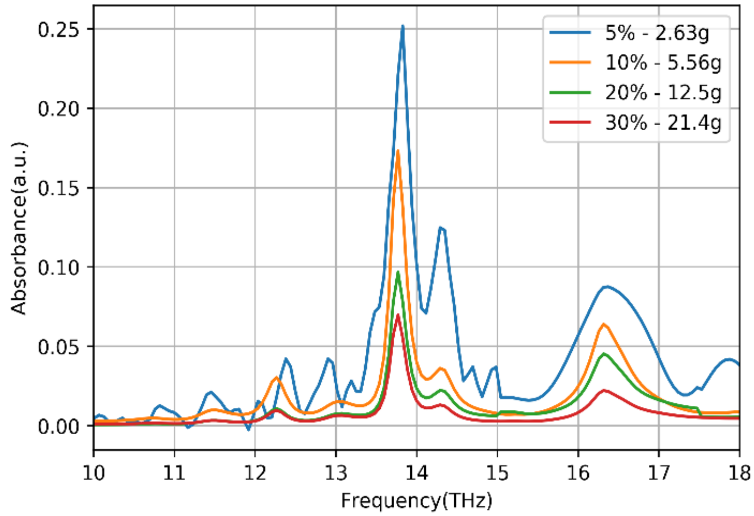


Figure 6.2: the terahertz (THz) absorption spectra of salt constituents' characteristics in aqueous solutions in the frequency range of 0.8 THz to 18 THz

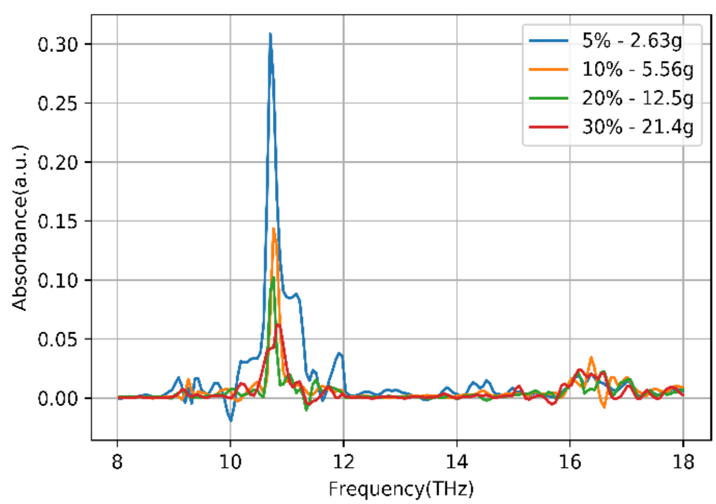


Figure 6.3: Terahertz (THz) absorption spectra of glucose constituents characteristics in aqueous solutions in the frequency range of 0.8 THz to 18 THz

a more discernible response compare to glucose and this is clearly discovered by THz waves. Furthermore, this prominent and distinguished results showed a distinctive characteristics and functional properties of both sugar and glucose concentrations in aqueous solutions since sugar is mainly compound of various synthesis whereas, glucose is considered to be pure. These results reveal the significant influence of added ingredients in the aqueous solutions and interestingly, and also provide a promising method of the rapid and effective identification of elements for the various concentrations. However, the main objective of this study is also to establish a computationally competent and reliable method for estimating water quality variables using THz waves that reduces labour and the cost of accurately measuring these labour these various parameters.

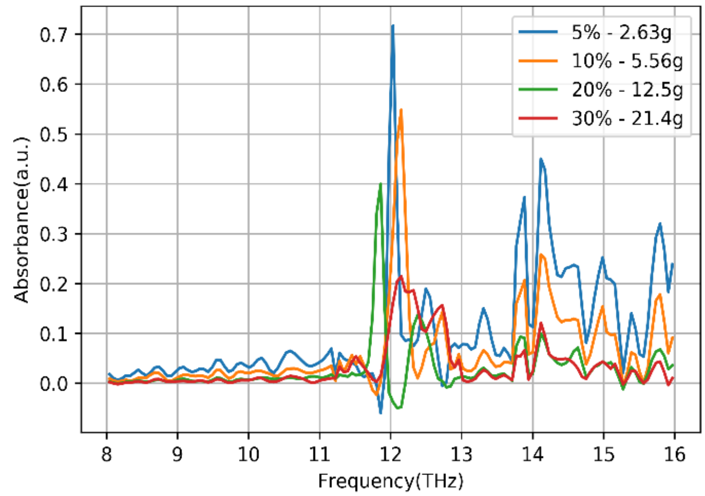


Figure 6.4: Terahertz (THz) absorption spectra of sugar constituents characteristics in aqueous solutions in the frequency range of 0.8 THz to 18 THz

For this purpose, ML algorithm [311] has been developed to identify any unknown irregularities or anomalies in pure distilled water. In addition, it is also aimed to detect the exact amount of concentrations of mysterious impurities added to the distilled water. Thus, an effective, automated, and precise quantitative detection of harmful contaminants at molecular level in water is utmost of significance to provide early warnings to protect public health.

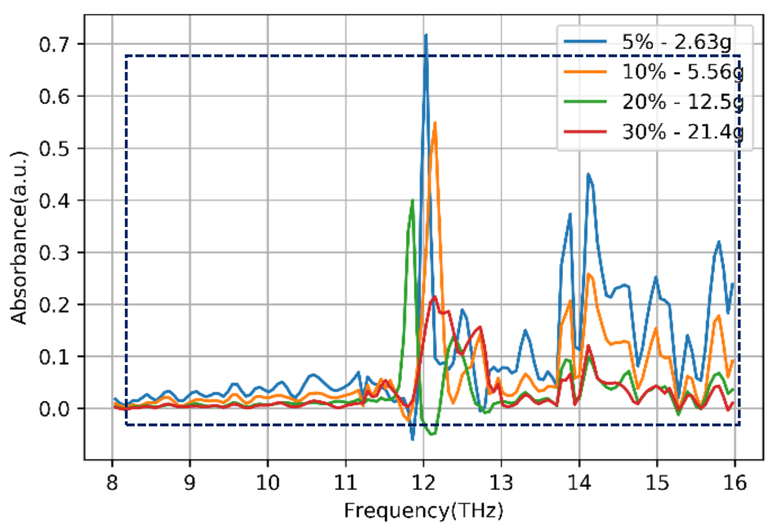


Figure 6.5: Identification of target response region (TRR) for the feature extraction process.

6.2.1 Feature Extraction Method

While taking the measurement, it was noticed that observations collected using FTIR setup were appeared to be little irregular and unwanted excessive variations. The occurrence of this notice-

Table 6.2: Significant Feature Extraction Techniques Using Time and Frequency Domain Features.

Time-Domain Features		Frequency Domain Features	
Minimum	2	Special Entropy	3
Maximum	2	Spectral Power	9
Mean	2		
Standard Deviation	2		
Skewness	2		
Kurtosis	2		
25^{th} percentile (Q1)	2		
75^{th} percentile (Q 3)	2		
Range	2		
Root Mean Square	2		
Quartile	2		
Total features	22	Total features	12

ably undesired and counterfeit observations may have given the fictitious information about any impurities or concentrations level added into the aqueous solutions. Furthermore, useful observations would also have a fruitful impact on overall classification outcome. In addition, presence of any imprecise minerals or chemicals in water can be highly harmful for overall public health. In this regard, it was significant to discover the sensitive SFR as shown in Figure 6.5, in the THz region with minimum intrusion of any external factors in the observations, contributing to obtain the maximum information about the smallest particles of constituents in water. For this purpose, specific region (SR) was established for each constituents' samples ranging from 8 THz to 17 THz out of whole region. Within this specific region, absorption spectra peaks for low and high concentrations can be easily discerned with little overlap [315].

Since the observations collected from setup was in frequency domain, so, it was converted to time-domain region using an IFFT to initiate the possibility of acquiring the statistical features of observations. Out of 338 features observations, 34 valuable features were extracted collectively by looking at both frequency and time domain features, as shown in Table 6.2. Time-domain features such as mean, STD, skewness and kurtosis were useful for distribution of data, discovering any irregularities of examined area, and obtaining an evenness to a distribution of data, respectively [261, 316]. Q3 and Q1 showed how the observation data were dispersed in the two sides of the median [317, 318]. The statistical domain features proven to be helpful for choosing most relevant and meaningful features, contributing to the accurate identification and concentrations quantities in aqueous solutions [261, 317, 318]. In this study, frequency domain features were also employed such as special entropy and spectral power. The block diagram of the proposed classification system for different days based on multi-domain features extraction approach is shown in Figure 6.6.

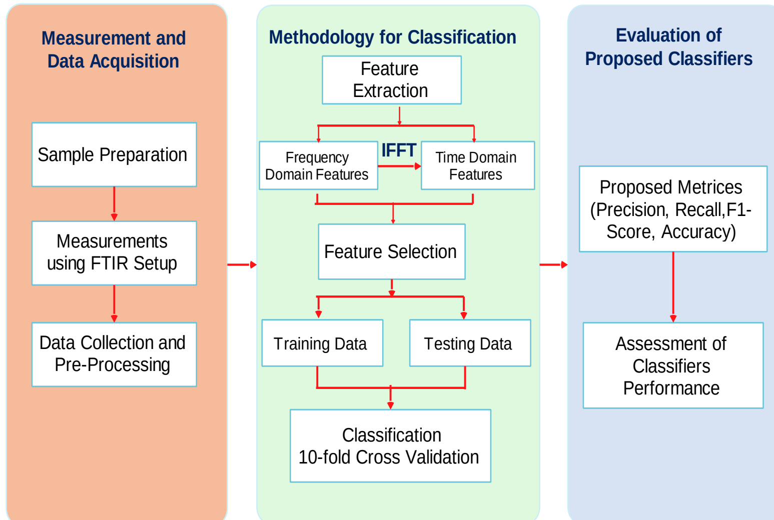


Figure 6.6: The methodological approach of proposed algorithm implementation process.

6.2.2 Classification Methodology

In this research, four classification techniques were considered used namely, SVM, KNN and random forest (RF), and decision tree (D-tree) [319]. In this study, considering the measurements obtained using THz waves, two scenarios have been considered and developed a classifier model for them. The concept behind the formulation of these two scenarios is to adapt the real-life situations where the purity of water is extremely essential for the safety of human health. Considering this significant issue, the performance of all four classifiers were analysed and tested for accurate identification of impurities and to trace the specific amount of constituents' concentrations in aqueous solutions for the given data. For this purpose, the measured dataset was randomly separated into training and testing with division of 70% and 30%, respectively.

For this purpose, Python SciKit library was used as it has been widely utilized in data-science discipline [320]. All the measurements data was converted into CSV format so that they can be easily processed by Scikit library. The dataset for different constituents with various concentrations are properly labelled to execute the supervised ML technique. In this regard, 10-fold cross validation technique was considered to critically analyse ML algorithms where each dataset are examined as test data and the remaining were taken as training data and this process was continued until all dataset are tested, resulting in the average results across all the repetitions.

In comparison to other ML techniques, the KNN algorithm is well-known for its simplicity and ease of operation [321]. This technique operates by evaluating the testing data to the training data. In this scenario, K sample are assigned to a feature of training data and subsequently, testing data is allocated to k sample that closely matches the new data. Thus, tuning this fundamental parameter of k-sample plays a significant role in achieving the ultimate performance of this classifier [321–323]. Furthermore, the SVM operates mainly on two classes and is formally defined by dividing hyper-plane as a discriminatory classifier. The hyper-plane acts as a decision borderline for classification of datasets between two classes. Equation 6.2 represents how the SVM operates [323].

$$\begin{aligned} \bar{w} \cdot u + \bar{b} > 0 \\ \bar{w}u + \bar{b} < 0 \end{aligned} \tag{6.2}$$

In above equation,‘w’ indicates the weight vector, ’u’ displays the input vector and b denotes a constraint. Furthermore, random forest is a set of trees for making decisions [319]. Every tree allows performance prediction by searching for features found during the training process. The majority of prediction is the final prediction for the Random Forest [319].

Feature Selection

In applications such as performing the measurements and dealing with various instruments, possibility of some superfluous and extraneous features is increased which may result in lowering the classification performance. Therefore, it was essential to eliminate those features in order to enhance the classification performance of proposed classifiers as well as reducing the computational costs for deployment. To do so, three feature selection techniques namely, sequential forward selection (SFS), and Relief based selection algorithm (Relief-F) which are widely used are considered to accomplish the feature selection procedure [325]. In SFS method, at the start, empty features are being replaced by some noticeable features which helps to enhance the overall accuracy [325]. Compare to SFS, Relief-F can present a relatively effective approach by considering the function relationships for evaluating the weights of features for appropriate classification and selection instead of relying on different classifiers [326].

Just as precision and recall, individually, are incapable of covering all key aspects of accuracy, thus, F1-score employ the cumulative mean approach to show their performances. By this way, all aspects are considered and demonstrate the overall accuracy. The higher the score, the better the accuracy. Applying these feature selections has considerably yield an improvement of 5%, 4%, 7% and 3% in RF, SVM, D-Tree and KNN, respectively. Furthermore, the additional advantage of feature selection is the further reduction of overall number of features needed for the optimal set, hence computation weights is also optimized for optimal results.

Evaluation of Classifiers Performance Metrices

In this section, the performance of all proposed classifiers was evaluated by using four commonly metrices such as, accuracy, precision, recall (also known as true positive), and F1-score [327]. Table 6.3 presents the list of classifier performance metrices. Here, precision

Table 6.3: Metrices used for the performance evaluation of Classifiers.

Metrices	Formulas
Precision	$\frac{TP}{TP+FP}$
Recall	$\frac{TP}{TP+FN}$
F1-score	$2x\frac{\text{Precision* Recall}}{\text{Precision + Recall}}$

metric is employed to evaluate the precision of one of the classifications relative to all other classifications. In addition, recall or sensitivity values shows the possibility of occurring accurate classification of categorised classes from the remaining classes. Finally, F1-score is normally employed to obtain the average between the Precision and Recall metrices. In this study, the key objective of using these commonly agreed metrices was primarily to detect any potential mis-classification, resulting in inaccurate details about the presence of impurities in aqueous solutions [327].

6.3 Results and Discussions

This section presents the metrics evaluation of classifiers technique using various feature selection techniques. It is perceived that after selecting the relevant features, execution time taken by classifiers for performing ten-fold cross-validation was considerably reduced. The ten-fold cross-validation is also more suitable for the given phenomena because the data-set is not very large and is often the reality with water quality data-sets. In cross-validation, the data is separated into k subsets and is repeated overall the available data-sets, given that K-1 subsets as training set and 1 subset as testing set. Though, due to iterations, this method is considered as computationally intensive technically challenging, however, it is seemingly suitable for the given data. Table 6.4 depicted the quality metrices performance for all proposed classifiers ranging from 0 to 1, indicating the estimation of impurities solutions detection added to the aqueous solutions.

By analysing the results, it can be noticed that RF showed a higher accuracy 84.74% for the identification of unknown ingredients in aqueous solutions followed by KNN, D-Tree and SVM, D-Tree and KNN, with 80.08%, 78.81%, and 74.57%, respectively. Intriguingly, the assessment of metrices for the sugar displayed an effective performance, showing 1 for all classifiers except RF, revealing that sugar is compound of other ingredients. The obtained results by KNN model also shows adequate performance considering the absorption spectra of glucose and sugar in different concentrations as both glucose and sugar molecules broadened in aqueous solutions. Furthermore, despite the distinctive complexities and chemical dynamics emanating from bio-

Table 6.4: Classification Performance of all three Classifiers using Tenfold Cross Validation

Classification Performance by Applying Tenfold Cross Validation					
Classifier Algorithm	Solvent	Quality Metrices			Accuracy %
		Precision	Recall	F1-Score	
Random Forest	Glucose	0.74	0.82	0.78	84.74%
	Salt	0.61	0.64	0.62	
	Sugar	0.67	0.71	0.69	
SVM	Glucose	0.61	0.64	0.62	74.57%
	Salt	0.67	0.64	0.65	
	Sugar	1.00	1.00	1.00	
D-Tree	Glucose	0.67	0.71	0.69	78.81%
	Salt	0.73	0.69	0.71	
	Sugar	1.00	1.00	1.00	
KNN	Glucose	0.68	0.74	0.71	80.08%
	Salt	0.75	0.69	0.72	
	Sugar	1.00	1.00	1.00	

molecular vibrations and constituents of distinct solvents, the evaluation of classifiers can be deemed as relatively efficient and is certainly above the alarming stage. Considering a real-life scenario, the proposed classifier methodology can be substantial by using the amalgamation of highly sensitive and good penetration feature of THz with ML approach to detecting the contagious contaminants in pure water.

The proposed study, in addition to discovering unknown contaminants in aqueous solutions, also quantifies and unravel the estimate prediction of quantity of contaminants added in aqueous solutions. For this purpose, classifiers model was developed, and their efficiency was assessed using the quality metrices. Upon a close inspection of results attained in Table 6.5, 6.6, and 6.7, it was noticed that RF showed relatively enhanced performance compared to other algorithms, showing 97%, 95% and 85.24% for salt, glucose and sugar concentrations, respectively. Also, despite having virtually molecular configuration and morphological structure of glucose and sugar, all classifiers as observed, proved to have reasonable predictions of different concentrations levels for both glucose and sugar in aqueous solutions, ranging from 87.02% to 96.61% as shown in Table 6.5. For salt concentrations, RF outperformed other algorithms, showing a prediction accuracy of 97.98% for distinct concentrations in aqueous solutions. Hence, it was concluded that the Random Forest yields considerable reliability and promising accuracy results in both scenarios, recognizing the substances as well as its precise concentrations solutions in aqueous solutions compared to other classifiers.

However, some limitations can adversely affect the machine learning algorithms, resulting in degrading the overall performance. This unintended situation appeared to have rarely occurred because of selecting inadequate variables due to its high intricacy. Nonetheless, machine learning based models are still a feasible substitute to the physically dependent modelling in

Table 6.5: Classification Performance of KNN by Applying Tenfold Cross Validation

Performance of KNN by Applying Tenfold Cross Validation					
Solvent	Conc.	Quality Metrices			Accuracy %
		Precision	Recall	F1-Score	
Salt	5%	1.00	0.84	0.91	81.53%
	10%	1.00	0.77	0.87	
	20%	0.80	0.80	0.80	
	30%	0.61	0.85	0.71	
Sugar	5%	1.00	0.97	0.98	96.61%
	10%	1.00	0.97	0.98	
	20%	0.93	0.96	0.95	
	30%	0.93	0.97	0.95	
Glucose	5%	1.00	0.86	0.93	87.02%
	10%	1.00	0.87	0.80	
	20%	0.71	0.91	0.80	
	30%	0.85	0.84	0.84	

Table 6.6: Classification Performance of SVM by Applying Tenfold Cross Validation

Performance of SVM by Applying Tenfold Cross Validation					
Solvent	Conc.	Quality Metrices			Accuracy %
		Precision	Recall	F1-Score	
Salt	5%	1.00	0.84	0.91	86.59%
	10%	1.00	0.81	0.89	
	20%	0.88	0.90	0.88	
	30%	0.87	0.93	0.89	
Sugar	5%	0.86	0.97	0.92	93.11%
	10%	0.95	0.89	0.92	
	20%	0.94	0.98	0.96	
	30%	1.00	0.87	0.93	
Glucose	5%	0.92	0.95	0.94	96.88%
	10%	1.00	0.96	0.98	
	20%	0.89	0.95	0.92	
	30%	1.00	0.97	0.99	

Table 6.7: Classification Performance of D-Tree by Applying Tenfold Cross Validation

Performance of D-Tree by Applying Tenfold Cross Validation					
Solvent	Conc.	Quality Metrices			Accuracy %
		Precision	Recall	F1-Score	
Salt	5%	1.00	0.84	0.91	82.10%
	10%	1.00	0.78	0.88	
	20%	0.81	0.81	0.81	
	30%	0.62	0.85	0.72	
Sugar	5%	1.00	0.97	0.98	93.01%
	10%	1.00	0.97	0.98	
	20%	0.93	0.96	0.95	
	30%	0.93	0.97	0.95	
Glucose	5%	0.95	0.97	0.96	95.51%
	10%	1.00	0.96	0.98	
	20%	0.98	0.95	0.93	
	30%	1.00	0.97	0.99	

predicting the realistic scenarios, where small error can be fatal to the public health and safety. Keeping this mind, in this study, the strong aim of applying cross-validation technique was to evaluate the consistency of proposed classifiers by minutely assessing the absorption spectra characteristics of different substances concentrations in aqueous solutions, providing real-time monitoring of unknown substance, and can detect early symptoms of contamination's in water. Furthermore, these preliminary results obtained from the amalgamation of ML with THz waves have the potential to curtail any microbiological contaminants in aqueous solutions and mitigate their harmful effects on human health.

6.4 Summary

In this research study, the use of non-invasive THz feature and ML enabled optimized technological solution was presented to detect various substances and their distinct concentrations in aqueous solutions. In this process, the FTIR system measured the absorption spectra and characteristics of salt, sugar and glucose solutions with varying levels for two hours and collected 338 data points for every specimen and regarded them as features. Since the observations were recorded at laboratory, there might be the possibility of some distortion in measurements. To prevent this, we performed features selection to discard any spurious that may yield forged observations of substance concentrations in aqueous solutions, given the public protection. The selection of meaningful and significant features drastically enhanced the classifier performance for detecting the substance solutions in aqueous solutions. Furthermore, the comprehensive cross-validation methodology exhibited in most cases, RF model showed reliability, and achieved highest classification accuracy in identifying the salt solutions and its quantity in aqueous solutions, compared to other classifiers. Moreover, KNN, D-Tree and SVM displayed substantial performance particularly for sugar and glucose concentrations in aqueous solutions.

These preliminary results showed a notable relationship of THz waves with ML techniques. It also fully reveals the significant influence of ML and its process reliability in terms of detecting the substance solutions as well as their concentrations in aqueous solutions. The outcome of this work has the potential to a play vital role by providing unprecedented and cost-effective opportunity in real-time monitoring to enhance a detection of impurities in water and potentially contribute to the protection of public health.

Chapter 7

Conclusion, Future Work and Challenges

7.1 Conclusion

The THz technology has captivated interests of numerous scientists and driven many researchers, as scientists, horticulturists, plant physiologists, and biochemists from diverse disciplines due to strong penetration and non-invasive characteristics and pervasive properties of THz. Due to these distinctive and unique features, it has attained significant contributions in diverse field applications such as the imaging of unseen hazard items, material characterisation, diagnostic applications including treatment of skin and dental care, effective and quality control of food, and telecommunication. Although, the study of THz in plant science sector is at infant stage and its potentials to propagate satisfactorily through plants is still of one of the least examined research areas until now.

In this thesis, the overarching aim was to highlight the notable contributions and advancements of THz technology and its strong utility in monitoring the health status of living plant leaves by studying the biological and physiological traits as well as observe the internal morphological characteristics of plant leaves at cellular level. In this regard, a novel and non-invasive technique for characterising the WC, and in turn the health of plant leaves was proposed using THz waves. Since plant leaves comprise of a composite biological structure of tissues, distinct bio-molecules like cellulose and synthesis compounds including proteins, carbohydrates and many other molecular weight compounds. However, the main focus was to determine the precise amount of WC of various leaves because it is not only an essential component but an important nutrient to the process of photosynthesis, and transpiration in the overall process of growth.

Based on the aforementioned aim, an experiment was performed on eight different leaves and the electromagnetic properties of all leaves were determined for four consecutive days through the measured scattering parameters using a material characterization Kit. The electromagnetic parameters are measured in simple, fast, and non-invasive manner using THz material characterisation kit. Moreover, the weight and thickness of the leaves were also recorded at the same time. observed that the leaves became increasingly transparent to the THz waves through the course of four days experiment, Similar decaying trends were observed in the peak values of the real part of the extracted relative permittivity as the decreasing weight due to loss of WC. Moreover, The structural integrity and configuration of leaves were also considered by employing two polytetrafluoroethylene (PTFE) caps which were fitted internally to the waveguide. In this process, a notable correlation between electromagnetic parameters with WC in leaves was determined i.e., change in the WC of leaves was reflected in the electromagnetic parameters at certain frequencies.

Furthermore, In this study, ML was introduced with an integration of THz for the proactive and precise detection of WC in plant leaves to drive and sustain overall crops productivity in an advanced agricultural system. In this regard, it was aimed to assess the performance of the proposed ML algorithm with the incorporation of THz for real-time applications in monitoring any diminutive variations of WC in plants leaves to help in developing digital agricultural systems.The findings obtained from an attempt to integrate ML using THz demonstrated the strong potential for assessing the real-time information and precise estimation of WC in various plants' leaves. The proposed approach showed a strong prospect to provide prolific recommendations and insights for growers to take proactive actions in relations to plants health monitoring. The results in chapter 4 demonstrated that SVM outperformed other classifiers using tenfold and leave-one-observations-out cross-validation for different days classification with an overall accuracy of 98.8%, 97.15%, and 96.82% for Coffee, pea shoot, and baby spinach leaves respectively. In addition, using SFS technique, coffee leaf showed a significant improvement of 15%, 11.9%, 6.5% in computational time for SVM, KNN and D-tree. For pea-shoot, 21.28%, 10.01%, and 8.53% of improvement was noticed in operating time for SVM, KNN and D-Tree classifiers, respectively. Lastly, baby spinach leaf exhibited a further improvement of 21.28% in SVM, 10.01% in KNN, and 8.53% in D-tree in overall operating time for classifiers. These improvements in classifiers produced significant advancements in classification accuracy, indicating a more precise quantification of WC in leaves.

In addition, in this work, a feasible, non-invasive method was also proposed for the accurate detection MC in fruits at cellular level in an automated fashion to maintain a healthy sensory characteristic of fruits. Result have shown that our state-of-the-art technique can be strong prospect for automatic assessment of fruit freshness , which in turn, can help in reducing the health and purification expenses, and optimize economic benefits by maintaining the nutrients level and MC in fruits. The results in chapter 5 illustrated that the performance of SVM exceeded other classifiers results using 10-fold validation and leave-one-observation-out-cross-validation techniques. Moreover, all three classifiers exhibited 100% accuracy for day 1 and 4 with 80% MC value (freshness) and 2% MC value (staleness) of both fruits' slices, respectively. Similarly, for day 2 and 3, an accuracy of 95% was achieved with intermediate MC values in both fruits' slices. This study will pave a new direction for the real-time quality evaluation of fruits in a non-invasive manner by incorporating ML with THz sensing at a cellular level. It also has a strong potential to optimize economic benefits by the timely detection of fruits quality in an automated fashion.

Lastly, the thesis suggested a technological solution for for the water quality detection using THz waves. Since water is considered to be the most essential and vital resources to maintain healthy status of plant leaves for the effective and sustainable progress in agricultural field, therefore, the timely delivery to the vegetation field and efficient use to protect the public and environment health with no intrusion of harmful impurities, safe, reliable manner is one of huge challenge amid to the ongoing climate transformations. To address this significant issue, non-invasive technological solution was presented that can provide the approximate prediction and detection of even the smallest of contaminants in distilled water due to high sensitivity and non-destructive nature and can also produce high optical throughput. The ML driven technique employed a FTIR system that operates in a THz frequency range of interest operating from 1 THz to 20 THz to precisely determine the various solvents constituents' characteristics in aqueous solutions non-invasively. The suggested technique has a strong potential to proactively determine the presence of any added impurities in aqueous solutions in real time to protect the environment, include early alerts to protect the public health, and reduce any superfluous costs.

The results demonstrated that RF obtained a higher accuracy of 84.74% in identifying the substance in aqueous solutions. Moreover, it was also found that RF with 97.98%, outperformed other classifiers for estimation of salts concentration added in aqueous solutions. However, for sugar and glucose concentrations, SVM exhibited a higher accuracy of 93.11% and 96.88%, respectively, compared to other classifiers. Thus, proposed technique incorporating ML with THz waves, may be significant in providing an efficient, cost-effective and real-time monitoring for water quality detection system.

7.2 Future Work

In line with the work presented, the future research directions which would make potential and natural progression to complete the studies in the thesis are summarised:

7.2.1 Mathematical modelling of leaf

In this study, we have obtained electromagnetic measurements of various leaves using scattering parameters which indicate the health status of leaves by showing distinct response. This work can be further extended and validated by obtaining the simulation model of particular leaf and observe the scattering parameter by applying THz waves. This proposed approach can further help to enhance both reliability and credibility of technique used for determining the electromagnetic parameters in THz region to minutely observe the inter-morphological biological traits and characteristics of different leaves at cellular level. By comparing the practical measurement of plant leaves with in-line mathematical model of leaf will provide rich recommendations and insights about the electromagnetic properties of leaf at THz with the ultimate scope of studying the behaviour of leaves in more detail.

7.2.2 Recognition of species

Although researchers have proposed various techniques to discover the distinctive features of plants leaves and have achieved considerable progress. However, a novel electromagnetic technique can be introduced using THz waves for the recognition of diverse plant leaves by investigating their structural behaviour and complex traits at cellular level. In this regard, It is envisioned that this innovative approach will not only help to accurately identify the spectral plant features which may correlate with specific plant stresses, and would be integral to closely monitor the future development life support systems in current spaceflight climate.

7.2.3 Interaction of THz with more specimens

We have performed measurements for plants leaves and for the detection of impurities in water. These specimens can be further increased by adding more specimens to validate the proposed technique. With the comprehensive measurements of additional specimens, interaction of THz with diverse specimens can be established to study and investigate the behaviour and their response in different THz region. Therefore, it would be appropriate to perform the experiments on further samples to get the parameters of interest.

7.2.4 Safety issues: Heating of THz wave is significant concern

While using the MCK to determine the material properties of different living plants leaves from scattering parameters, overheating of system has been a major concern. The occurrence of excessive heating of system that generate THz waves may have precipitated some random noise, resulting in fictitious findings of the species and have ultimately affect the health of leaves. Therefore, the study of THz wave heating affects on the health of plants can be another substantial prospect for future research direction as it may reveal its consequences which could be detrimental to tissue of leaves, eventually, affecting the crops productivity and economy.

7.2.5 Challenges

The THz non-destructive detection technology in food inspection and water (such as juice, beverages, and alcohol) contamination detection is emerging as a new area of study. With unique superiority, the THz technology has attracted many researchers in this field. This field has made great progress, however, it is still at an initial stage and there are still many challenges that need to be addressed. Some of the most important challenges are given as follows:

Distraction and Absorption of Water

THz non-destructive detection for food and water contamination detection is limited by water thanks to the firm absorption of THz radiation by water. Therefore, a major challenge is the overwhelming fading of THz radiation by water molecules. For example, in food detection THz technology is not appropriate for detecting high humidity products with a thickness greater than 1mm. The inability of THz technology to detect the bio-molecular interactions in solution is the major hurdle facing terahertz further applications. Methods for overcoming this challenge are given in the following research.

One method that applies to both food and water contamination is THz-TDS [138]. The THz-TDS system can obtain the amplitude and phase information of the THz pulse simultaneously [196]. Moreover, by applying the Fourier transform on the time waveform, the optical parameters of the specific sample can be directly acquired such as the absorption coefficient and refractive index. For food detection, many researchers proposed several methods using a broadband THz system or performing the measurement in the low-frequency terahertz region [147]. For water contamination detection, a method with reflective pulsed terahertz tomography is proposed in [184].

Low Penetration Depth

Another challenge of THz technology is the low penetration depth of THz radiation, especially measuring the liquid and meat products. This problem can be addressed using the penetration-enhancing agents, the use of graphene composite, and strengthening the intensity of terahertz source. In [49], it presented that the penetration depth can be enhanced using penetration-enhancing agents. In [133], it was demonstrated that a new graphene composite with a double circular metal ring array deposited on graphene can boost the terahertz absorption. THz power enhancement is also an effective way to increase penetration depth [186].

Scattering Effects

Scattering effects is also a common problem in THz-TDS transmission measurements of solid state samples, especially for irregular granular samples. To reduce or remove the scattering effects, plasma and meta-material are used. In addition, some better algorithms can also be used, such as wavelet transform and Monte Carlo method, to extract spectral data to minimize the scattering effects. A novel method by coating plasma and meta-material is presented to reduce the scattering effects of target, which has a better performance than coated with plasma [142]. Also, a modified Monte Carlo method is utilized to exclude or reduce the scattering effects [157].

More Compact Systems at Higher Powers

With the requirements of integration and miniaturization, the demand for significant power and the higher dynamic range of THz application system is of immense-importance. To overcome this challenge, it needs an extensive literature review and more adequate and viable techniques. Due to the available THz sources operating at high power, substances and materials which partially absorb THz radiation or wider bodies can be explored.

Appendix A

Snapshot of Systems

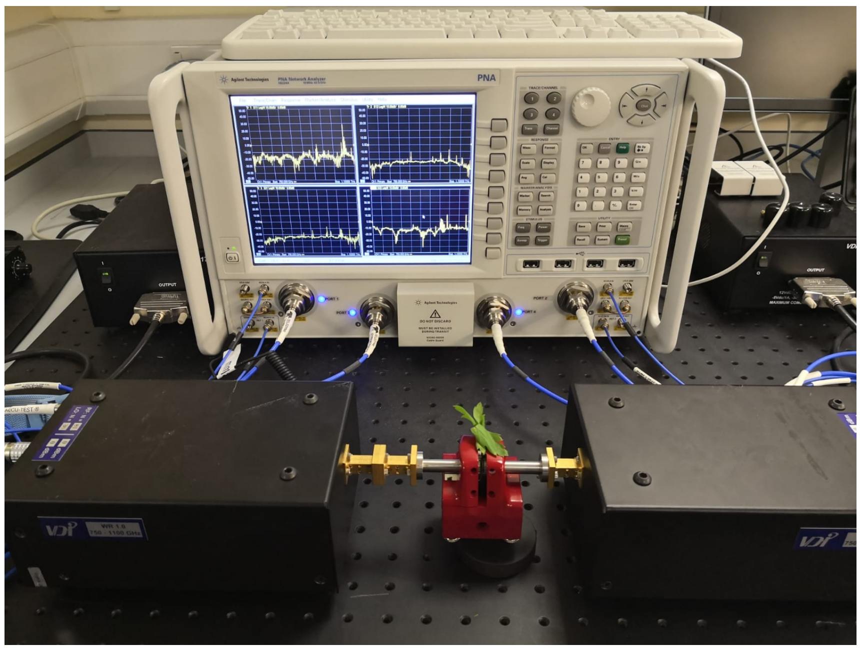


Figure A.1: Snapshot of real system measuring the WC of leaves using THz waves in the frequency range of 0.75 to 1.1 THz.

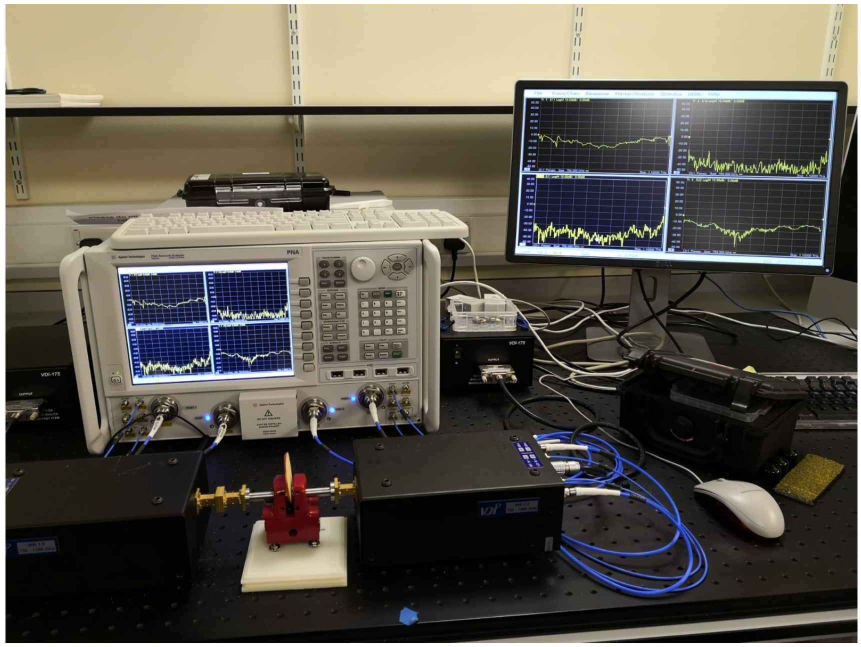


Figure A.2: Snapshot of system measuring the MC in fruits using THz waves in the frequency range of 0.75 to 1.1 THz.

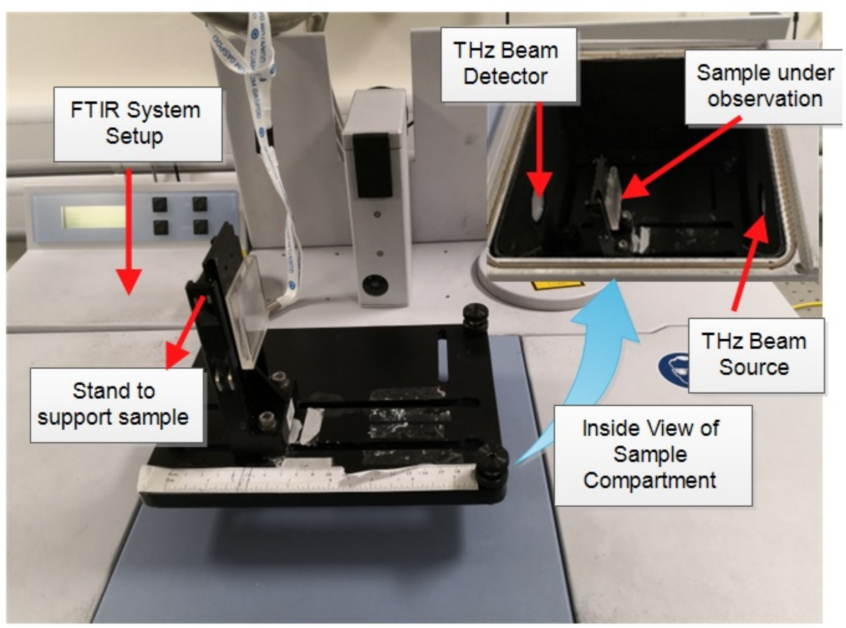


Figure A.3: Snapshot of FTIR system measuring the impurities in salt, sugar and glucose concentrations solutions using THz waves

Bibliography

[1] M. El Beyrouthya, “Nanotechnologies: Novel Solutions for Sustainable Agriculture”. *Adv. Crop Sci. Technol., vol. 02*, no. 03, p. 8863, 2014.

[2] Khaled D., Nuria Novas Castellano, José A. Gazquez Parra, RM. Garcia Salvador and Manzano-Agugliaro F, “Fruit and Vegetable Quality Assessment via Dielectric Sensing.” *Sensors*, (2015).

[3] Zahid, A., Abbas, H.T., Ren, A. et al., “Machine learning driven non-invasive approach of water content estimation in living plant leaves using terahertz waves ”, *Plant Methods 15*, 138 (2019).

[4] A. Zahid, H. T. Abbas, H. Heidari, M. Imran, A. Alomainy and Q. H. Abbasi, “Monitoring the Variability of Water Dynamics in Plant Leaves at Cellular Level Using Terahertz Sensing”, 2019 *Second International Workshop on Mobile Terahertz Systems (IWMTS), Bad Neuenahr, Germany*, 2019, pp. 1-4,

[5] A. Zahid, H. T. Abbas, H. Heidari, M. Imran, A. Alomainy and Q. H. Abbasi, "Electromagnetic Properties of Plant Leaves at Terahertz Frequencies for Health Status Monitoring," 2019 IEEE MTT-S International Microwave Biomedical Conference (IMBioC), Nanjing, China, 2019, pp. 1-4, doi: 10.1109/IMBIOC.2019.8777782.

[6] S. S. Mukhopadhyay, “Nanotechnology in agriculture: Prospects and constraints,” *Nanotechnol. Sci. Appl., vol. 7,* no. 2, pp. 63–71, 2014.

[7] A. Afsharinejad, A. Davy, P. O’Leary, and C. Brenann, “Transmission through single and multiple layers of plant leaves at THz frequencies,” *in GLOBECOM 2017-2017 IEEE Global Communications Conference*, pp. 1–6, IEEE, 2017.

[8] A. Afsharinejad, A. Davy, and M. Naftaly, “Variability of terahertz transmission measured in live plant leaves,” *IEEE Geoscience and Remote Sensing Letters, vol. 14,* no. 5, pp. 636–638, 2017.

[9] T. K¨urner and S. Priebe, “Towards THz communications-status in research, standardization and regulation,” *Journal of Infrared, Millimeter, and Terahertz Waves, vol. 35,* no. 1, pp. 53–62, 2014.

[10] I. E. Carranza, J. P. Grant, J. Gough, and D. Cumming, "Terahertz metamaterial absorbers implemented in cmos technology for imaging applications: scaling to large format focal plane arrays," IEEE Journal of Selected Topics in Quantum Electronics, vol. 23, no. 4, pp. 1–8, 2017.

[11] Q. H. Abbasi, M. U. Rehman, K. Qaraqe, and A. Alomainy, "Advances in body-centric wireless communication: Applications and state-of-the-art". *Institution of Engineering and Technology*, 2016.

[12] J. M. Jornet and I. F. Akyildiz, "Channel modeling and capacity analysis for electromagnetic wireless nanonetworks in the terahertz band," *IEEE Transactions on Wireless Communications, vol. 10, no. 10*, pp. 3211–3221, 2011.

[13] C. J¨ordens, M. Scheller, B. Breitenstein, D. Selmar, and M. Koch, "Evaluation of leaf water status by means of permittivity at terahertz frequencies," *Journal of biological physics, vol. 35, no. 3,* pp. 255–264, 2009.

[14] R. Gente and M. Koch, "Monitoring leaf water content with THz and sub-THz waves," *Plant methods, vol. 11, no. 1*, p. 15, 2015.

[15] A. Afsharinejad, A. Davy, and B. Jennings, "Dynamic channel allocation in electromagnetic nanonetworks for high resolution monitoring of plants," *Nano Communication Networks, vol. 7*, pp. 2–16, 2016.

[16] N. Born, D. Behringer, S. Liepelt, S. Beyer, M. Schwerdtfeger, B. Ziegenhagen, and M. Koch, "Monitoring plant drought stress response using terahertz time-domain spectroscopy," *Plant physiology, vol. 164, no. 4*, pp. 1571–1577, 2014.

[17] B. Breitenstein, M. Scheller, M. K. Shakfa, T. Kinder, T. M¨uller-Wirts, M. Koch, and D. Selmar, "Introducing terahertz technology into plant biology: A novel method to monitor changes in leaf water status," *Journal of Applied Botany and Food Quality, vol. 84, no. 2*, pp. 158–161,2011.

[18] V. Torres, I. Palacios, J. C. Iriarte, I. Liberal, L. G. Santesteban, C. Miranda, J. B. Royo, and R. Gonzalo, "Monitoring water status of grapevine by means of THz waves," *Journal of Infrared, Millimeter, and Terahertz Waves, vol. 37*, no. 5, pp. 507–513, 2016.

[19] A. Afsharinejad, A. Davy, B. Jennings, and C. Brennan, "Performance analysis of plant monitoring nanosensor networks at THz frequencies," *IEEE Internet of Things Journal, vol. 3, no. 1*, pp. 59–69, 2016.

[20] "Swissto12." http://www.swissto12.com/, *2017*. [Online; accessed 15th December 2017].

[21] X. Yang et al., "Biomedical Applications of Terahertz Spectroscopy and Imaging", *Trends in Biotechnology, Volume 34*, Issue 10, 2016,

[22] Q. H. Abbasi, M. Ur Rehman, K. Qaraqe and A. Alomainy (eds.), "Advances in Body-Centric Wireless Communication", *Applications and State-of-the-art*, The IET, July 2016.

[23] P. Boronin, et al., "Capacity and throughput analysis of nanoscale machine communication through transparency windows in the terahertz band", *Nano Communication Networks, 5(3)*,:72–82, 2014.

[24] Y. Li, Y. Wang, Q. Cao, J. Cao and D. Qiao, "A Self-Powered Vibration Sensor With Wide Bandwidth", *IEEE Transactions on Industrial Electronics, vol. 67, no. 1*, pp. 560-568, 2020

[25] Borgatta, J. et al., "Copper based nanomaterials suppress root fungal disease in watermelon (Citrullus lanatus): role of particle morphology, composition and dissolution behaviour", ACS Sustain. Chem. Eng. 6, 14847–14856, 2018.

Zahid, A., et al. "Machine learning driven non-invasive approach of water content estimation in living plant leaves using terahertz waves",

[26] Zhu, Q. et al, "FRET-based glucose imaging identifies glucose signalling in response to biotic and abiotic stresses in rice roots. J", *Plant Physiol. 215, 65–72*, 2017.

[27] Alomainy, A., et al. "Nano-Electromagnetic Communication at Terahertz and Optical Frequencies: Principles and Applications", *Institution of Engineering and Technology*, 2019.

[28] Richardson, J. J. & Liang, K, "Nano-biohybrids: in vivo synthesis of metal-organic frameworks inside living plants", *Small 14*, 2018.

[29] Jordens C, Scheller M, Breitenstein B, Selmar D, Koch M, "Evaluation of leaf water status by means of permittivity at terahertz frequencies", *J Biol Phys. 2009;* 35(3):255–64.

[30] Liew OW, Chong PCJ, Li B, Asundi AK, "Signature optical cues: emerging technologies for monitoring plant health", *Sensors. 2008*, 8:3205–39.

[31] Torres V, Palacios I, Iriarte JC, Liberal I, Santesteban LG, Miranda C, Royo JB, Gonzalo R, "Monitoring water status of grapevine by means of THz waves", *J Infrared Millim Terahertz Waves. 2016*, 37(5):507–13.

[32] Santesteban LG, Palacios I, Miranda C, Iriarte JC, Royo JB, Gonzalo R, "Terahertz time domain spectroscopy allows contactless monitoring of grapevine water status", *Front Plant Sci. 2015*, 6(June):1–9.

[33] Song Z, et al, “Temporal and spatial variability of water status in plant leaves by Terahertz imaging,” *IEEE Trans Terahertz Sci Technol. 2018*, 8(5):520–7

[34] A. Ren et al., “Machine Learning Driven Approach Towards the Quality Assessment of Fresh Fruits Using Non-Invasive Sensing”, *IEEE Sensors Journal, vol. 20, no. 4*, pp. 2075-2083, 2020.

[35] Liakos KG, Busato P, Moshou D, Pearson S, Bochtis D. “Machine learning in agriculture: a review”, *Sensors. 2018*, 18:2674

[36] Wang, K., Sun, D. W., & Pu, H. (2017). “Emerging non-destructive terahertz spectroscopic imaging technique: Principle and applications in the agri-food industry”, *Trends in Food Science and Technology*,

[37] Cleveland, R. F., and Ulcek, J. J. L.(1999), “Questions and Answers about Biological Effects and Potential Hazards of Radio frequency Electromagnetic Fields (4th Edition)”, *Federal Communications Commission Office of Engineering & Technology*

[38] Clark, J., (2000)., “Ionisation Energy”, *Archived from the original."* on 2011-11-26. Retrieved 2011-12-07.

[39] Mittleman, D. M. (2017)., “Perspective: Terahertz science and technology”, *In Journal of Applied Physics (Vol. 122)*.

[40] Suganya, R., Krishnaveni, N., & Senthil, T. S. (2015). “Synthesis and characterization of zinc oxide nanocrystals from chemical and biological methods and its photocatalytic activities”, *International Journal of ChemTech Research, 8(11)*, 490–496.

[41] Hebling, J., Yeh, K.-L., Hoffmann, M. C., Bartal, B., Nelson, K. A. (2008), “Generation of high-power terahertz pulses by tilted-pulse-front excitation and their application possibilities”, *Journal of the Optical Society of America B, 25(7)*, B6.

[42] Dai, J., Liu, J., Zhang, X. C. (2011)., “Terahertz wave air photonics: "Terahertz wave generation and detection with laser-induced gas plasma”, *IEEE Journal on Selected Topics in Quantum Electronics, 17(1)*, 183–190.

[43] Vicario, C., Monoszlai, B., & Hauri, C. P. (2014). “GV/m single-cycle terahertz fields from a laser-driven large-size partitioned organic crystal”, *Physical Review Letters*, 112(21).

[44] Hoffmann, M. C., Fülöp, J. A. (2011)., “Intense ultrashort terahertz pulses: Generation and applications”, *Journal of Physics D: Applied Physics.*,

[45] Woodward, R. M., Cole, B. E., Wallace, V. P., Pye, R. J., Arnone, D. D., Linfield, E. H., & Pepper, M. (2002). “Terahertz pulse imaging in reflection geometry of human skin cancer and skin tissue”, *Physics in Medicine and Biology, 47(21)*, 3853–3863.

[46] Huang, S. Y., Wang, Y. X. J., Yeung, D. K. W., Ahuja, A. T., Zhang, Y. T., Pickwell-Macpherson, E. (2009)., “Tissue characterization using terahertz pulsed imaging in reflection geometry”, *Physics in Medicine and Biology, 54(1)*, 149–160.

[47] Heshmat, B., Andrews, G. M., Naranjo-Montoya, O. A., Castro-Camus, E., Ciceri, D., Sanchez, A. R., Raskar, R. (2017)., “Terahertz scattering and water absorption for porosimetr”, *Optics Express, 25(22)*, 27370–27382.

[48] Hintzsche, H., Jastrow, C., Kleine-Ostmann, T., Kärst, U., Schrader, T., Stopper, H. (2012)., “Terahertz Electromagnetic Fields (0.106 THz) Do Not Induce Manifest Genomic Damage”, *In Vitro. PLoS ONE, 7(9).*

[49] Cheon, H., Yang, H. J., Son, J. H. (2017)., “Toward Clinical Cancer Imaging Using Terahertz Spectroscopy”, *IEEE Journal of Selected Topics in Quantum Electronics*, 23(4).

[50] Shen, Y. C. (2011). “Terahertz pulsed spectroscopy and imaging for pharmaceutical applications: A review. International Journal of Pharmaceutics”, *417(1–2)*, 48–60.

[51] Chopra, N., Yang, K., Abbasi, Q. H., Qaraqe, K. A., Philpott, M., Alomainy, A. (2016)., “THz Time-Domain Spectroscopy of Human Skin Tissue for In-Body Nanonetworks”,IEEE Transactions on Terahertz Science and Technology, 6(6), 803–809.

[52] Piro, G., Yang, K., Boggia, G., Chopra, N., Grieco, L. A., & Alomainy, A. (2015)., “Terahertz Communications in Human Tissues at the Nanoscale for Healthcare Applications”, *IEEE Transactions on Nanotechnology, 14(3)*, 404–406.

[53] Abbasi, Q. H., El Sallabi, H., Chopra, N., Yang, K., Qaraqe, K. A., Alomainy, A. (2016), “Terahertz Channel Characterization Inside the Human Skin for Nano-Scale Body-Centric Networks”, *IEEE Transactions on Terahertz Science and Technology*, 6(3), 427–434.

[54] Abbasi, Q. H., Yang, K., Chopra, N., Jornet, J. M., Abuali, N. A., Qaraqe, K. A., Alomainy, A. (2016)., “Nano-Communication for Biomedical Applications: A Review on the State-of-the-Art from Physical Layers to Novel Networking Concepts”, *IEEE Access*.

[55] Iqbal, M. O., Rahman, M. M. U., Imran, M. A., Alomainy, A., Abbasi, Q. H. (2018), “Modulation Mode Detection & Classification for in-Vivo Nano-Scale Communication Systems Operating in Terahertz Band”, *1–9*.

[56] Kleine-Ostmann, T., Nagatsuma, T. (2011)., “A review on terahertz communications research”, *Journal of Infrared, Millimeter, and Terahertz Waves.*

[57] Zhan, H., Zhao, K., Bao, R., & Xiao, L. (2016)., “Monitoring PM2.5 in the Atmosphere by Using Terahertz Time-Domain Spectroscopy”, *Journal of Infrared, Millimeter, and Terahertz Waves, 37(9)*, 929–938.

[58] Dworak, V., Augustin, S., Gebbers, R. (2011)., “Application of Terahertz radiation to soil measurements: Initial results”, *Sensors, 11(10)*, 9973–9988.

[59] Gowen, A. A., O’Sullivan, C., O’Donnell, C. P. (2012)., “Terahertz time domain spectroscopy and imaging: Emerging techniques for food process monitoring and quality control”, *Trends in Food Science and Technology.*,

[60] Liu, H.-B., Zhong, H., Karpowicz, N., Chen, Y., Zhang, X.-C. (2007)., “Terahertz Spectroscopy and Imaging for Defense and Security Applications”, *Proceedings of the IEEE, 95(8)*, 1514–1527.

[61] Finneran, I. A., Welsch, R., Allodi, M. A., Miller, T. F., Blake, G. A. (2016)., “Coherent two-dimensional terahertz-terahertz-Raman spectroscopy”, *Proceedings of the National Academy of Sciences*, 113(25), 6857–6861.

[62] Lu, J., Li, X., Hwang, H. Y., Ofori-Okai, B. K., Kurihara, T., Suemoto, T., Nelson, K. A. (2017). “Coherent Two-Dimensional Terahertz Magnetic Resonance Spectroscopy of Collective Spin Waves”, *Physical Review Letters*, 118(20).

[63] Lu, J., Zhang, Y., Hwang, H. Y., Ofori-Okai, B. K., Fleischer, S., Nelson, K. A. (2016). “Nonlinear two-dimensional terahertz photon echo and rotational spectroscopy in the gas phase”, *Proceedings of the National Academy of Sciences, 113(42)*, 11800–11805.

[64] Mathanker, S. K., Weckler, P. R., Wang, N. (2013). “X-Ray Applications in Food and Agriculture: A Review”, *Transactions of the ASABE*, 1227–1239.

[65] Heyden, M., Sun, J., Funkner, S., Mathias, G., Forbert, H., Havenith, M., Marx, D. (2010)., “Dissecting the THz spectrum of liquid water from first principles via correlations in time and space”, *Proceedings of the National Academy of Sciences, 107(27)*, 12068–12073.

[66] Ogawa, Y., Hayashi, S., Otani, C., Kawase, K. (2008)., “Terahertz Sensing for Ensuring the Safety and Security”, *Progress In Electromagnetics Research Symposium, Hangzhou, China*, 409–413.

[67] Yan, Z.-K., Zhang, H.-J., & Ying, Y.-B. (2007)., “Research progress of Terahertz wave technology in quality measurement of food and agricultural products”, *Guang Pu Xue Yu Guang Pu Fen Xi Guang Pu, 27(11)*, 2228–2234.

[68] Yin, M., Tang, S., & Tong, M. (2016). “The application of terahertz spectroscopy to liquid petrochemicals detection: A review”, *Applied Spectroscopy Reviews, 51(5)*, 379–396.

[69] Yang, Y., Mandehgar, M., & Grischkowsky, D. R. (2011). “Broadband THz pulse transmission through the atmosphere”, *IEEE Transactions on Terahertz Science and Technology, 1(1)*, 264–273.

[70] Yang, L., Sun, H., Weng, S., Zhao, K., Zhang, L., Zhao, G., ... Jia’er, C. (2008)., “Terahertz absorption spectra of some saccharides and their metal complexes. Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy”, *69(1)*, 160–166.

[71] Suzuki, T., Ogawa, Y., & Kondo, N. (2011). “Characterization of pesticide residue, Cis-permethrin by terahertz spectroscopy”, *Engineering in Agriculture, Environment and Food, 4(4)*, 90–94.

[72] Xiao-Li, Z., & Jiu-Sheng, L. (2011). “Diagnostic techniques of talc powder in flour based on the THz spectroscopy”, *In Journal of Physics: Conference Series (Vol. 276)*,

[73] Xu, J., Plaxco, K. W., & Allen, S. J. (2006). “Collective dynamics of lysozyme in water: Terahertz absorption spectroscopy and comparison with theory”, *Journal of Physical Chemistry B, 110(47)*, 24255–24259.

[74] Cui, Y., Mu, K., Wang, X., Zhang, Y., Zhang, C. (2009). “Measurement of mixtures of melamine using THz ray,August 2009)”, *73851E.*,

[75] Ahi, K., Anwar, M. (2016)., “Advanced terahertz techniques for quality control and counterfeit detection”, *Vol. 9856, p. 98560G*.

[76] Chattopadhyay, G., Reck, T. J., Jung-Kubiak, C., Siles, J. V., Lee, C., Lin, R., Mehdi, I. (2013)., “Silicon micromachining for terahertz component development”, *In 2013 IEEE MTT-S International Microwave and RF Conference, IMaRC 2013.*

[77] Malhotra, I., Jha, K. R., Singh, G. (2018).,“Terahertz antenna technology for imaging applications: A technical review”, *International Journal of Microwave and Wireless Technologies.*

[78] Tonouchi, M. (2007)., “Cutting-edge terahertz technology”, *Nature Photonics*,

[79] Ghann, W., Uddin, J. (2017)., “Terahertz Spectroscopy - A Cutting Edge Technology”, *3–20*.

[80] Lewis, R. A. (2014)., “A review of terahertz sources”, *Journal of Physics D: Applied Physics,* 47(37).

[81] Zolotarev, V. (2007)., “Spectral emissivity of a Globar lamp in the 2-50-μm region”, *Journal of Optical . . . , 74(June)*, 378–384.

[82] Knyazev, B. A., Kulipanov, G. N., Vinokurov, N. A. (2010). Novosibirsk., “terahertz free electron laser: Instrumentation development and experimental achievements”, *Measurement Science and Technology, 21(5).*

[83] Ozyuzer, L., Koshelev, A. E., Kurter, C., Gopalsami, N., Li, Q., Tachiki, M., Welp, U. (2007)., “Emission of coherent THz radiation from superconductors”, *Science, 318(5854)*, 1291–1293.

[84] Hübers, H. W., Pavlov, S. G., Shastin, V. N. (2005). “Terahertz lasers based on germanium and silicon. Semiconductor Science and Technology”, *20(7).*

[85] Wu, X. L., Xiong, S. J., Liu, Z., Chen, J., Shen, J. C., Li, T. H., Chu, P. K. (2011). “Green light stimulates terahertz emission from mesocrystal microspheres”, *Nature Nanotechnology, 6(2)*, 103–106.

[86] Kress, M., Löffler, T., Eden, S., Thomson, M., Roskos, H. G. (2004). “Terahertz-pulse generation by photoionization of air with laser pulses composed of both fundamental and second-harmonic waves”, *Optics Letters, 29(10)*, 1120.

[87] Bogue, R. (2018)., “Sensing with terahertz radiation: A review of recent progress”,*Sensor Review, 38(2)*, 216–222.

[88] Fedorova, K. A., Gorodetsky, A., Rafailov, E. U. (2017)., “Compact all-quantum-dot-based tunable THz laser source”, *IEEE Journal on Selected Topics in Quantum Electronics*, 23(4), 1–5.

[89] Wu, B., Zhang, Z., Cao, L., Fu, Q., & Xiong, Y. (2018)., “Electro-optic sampling of optical pulses and electron bunches for a compact THz-FEL source”, *Infrared Physics and Technology, 92*, 287–294.

[90] Liu, S. (2018). "High peak power THz source for ultrafast electron diffraction." *AIP Advances, 8(1).*

[91] Nasse, M. J., Schuh, M., Naknaimueang, S., Schwarz, M., Plech, A., Mathis, Y. L., Müller, A. S. (2013). “FLUTE: A versatile linac-based THz source”., *In Review of Scientific Instruments (Vol. 84).*

[92] Morohashi, I., Irimajiri, Y., Kumagai, M., Kawakami, A., Sakamoto, T., Sekine, N., Hosako, I. (2016).,"Terahertz source with broad frequency tunability up to 3.8 THz using MZ-FCG for frequency reference in phase-locking of THz source devices", *2016 IEEE International Topical Meeting on Microwave Photonics, MWP 2016*, 14, 126–128.

[93] Tan, P., Fu, Q., Li, Y. B., Xiong, Y. Q., Qin, B., Fan, T. Q. Y. M. W., & Li, D. (2012)., "Effects of Metal Mirrors Reflectivity and Aberrations on Thz Fel Radiation Performance", *Proceedings of IPAC2012, New Orleans*, Louisiana, USA, 1729–1731.

[94] Ye, H. B., Zhang, Y. H., & Shen, W. Z. (2006)., "Carrier transport and optical properties in GaAs far-infrared/terahertz mirror structures", *Thin Solid Films, 514(1–2)*, 310–315.

[95] Krumbholz, N., Gerlach, K., Rutz, F., Koch, M., Piesiewicz, R., Kürner, T., Mittleman, D. (2006), "Omni-directional terahertz mirrors: A key element for future terahertz communication systems", *Applied Physics Letters*, 88(20).

[96] Zhang, H., Xu, X., Fan, F., & Chang, S. (2015). "Study on a tunable narrow-band filter based on magnetic defects in photonic crystal in the terahertz region", *Optical Engineering, 54(4)*, 047104.

[97] Hristov, H. D., Rodriguez, J. M., Grote, W. (2012)., "The grooved-dielectric Fresnel zone plate: An effective terahertz lens and antenna", *Microwave and Optical Technology Letters, 54(6)*, 1343–1348.

[98] Jiang, X.-Y., Ye, J.-S., He, J.-W., Wang, X.-K., Hu, D., Feng, S.-F., Zhang, Y. (2013)., "An ultrathin terahertz lens with axial long focal depth based on metasurfaces", *Optics Express, 21(24)*, 30030.

[99] Scherger, B., Jördens, C., & Koch, M. (2011). "Variable-focus terahertz lens",*Optics Express, 19(5)*, 4528.

[100] Furlan, W. D., Ferrando, V., Monsoriu, J. A., Zagrajek, P., Czerwińska, E., Szustakowski, M. (2016)., "3D printed diffractive terahertz lenses", *Optics Letters, 41(8)*, 1748.

[101] Siemion, A., Siemion, A., Makowski, M., Suszek, J., Bomba, J., Czerwiński, A., Sypek, M. (2012)., "Diffractive paper lens for terahertz optics", *Optics Letters, 37(20)*, 4320.

[102] Partanen, A., Väyrynen, J., Hassinen, S., Tuovinen, H., Mutanen, J., Itkonen, T., Venäläinen, T. (2012). "Fabrication of terahertz wire-grid polarizers", *Applied Optics, 51(35)*, 8360.

[103] Cheng, L.-J., Liu, L. (2013)., "Optical modulation of continuous terahertz waves towards cost-effective re-configurable quasi-optical terahertz components", *Optics Express, 21(23)*, 28657.

[104] Zubair, A., Tsentalovich, D. E., Young, C. C., Heimbeck, M. S., Everitt, H. O., Pasquali, M., & Kono, J. (2016)., “Carbon nanotube fiber terahertz polarizer”, *Applied Physics Letters*, 108(14).

[105] Ikamas, K., Cibiraite, D., Lisauskas, A., Bauer, M., Krozer, V., Roskos, H. G. (2018)., “Broadband Terahertz Power Detectors Based on 90-nm Silicon CMOS Transistors with Flat Responsivity Up to 2.2 THz”, *IEEE Electron Device Letters, 39(9)*, 1413–1416.

[106] Rogalski, A., & Sizov, F. (2011)., “Terahertz detectors and focal plane arrays”, *Opto-Electronics Review*.

[107] Lee, A. W. M., Williams, B. S., Kumar, S., Hu, Q., Reno, J. L. (2006). “Real-time imaging using a 4.3-THz quantum cascade laser and a 320 × 240 microbolometer focal-plane array”, *IEEE Photonics Technology Letters, 18(13)*, 1415–1417.

[108] Wang, N., Javadi, H., & Jarrahi, M. (2016). “Heterodyne terahertz detection with plasmonic photomixers”, *In IEEE MTT-S International Microwave Symposium Digest (Vol. 2016–Augus*, pp. 1–3).

[109] Han, S.-P., Ko, H., Park, J.-W., Kim, N., Yoon, Y.-J., Shin, J.-H., … Park, K. H. (2013)., “InGaAs Schottky barrier diode array detector for a real-time compact terahertz line scanner”, *Optics Express, 21(22)*, 25874.

[110] Komiyama, S. (2011)., “Single-photon detectors in the Terahertz range”, *IEEE Journal on Selected Topics in Quantum Electronics, 17(1)*, 54–66.

[111] Betancourt Arellano, R. S., Nahuelhual Muñóz, L. (2017)., “Servicios ecosistémicos y bienestar local: Caso de estudio sobre productos de medicina natural en Panguipulli, sur de Chile”, *In Ecologia Austral* (Vol. 27, pp. 99–112).

[112] Akyildiz, I. F., Jornet, J. M., & Han, C. (2014)., “Terahertz band: Next frontier for wireless communications”, *Physical Communication, 12*, 16–32.

[113] Hack, E., Valzania, L., Gäumann, G., Shalaby, M., Hauri, C. P., Zolliker, P. (2016)., “Comparison of thermal detector arrays for off-axis THz holography and real-time THz imaging”, *Sensors (Switzerland), 16(2)*.

[114] Altan, H., Sahin, A. B., Çolak, T., Yildirim, I. O., Demirsoy, B., Güray, C., Telli, A. (2018)., “TRay: active THz imaging autonomous biometric secured access control and tracking gate system”, *In Millimetre Wave and Terahertz Sensors and Technology* XI (Vol. 10800, p. 10).

[115] Baek, S. H., Kang, J. H., Hwang, Y. H., Ok, K. M., Kwak, K., Chun, H. S. (2016)., “Detection of Methomyl, a Carbamate Insecticide, in Food Matrices Using Terahertz Time-Domain Spectroscopy”, *Journal of Infrared, Millimeter, and Terahertz Waves."*, 37(5), 486–497.

[116] Baek, S. H., Lim, H. Bin, Chun, H. S. (2014)., “Detection of melamine in foods using terahertz time-domain spectroscopy”, *Journal of Agricultural and Food Chemistry*, 62(24), 5403–5407.

[117] Corsi, C, Sizov, F., “THz and Security Applications”, 2014, *Springer Netherlands*, 1874-6500.

[118] Chen, T., Zhang, Q., Li, Z., Yin, X., Hu, F. (2018)., “Experimental and theoretical investigations of tartaric acid isomers by terahertz spectroscopy and density functional theory”, *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy, 205*, 312–319.

[119] Cheon, H., Yang, H. J., Lee, S. H., Kim, Y. A., Son, J. H. (2016)., “Terahertz molecular resonance of cancer DNA”, *Scientific Reports, 6.*

[120] Chua, H. S., Obradovic, J., Haigh, A. D., Upadhya, P. C., Hirsch, O., Crawley, D., Linfield, E. H. (2005)., “Terahertz time-domain spectroscopy of crushed wheat grain”, *In IEEE MTT-S International Microwave Symposium Digest* (Vol. 2005, pp. 2103–2106).

[121] Da Silva, V. H., Vieira, F. S., Rohwedder, J. J. R., Pasquini, C., Pereira, C. F. (2017)., “Multivariate quantification of mebendazole polymorphs by terahertz time domain spectroscopy (THZ-TDS)”, *Analyst, 142(9), 1519–1524.*,

[122] Damian, V., Logofătu, P. C., Vasile, T. (2016)., “3D THz hyperspectrum applied in security check-in”, *(December 2016)*, 100100Y.

[123] Defernez, M., Kemsley, E. K., Wilson, R. H. (1995)., “Use of infrared-spectroscopy and chemometrics for the authentication of fruit purees”, *Journal of Agricultural and Food Chemistry*, 43(1), 109–113.

[124] Federici, J. F. (2012)., “Review of moisture and liquid detection and mapping using Terahertz imaging”, *Journal of Infrared, Millimeter, and Terahertz Waves* 33(2), 97–126.

[125] Federici, J. F., Schulkin, B., Huang, F., Gary, D., Barat, R., Oliveira, F., Zimdars, D. (2005)., “THz imaging and sensing for security applications - Explosives, weapons and drugs”, *Semiconductor Science and Technology."*

[126] Federici, J., Moeller, L. (2010)., “Review of terahertz and subterahertz wireless communications”, *Journal of Applied Physics.*

[127] Fitch, M.J., Osiander, R. (2004), “Terahertz waves for communications and sensing”,*Johns Hopkins Apl Technical Digest, 25*, 348-355.

[128] Fischer, B., Hoffmann, M., Helm, H., Modjesch, G., Jepsen, P. U. (2005), “Chemical recognition in terahertz time-domain spectroscopy and imaging”, *Semiconductor Science and Technology*,

[129] Galoda, S., Singh, G. (2007)., “Fighting terrorism with terahertz”, *IEEE Potentials, 26(6)*, 24–29.

[130] Gorenflo, S., Tauer, U., Hinkov, I., Lambrecht, A., Buchner, R., Helm, H. (2006)., “Dielectric properties of oil-water complexes using terahertz transmission spectroscopy”, *Chemical Physics Letters, 421(4–6)*, 494–498.

[131] Guan, L. M., Zhan, H. L., Miao, X. Y., Zhu, J., Zhao, K. (2017). “Terahertz-dependent evaluation of water content in high-water-cut crude oil using additive-manufactured samplers”, *Science China: Physics, Mechanics and Astronomy*, 60(4), 44211.

[132] Hao, G., Liu, J., Hong, Z. (2011)., “Determination of soluble solids content in apple products by terahertz time-domain spectroscopy”, *In International Symposium on Photoelectronic Detection and Imaging 2011: Terahertz Wave Technologies and Applications*, (Vol. 8195, p. 819510).

[133] He, G., Stiens, J. (2018), “Enhanced Terahertz Absorption of Graphene Composite Integrated with Double Circular Metal Ring Array”, *Plasmonics, 13(5)*, 1705–1710.

[134] Hesler, J. L., Liu, L., Xu, H., Duan, Y., Weikle, R. M., (2008)., “The development of quasi-optical THz detectors”, *In 33rd International Conference on Infrared and Millimeter Waves and the 16th International Conference on Terahertz Electronics*, 2008, IRMMW-THz 2008.

[135] Horvat, J., Lewis, R. a. (2009)., “Peeling adhesive tape emits electromagnetic radiation at terahertz frequencies”, *Optics Letters, 34(14)*, 2195.

[136] Hosako, I., Oda, N. (2011)., “Terahertz imaging for detection or diagnosis”, *SPIE Newsroom, 10(2)*, 003651.

[137] Hua, Y., Zhang, H. (2010)., “Qualitative and quantitative detection of pesticides with terahertz time-domain spectroscopy”, *IEEE Transactions on Microwave Theory and Techniques*, 58(7 PART 2), 2064–2070.

[138] Huang, Y., Xu, W., Qin, J., Xie, L., Ying, Y. (2018). “Monitoring high-absorption aqueous solution with multiple attenuated total reflection terahertz time-domain spectroscopy”, *2018 ASABE Annual International Meeting. St. Joseph*, MI: ASABE.

[139] Ikamas, K., Nevinskas, I., Krotkus, A., Lisauskas, A. (2018). “Silicon Field Effect Transistor as the Nonlinear Detector for Terahertz Autocorellators”, *Sensors, 18(11)*, 3735.

[140] Jepsen, P. U., Cooke, D. G., Koch, M. (2011)., “Terahertz spectroscopy and imaging - Modern techniques and applications”, *Laser and Photonics Reviews, 5(1)*, 124–166.

[141] Jepsen, P. U., Møller, U., Merbold, H. (2007)., “Investigation of aqueous alcohol and sugar solutions with reflection terahertz time-domain spectroscopy”, *Optics Express, 15(22)*, 14717.

[142] Ji, J., Jiang, J., Chen, J., Du, F., Huang, P. (2018). “Scattering reduction of perfectly electric conductive cylinder by coating plasma and metamaterial”, *Optik*, 161, 98–105.

[143] Jiusheng, L. I. (2010)., “Optical parameters of vegetable oil studied by terahertz time-domain spectroscopy”, *Applied Spectroscopy, 64(2)*, 231–234.

[144] Jördens, C. (2008), “Detection of foreign bodies in chocolate with pulsed terahertz spectroscopy”, *Optical Engineering, 47(3)*, 037003.

[145] Jördens, C., Scheller, M., Breitenstein, B., Selmar, D., Koch, M. (2009)., “Evaluation of leaf water status by means of permittivity at terahertz frequencies”, *Journal of Biological Physics, 35(3)*, 255–264.

[146] Kaatze, U., & Hübner, C. (2010)., “Electromagnetic techniques for moisture content determination of materials”, *Measurement Science and Technology*.

[147] Kadomura, Y., Yamamoto, N., Tominaga, K. (2018)., “THz Dynamics of Hydrated Phospholipid Studied by Broadband Dielectric Spectroscopy”, *In 2018 43rd International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz)* (pp. 1–2).

[148] Khaled, D. El, Novas, N., Gazquez, J. A., Garcia, R. M., Manzano-Agugliaro, F. (2015)., “Fruit and vegetable quality assessment via dielectric sensing”, *Sensors* (Switzerland).

[149] Kitahara, H., Yagi, T., Mano, K. (2005). “Dielectric characteristics of water solutions of ethanol in the terahertz region”, *Journal of the Korean . . . , 46(1)*, 82–85. Retrieved from

[150] Koudou, B. (2014)., “Elimination of lymphatic filariasis: The endgame and challenges for meeting the 2020 target”, *International Journal of Infectious Diseases,*, 21, 42.

[151] Labuza, T. P., Hyman, C. R. (1998)., “Moisture migration and control in multi-domain foods”, *Trends in Food Science and Technology, 9(2)*, 47–55.

[152] Lazebnik, M., Popovic, D., McCartney, L., Watkins, C. B., Lindstrom, M. J., Harter, J., Hagness, S. C. (2007). "A large-scale study of the ultra wideband microwave dielectric

properties of normal, benign and malignant breast tissues obtained from cancer surgeries", *Physics in Medicine and Biology, 52(20)*, 6093–6115.

[153] Lee, D. K., Kang, J. H., Lee, J. S., Kim, H. S., Kim, C., Hun Kim, J., Seo, M. (2015)., "Highly sensitive and selective sugar detection by terahertz nano-antennas", *Scientific Reports*, 5.

[154] Liu, J. (2017)., "Terahertz spectroscopy and chemometric tools for rapid identification of adulterated dairy product", *Optical and Quantum Electronics*, 49(1).

[155] Liu, W., Liu, C., Chen, F., Yang, J., Zheng, L. (2016). "Discrimination of transgenic soybean seeds by terahertz spectroscopy", *Scientific Reports, 6*.

[156] Liu, Y., Ying, Y. (2005)., "Use of FT-NIR spectrometry in non-invasive measurements of internal quality of "Fuji" apples", *Postharvest Biology and Technology, 37(1)*, 65–71.

[157] Malevich, V. L., Sinitsyn, G. V., Sochilin, G. B., Rosanov, N. N. (2018)., "Manifestations of Radiation Scattering in the Method of Pulsed Terahertz Spectroscopy", *Optics and Spectroscopy, 124(6)*, 889–894.

[158] Massaouti, M., Daskalaki, C., Gorodetsky, A., Koulouklidis, A. D., Tzortzakis, S. (2013), "Detection of harmful residues in honey using terahertz time-domain spectroscopy", *Applied Spectroscopy, 67(11)*, 1264–1269.

[159] Mazin Abdul-Munaim, A., Reuter, M., Mazin Abdulmunem, O., C Balzer, J., Koch, M., G. Watson, D. (2016)., "Using Terahertz Time-Domain Spectroscopy to Discriminate among Water Contamination Levels in Diesel Engine Oil", *Transactions of the ASABE, 59(3)*, 795–801.

[160] Mickan, S. P., Shvartsman, R., Munch, J., Zhang, X. C., Abbott, D. (2004)., "Low noise laser-based T-ray spectroscopy of liquids using double-modulated differential time-domain spectroscopy", *In Journal of Optics B: Quantum and Semiclassical Optics*, (Vol. 6).

[161] Mieloszyk, M., Majewska, K., Ostachowicz, W. (2018). "Composite samples with different contaminations analysed with THz spectrometry", *In Proceedings of SPIE - The International Society for Optical Engineering (Vol. 10600*, p. 106001R–10600–12).

[162] Momivama, H., Sasaki, Y., Yosbimine, I., Nagano, S., Yuasa, T., Otani, C. (2018)., "Study of 3D imaging using a CW diode terahertz source for practical applications", *In 2018 43rd International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz)* (pp. 1–2).

[163] Naftaly, M. (2013). “Meteorology Issues and Solutions in THz Time-Domain Spectroscopy: Noise, Errors, Calibration”, *IEEE Sens.J.*, Vol. 13(1), 8-17.

[164] Naftaly, M. Miles, R. E. (2007). “Terahertz time-domain spectroscopy for material characterization”, *In Proceedings of the IEEE Vol. 95,* pp. 1658–1665.

[165] Nagai, M., Yada, H., Arikawa, T., Tanaka, K. (2006). “Terahertz time-domain attenuated total reflection spectroscopy in water and biological solution”, *International Journal of Infrared and Millimeter Waves, 27(4)*, 505–515.

[166] Nazarov, M. M., Cherkasova, O. P., Shkurinov, A. P. (2016)., “Study of the dielectric function of aqueous solutions of glucose and albumin by THz time-domain spectroscopy”, *Quantum Electronics, 46(6)*, 488–495.

[167] Nie, P., Qu, F., Lin, L., Dong, T., He, Y., Shao, Y., Zhang, Y. (2017)., “Detection of water content in rapeseed leaves using terahertz spectroscopy”, *Sensors (Switzerland)*, 17(12).

[168] Norouzian, F., Du, R., Hoare, E. G., Gardner, P., Constantinou, C., Cherniakov, M., Gashinova, M. (2018)., “Low-THz Transmission Through Water-Containing Contaminants on Antenna Radome”, *IEEE Transactions on Terahertz Science and Technology*, 8(1), 63–75.

[169] Novikova, A., Markl, D., Zeitler, J. A., Rades, T., Leopold, C. S. (2018)., “A nondestructive method for quality control of the pellet distribution within a MUPS tablet by terahertz pulsed imaging”, *European Journal of Pharmaceutical Sciences, 111*, 549–555.

[170] Ogawa, Y. (2006)., “Feasibility on the quality evaluation of agricultural products with terahertz electromagnetic wave”, *2006 ASABE Annual International Meeting Oregon Convention Center. St. Joseph*, MI: ASABE.

[171] Ok, G., Choi, S. W., Park, K. H., Chun, H. S. (2013)., “Foreign object detection by subterahertz quasi-Bessel beam imaging”, *Sensors (Switzerland), 13(1)*, 71–85.

[172] Ok, G., Park, K., Kim, H. J., Chun, H. S., Choi, S.-W. (2014)., "High-speed terahertz imaging toward food quality inspection”, *Applied Optics, 53(7)*, 1406.

[173] Ok, G., Shin, H. J., Lim, M.-C., Choi, S.-W. (2018)., “Large-scan-area sub-terahertz imaging system for nondestructive food quality inspection”, *Food Control, 96*, September 2018, 383–389.

[174] Parasoglou, P., Parrott, E. P. J., Zeitler, J. A., Rasburn, J., Powell, H., Gladden, L. F., Johns, M. L. (2009)., “Quantitative moisture content detection in food wafers”, *In 34th International Conference on Infrared, Millimeter, and Terahertz Waves, IRMMW-THz 2009*.

[175] Parasoglou, P., Parrott, E. P. J., Zeitler, J. A., Rasburn, J., Powell, H., Gladden, L. F., Johns, M. L. (2010)., "Quantitative water content measurements in food wafers using terahertz radiation", *IEEE Trans. THz Sci. Technol., 3(4)*, 172–182.

[176] Park, S. J., Hong, J. T., Choi, S. J., Kim, H. S., Park, W. K., Han, S. T., Ahn, Y. H. (2014)., "Detection of microorganisms using terahertz metamaterials", *Scientific Reports, 4*.

[177] Qu, F., Lin, L., Cai, C., Dong, T., He, Y., & Nie, P. (2018)., "Molecular Characterization and Theoretical Calculation of Plant Growth Regulators Based on Terahertz Time-Domain Spectroscopy", *Applied Sciences, 8(3)*, 420.

[178] Qu, F., Nie, P., Lin, L., Cai, C., & He, Y. (2018), "Review of theoretical methods and research aspects for detecting leaf water content using terahertz spectroscopy and imaging", *International Journal of Agricultural and Biological Engineering, 11(5)*, 27–34.

[179] Razeghi, M., Quanyong, L. U., Manna, S., Donghai, W. U., & Slivken, S. (2016)., "Breakthroughs bring THz spectroscopy, sensing closer to mainstream", *Photonics Spectra* (Vol. 50).

[180] Redo-Sanchez, A., Salvatella, G., Galceran, R., Roldós, E., García-Reguero, J. A., Castellari, M., & Tejada, J. (2011). "Assessment of terahertz spectroscopy to detect antibiotic residues in food and feed matrices", *Analyst, 136(8)*, 1733–1738.

[181] Ryu, W., Won, Y. H., & Lee, M. (2018). "Compact and portable terahertz imaging system using single-mode Fabry-Perot laser diode photomixer", *In Terahertz, RF, Millimeter, and Submillimeter-Wave Technology and Applications XI* (p. 52).

[182] Sizov, F., & Rogalski, A. (2010)., "THz detectors", *Progress in Quantum Electronics.*

[183] Stübling, E., Bauckhage, Y., Jelli, E., Fischer, B., Globisch, B., Schell, M., & Koch, M. (2017)., "A THz Tomography System for Arbitrarily Shaped Samples", *Journal of Infrared, Millimeter, and Terahertz Waves*.

[184] Sun, W., Wang, X., & Zhang, Y. (2012)., "A method to monitor the oil pollution in water with reflective pulsed terahertz tomography", *Optik, 123(21)*, 1980–1984.

[185] Takehara, D., Endo, M., Ishibashi, T., Shimizu, M., Kusanagi, S., Nozokido, T., & Bae, J. (2018)., "Dual-Polarization Imaging with Real-Time Capability Using a Terahertz Noise Source for Food Inspection", *In 2018 43rd International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz)*, (pp. 5–6).

[186] Turan, D., García, S. C. C., Castro-Camus, E., & Jarrahi, M. (2017)., "Terahertz power enhancement by improving metal adhesion layer of plasmonic photoconductive sources", *In 2017 Conference on Lasers and Electro-Optics (CLEO)*, (pp. 1–2).

[187] Venables, D. S., & Schmuttenmaer, C. A. (2000)., "Spectroscopy and dynamics of mixtures of water with acetone, acetonitrile, and methanol", *Journal of Chemical Physics, 113(24)*, 11222–11236.

[188] Wang, Y., Wang, Q., Zhao, Z., Liu, A., Tian, Y., & Qin, J. (2018)., "Rapid qualitative and quantitative analysis of chlortetracycline hydrochloride and tetracycline hydrochloride in environmental samples based on terahertz frequency-domain spectroscopy", *Talanta, 190*, 284–291.

[189] Wichmann, M., Scherger, B., Lippert, S., Kocic, N., Wietzke, S., Vieweg, N., ... Koch, M. (2011). "Terahertz lenses made by compression molding. IRMMW-THz 2011", *36th International Conference on Infrared, Millimeter, and Terahertz Waves, 50(15)*, 2256–2262.

[190] Xu, W., Xie, L., Ye, Z., Gao, W., Yao, Y., Chen, M., Ying, Y. (2015)., "Discrimination of Transgenic Rice containing the Cry1Ab Protein using Terahertz Spectroscopy and Chemometrics", *Scientific Reports*, 5.

[191] Xu, Y. (2006). "DNA damage: A trigger of innate immunity but a requirement for adaptive immune homeostasis", *Nature Reviews Immunology, 6(4)*, 261–270.

[192] Yang, K., Chopra, N., Abbasi, Q. H., Qaraqe, K. A., & Alomainy, A. (2017)., "Collagen Analysis at Terahertz Band Using Double-Debye Parameter Extraction and Particle Swarm Optimisation", *IEEE Access, 5*, 27850–27856.

[193] Yilmaz, T., Foster, R., & Hao, Y. (2014). "Towards accurate dielectric property retrieval of biological tissues for blood glucose monitoring", *IEEE Transactions on Microwave Theory and Techniques, 62(12)*, 3193–3204.

[194] Yu, X., Endo, M., Ishibashi, T., Shimizu, M., Kusanagi, S., Nozokido, T., & Bae, J. (2016). "Orthogonally polarized terahertz wave imaging with real-time capability for food inspection", *In Asia-Pacific Microwave Conference Proceedings*, APMC (Vol. 2).

[195] Zahid, A., Yang, K., Heidari, H., Li, C., Imran, M. A., Alomainy, A., & Abbasi, Q. H. (2018). "Terahertz characterisation of living plant leaves for quality of life assessment applications", *In URSI 2018 - Baltic URSI Symposium*. (pp. 117–120).

[196] Zhang, X., Lu, S., Liao, Y., & Zhang, Z. (2017). "Simultaneous determination of amino acid mixtures in cereal by using terahertz time domain spectroscopy and chemometrics", *Chemometrics and Intelligent Laboratory Systems, 164*, 8–15.

[197] Zulkifli, S. N., Rahim, H. A., & Lau, W. J. (2018), "Detection of contaminants in water supply: A review on state-of-the-art monitoring technologies and their applications", *Sensors and Actuators, B: Chemical*,

[198] M. Kah, N. Tufenkji and J. White, "Nano-enabled strategies to enhance crop nutrition and protection", *Nature Nanotechnology, vol. 14, no. 6*, pp. 532-540, 2019.

[199] E. Lombi, et al., "A One Health approach to managing the applications and implications of nanotechnologies in agriculture", *Nature Nanotechnology, vol. 14, no. 6*, pp. 523-531, 2019.

[200] G. Lowry, A. Avellan and L. Gilbertson, "Opportunities and challenges for nanotechnology in the agri-tech revolution",*Nature Nanotechnology, vol. 14, no. 6*, pp. 517-522, 2019.

[201] M. Zhang, et al., "Detection of stress in tomatoes induced by late blight disease in California, USA, using hyperspectral remote sensing", *Int. J. Appl. Earth Obs. Geoinf. 4*, 295–310, 2003.

[202] Nietzel, T. et al. "The fluorescent protein sensor roGFP2-Orp1 monitors in vivo H2 O2 and thiol redox integration and elucidates intracellular H2O2 dynamics during elicitor-induced oxidative burst in Arabidopsis", *New Phytol. 221, 1649–1664*, 2019.

[203] N. Hassan, C. Chou and M. Hassan, "eNEUTRAL IoNT: Energy-Neutral Event Monitoring for Internet of Nano Things", *IEEE Internet of Things Journal, vol. 6*, no. 2, pp. 2379-2389, 2019.

[204] Y. Li, Y. Wang, Q. Cao, J. Cao and D. Qiao, "A Self-Powered Vibration Sensor With Wide Bandwidth", *IEEE Transactions on Industrial Electronics, vol. 67, no. 1*, pp. 560-568, 2020.

[205] Borgatta, J. et al. "Copper based nanomaterials suppress root fungal disease in watermelon (Citrullus lanatus): role of particle morphology, composition and dissolution behavior", *ACS Sustain. Chem. Eng. 6, 14847–14856*, 2018.

[206] Zhu, Q. et al. "FRET-based glucose imaging identifies glucose signalling in response to biotic and abiotic stresses in rice roots", *J. Plant Physiol. 215, 65–72*, 2017.

[207] Alomainy, A., et al. "Nano-Electromagnetic Communication at Terahertz and Optical Frequencies: Principles and Applications", Institution of Engineering and Technology", 2019.

[208] Richardson, J. J. & Liang, K. Nano-biohybrids: "in vivo synthesis of metal-organic frameworks inside living plants", *Small 14*, 2018.

[209] Wu, H., Tito, N. & Giraldo, J. P. "Anionic cerium oxide nano-particles protect plant photosynthesis from abiotic stress by scavenging reactive oxygen species". *ACS Nano 11, 11283–11297*, 2017.

[210] Afsharinejad, A.; Davy, A.; O'Leary, P.; Brenann, C., "Transmission through Single and Multiple Layers of Plant Leaves at THz Frequencies", *In Proceedings of the GLOBECOM 2017—2017 IEEE Global Communications Conference, Singapore*, 4–8 December 2017; pp. 1–6.

[211] Woodward, R.M.; Cole, B.E.; Wallace, V.P.; Pye, R.J.; Arnone, D.D.; Linfield, E.H.; Pepper, M., "Terahertz pulse imaging in reflection geometry of human skin cancer and skin tissue", *Phys. Med. Biology 2002*,47, 3853

[212] Smith, R.L., "Diagnostic Modalities", *Clin. Orthop. Relat. Res.2004*, 427, S174

[213] Sun, Q.; He, Y.; Liu, K.; Fan, S.; Parrott, E.P.J.; Pickwell-MacPherson, E., "Recent advances in terahertz technology for biomedical applications", *Quant. Imaging Med. Surg.2017*,7, 345–355.

[214] Naftaly, M.; Miles, R.E., "Terahertz Time-Domain Spectroscopy for Material Characterization", *Proc. IEEE 2007*,95, 1658–1665.

[215] Kürner, T.; Priebe, S., "Towards THz Communications - Status in Research, Standardization and Regulation", *J. Infrared Millim. Terahertz Waves 2014*,35, 53–62.

[216] Afsharinejad, A.; Davy, A.; Naftaly, M., "Variability of Terahertz Transmission Measured in Live Plant Leaves", *IEEE Geosci. Remote Sens. Lett.2017* ,14, 636–638,

[217] Santesteban, L.G.; Palacios, I.; Miranda, C.; Iriarte, J.C.; Royo, J.B.; Gonzalo, R., "Terahertz time domain spectroscopy allows contactless monitoring of grapevine water status", *Front.Plant Sci.2015,* 6, 1–9,

[218] Gente, R.; Born, N.; Velauthapillai, A.; Balzer, J.C.; Koch, M., "Monitoring the water content of plant leaves with THz time domain spectroscopy", *In Proceedings of the 2015 40th International Conference on Infrared, Millimeter, and Terahertz waves (IRMMW-THz), Hong Kong, China*, 23–28 August 2015; pp. 1–2,

[219] Torres, V.; Palacios, I.; Iriarte, J.C.; Liberal, I.; Santesteban, L.G.; Miranda, C.; Royo, J.B.; Gonzalo, R., "Monitoring Water Status of Grapevine by Means of THz Waves", *J. Infrared Millim. Terahertz Waves 2016*,37, 507–513,

[220] Jördens, C.; Scheller, M.; Breitenstein, B.; Selmar, D.; Koch, M., "Evaluation of leaf water status by means of permittivity at terahertz frequencies", *J. Biol. Phys.2009*,35, 255–264.14.

[221] Gente, R.; Koch, M., "Monitoring leaf water content with THz and sub-THz waves", *Plant Methods 2015*,11, 15,

[222] Born, N.; Behringer, D.; Liepelt, S.; Beyer, S.; Schwerdtfeger, M.; Ziegenhagen, B.; Koch, M., "Monitoring Plant Drought Stress Response Using Terahertz Time-Domain Spectroscopy", *Plant Physiol.2014*,164, 1571–1577,

[223] Breitenstein, B.; Scheller, M.; Shakfa, M.K.; Kinder, T.; Müller-Wirts, T.; Koch, M.; Selmar, D., "Introducing terahertz technology into plant biology: A novel method to monitor changes in leaf water status", *J. Appl.Botany Food Quality 2011*,84, 158–161.

[224] Higa, S.; Kobori, H.; Tsuchikawa, S., "Mapping of Leaf Water Content Using Near-Infrared Hyper-spectral Imaging", *Appl. Spectrosc.2013*,67, 1302–1307.

[225] Song, Z.; Yan, S.; Zang, Z.; Fu, Y.; Wei, D.; Cui, H.L.; Lai, P., "Temporal and spatial variability of water status in plant leaves by Terahertz imaging", *IEEE Trans. Terahertz Sci. Technol.2018*,8, 520–527,

[226] Sardans, J.; Peñuelas, J.; Lope-Piedrafita, S., "Changes in water content and distribution in Quercus ilex leaves during progressive drought assessed by in vivo 1H magnetic resonance imaging", *BMC Plant Biol.2010* ,10, 1–12,

[227] Sarracanie, M.; Lapierre, C.D.; Salameh, N.; Waddington, D.E.J.; Witzel, T.; Rosen, M.S., "Low-Cost High-Performance MRI", *Sci. Rep.2015*,5, 1–9,

[228] Saha, S.C.; Grant, J.P.; Ma, Y.; Khalid, A.; Hong, F.; Cumming, D.R.S., "Terahertz Frequency-Domain Spectroscopy Method for Vector Characterization of Liquid Using an Artificial Dielectric", *IEEE Trans.Terahertz Sci. Technol.2012*,2, 113–122,

[229] Hadjiloucas, S.; Karatzas, L.S.; Bowen, J.W., "Measurements of leaf water content using terahertz radiation", *IEEE Trans. Microw. Theory Tech.1999* ,47, 142–149,

[230] De Cumis, U.S.; Xu, J.H.; Masini, L.; Degl'Innocenti, R.; Pingue, P.; Beltram, F.; Tredicucci, A.; Vitiello, M.S.;Benedetti, P.A.; Beere, H.E.; et al., "Terahertz confocal microscopy with a quantum cascade laser source", *Opt.Express 2012*,20, 21924–21931,

[231] Nicolson, A.M.; Ross, G.F., "Measurement of the Intrinsic Properties of Materials by Time-Domain Techniques", *IEEE Trans. Instrum. Meas.1970*,19, 377–382,

[232] Weir, W.B., "Automatic Measurement of Complex Dielectric Constant and Permeability", *Proc. IEEE 1974*,62, 33–36,

[233] Baker-Jarvis, J.; Vanzura, E.J.; Kissick, W.A., "Improved technique for determining complex permittivity with the transmission/reflection method", *IEEE Trans. Microw.Theory Tech.1990*,38, 1096–1103,

[234] Costa, F.; Borgese, M.; Degiorgi, M.; Monorchio, A., "Electromagnetic Characterisation of Materials by Using Transmission/Reflection (T/R) Devices", *Electronics 2017*,6, 95,

[235] Luukkonen, O.; Maslovski, S.I.; Tretyakov, S.A., "A Step wise Nicolson–Ross–Weir-Based Material Parameter Extraction Method", *IEEE Antennas Wirel. Propag. Lett.2011*,10, 1295–1298,

[236] Boughriet, A.H.; Legrand, C.; Chapoton, A., "Non-iterative stable transmission/reflection method for low-loss material complex permittivity determination", *IEEE Trans. Microw. Theory Tech.1997*,45, 52–57,

[237] Kim, S.; Guerrieri, J.R., "Low-Loss Complex Permittivity and Permeability Determination in Transmission/Reflection Measurements With Time-Domain Smoothing", *Prog. Electromagn. Res. M2015*,44, 69–79,

[238] Cao, Z.; Wang, Q.; Zheng, C. Best, "hyper spectral indices for tracing leaf water status as determined from leaf dehydration experiments", *Ecol. Indic.2015*, 54, 96–107,

[239] Khaled DE, Novas N, Gazquez JA, Garcia RM, Manzano-Agugliaro F, "Fruit and vegetable quality assessment via dielectric sensing", *Sensors 2015*; 15:15363–97.

[240] Zahid A, Abbas HT, Sheikh F, Kaiser T, Zoha A, Imran MA, Abbasi QH, "Monitoring health status and quality assessment of leaves using terahertz frequency", *In: 2019 IEEE symposium on antennas and propagation and USNC-URSI radio science meeting, Atlanta, GA, USA*, 07–12 July 2019.

[241] El Beyrouthya M., "Nanotechnologies: novel solutions for sustainable agriculture", *Adv Crop Sci Technol*. 2014;02(03):8863.

[242] Zahid A, Abbas HT, Imran MA, Qaraqe KA, Alomainy A, Cumming DR,Abbasi QH., "Characterization and water content estimation method of living plant leaves using terahertz waves", *Appl Sci. 2019*, ;9:2781.

[243] Afsharinejad A, Davy A, Naftaly M., "Variability of terahertz transmission measured in live plant leaves", *IEEE Geosci Remote Sens Lett.2017*, 14(5):636–8.

[244] Born N, et al. "Monitoring plant drought stress response using terahertz time-domain spectroscopy", *Plant Physiol. 2014*, 164(4):1571–7.

[245] Naftaly M, Miles RE., "Terahertz time-domain spectroscopy for material characterization", *Proc IEEE. 2007*;95(8):1658–65.

[246] Jordens C, Scheller M, Breitenstein B, Selmar D, Koch M. "Evaluation of leaf water status by means of permittivity at terahertz frequencies", *J Biol Phys*. 2009;35(3):255–64.

[247] Liew OW, Chong PCJ, Li B, Asundi AK. “Signature optical cues: emerging technologies for monitoring plant health”, *Sensors. 2008;8*, :3205–39.

[248] Torres V, Palacios I, Iriarte JC, Liberal I, Santesteban LG, Miranda C, Royo JB, Gonzalo R., “Monitoring water status of grapevine by means of THz waves”, *J Infrared Millim Terahertz Waves*, 2016;37(5):507–13.

[249] Santesteban LG, Palacios I, Miranda C, Iriarte JC, Royo JB, Gonzalo R., “Terahertz time domain spectroscopy allows contactless monitoring of grapevine water status”, *Front Plant Sci. 2015;* 6(June):1–9.

[250] Song Z, et al, “Temporal and spatial variability of water status in plant leaves by Terahertz imaging”, *IEEE Trans Terahertz Sci Technol*, 2018;8(5):520–7.

[251] Gente R, Born N, Velauthapillai A, Balzer JC, Koch M., “Monitoring the water content of plant leaves with thz time domain spectroscopy”, *In: 2015 40th international conference on infrared, millimeter, and terahertz waves (IRMMW-THz)*, Aug 2015. p. 1, 2.

[252] Gente R, Koch M., “Monitoring leaf water content with thz and sub-thz waves”, *Plant Methods.*, 2015;11(1):15.

[253] Breitenstein B, Scheller M, Shakfa MK, Kinder T, Müller-Wirts T, Koch M, Selmar D., “Introducing terahertz technology into plant biology: a novel method to monitor changes status”, *J Appl Bot Food Qual.2011*, ;84(2):158–61.

[254] Nie P, Qu F, Lin L, Dong T, He Y, Shao Y, Zhang Y. “Detection of water content in rapeseed leaves using terahertz spectroscopy”, *Sensors.2017;* 17(12):2830.

[255] Saha SC, Grant JP, Ma Y, Khalid A, Hong F, Cumming DRS., “Terahertz frequency-domain spectroscopy method for vector characterization of liquid using an artificial dielectric”, *IEEE Trans Terahertz Sci Technol.2012*,2(1):113–22.

[256] Sun Q, He Y, Liu K, Fan S, Parrott EPJ, Pickwell-MacPherson E., “Recent advances in terahertz technology for biomedical applications”, *Quant Imaging Med Surg.*, 2017;7(3):345.

[257] Ren A, Zahid A, Fan D, Yang X, Alomainy A, Imran MA, Abbasi QH., “State-of-the-art terahertz sensing for food and water security—a comprehensive review”, *Trends Food Sci Technol.2019* ;85:241–51. https://doi. org/10.1016/j.tifs.2019.01.019.

[258] Zahid A, Abbas HT, Heidari H, Imran MA, Alomainy A, Abbasi QH., “Monitoring the variability of water dynamics in plant leaves at cellular level using terahertz sensing”, *In: 2019 second international workshop on mobile terahertz systems (IWMTS),* 1–3 July 2019, Bad Neuenahr,Germany.

[259] Liakos KG, Busato P, Moshou D, Pearson S, Bochtis D., "Machine learning in agriculture: a review", *Sensors. 2018* ;18:2674.

[260] Cao Z, Wang Q, Zheng C. Best, "Hyper-spectral indices for tracing leaf water status as determined from leaf dehydration experiments", *Ecol Indic. 2015*, ;54:96–107.

[261] Mittleman DM, Jacobsen RH, Nuss MC., "T-ray Imaging", *IEEE J Sel Top Quantum Electron* 1996;2(3):679–92.

[262] Li H, Yuan D, Wang Y, Cui D, Cao L. Arrhythmia., "classification based on multi-domain feature extraction for an ECG recognition system", *Sensors.* 2016;16(10):1744.

[263] Von Storch H, Zwiers FW., "Statistical analysis in climate research", *Cambridge: Cambridge University Press*; 2001. ISBN 978-0-521-01230-0.

[264] Chen Hua, Chen Xiaofeng, Ma Shihua, Xiumei Wu, Yang Wenxing, Zhang Weifeng, Li Xiao., "Quantify glucose level in freshly diabetic's blood by terahertz time-domain spectroscopy.", *J Infrared Millim Terahertz Waves*. 2018;39(4):399–408.

[265] Yin X, Hadjiloucas S, Zhang Y., "Classification of THz pulse signals using two-dimensional cross-correlation feature extraction and non-linear classifers.", *Comput Methods Programs Biomed*. 2016;127:64–82.

[266] Dutta S, Chatterjee A, Munshi S., "Correlation techniques and least square support vector machine combine for frequency domain-based ECG beat classification", *Med Eng Phys 2010*, 32:1161–9.

[267] Berry E, Boyle RD, Fitzgerald AJ, Handley JW, "Time frequency analysis in terahertz pulsed imaging", *In: Bhanu B, Pavlidis I, editors. Computer vision: beyond the visible spectrum. Advances in pattern recognition.London:, Springer*; 2005. p. 290–329.

[268] Yin X, Ng BW, Abbott D, Ferguson B, Hadjiloucas S., "Application of auto regressive models of wavelet sub-bands for classifying terahertz pulse measurements.", *J Biol Syst. 2007* ;15(4):551–71.

[269] Haddadi R, Abdelmounim E, El Hanine M, Belaguid A, "Discrete wavelet transform based algorithm for recognition of QRS complexes", *In: 2014 international conference on multimedia computing and systems (ICMCS), Marrakech,* 2014, p. 375–9.

[270] Li Congcong, Wang Jie, Wang Lei, Luanyun Hu, Gong Peng, "Comparison of classification algorithms and training sample sizes in urban land classification with landsat thematic mapper imagery", *Remote Sens. 2014;* 6:964–83.

[271] Thanh Noi P, Kappas M., "Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery", *Sensors. 2018*, ;18(1):18.

[272] Gürbüz SZ, Erol B, Çağlıyan B, Tekeli B. "Operational assessment and adaptive selection of micro-Doppler features", *IET Radar Sonar Navig. 2015;* 9(9):1196–204.

[273] Pohjalainen J, Räsänen O, Kadioglu S., "Feature selection methods and their combinations in high-dimensional classification of speaker likability. Intelligibility Personal Traits Comput Speech Lang", *2015;29(1)*:145–71.

[274] Feizi-Derakhshi MR, Ghaemi M., "Classifying different feature selection algorithms based on the search strategies", *In: International conference on machine learning, electrical and mechanical engineering. 2014,* p.17–21.

[275] Urbanowicz RJ, Olson RS, Schmitt P, Meeker M, Moore JH., "Bench marking relief-based feature selection methods for bioinformatics data mining.2017", *arXiv preprint arXiv:1711* .08477.

[276] Perez-Castano Estefanıa, Ruiz-Samblas Cristina, Medina-Rodrıguez Santiago, Quiros-Rodrıguez Veronica, Jimenez-Carvelo Ana M, Valverde-Som Lucıa, Gonzalez-Casado Antonio, Cuadros-Rodrıguez Luis.,"Comparison of different analytical classification scenarios: application for the geographical origin of edible palm oil by sterolic (NP) HPLC fnger printing", *Anal Methods. 2015;*, 7:4192–201

[277] O. W. Liew, P. Chong, B. Li, and A. K. Asundi, "Signature optical cues: emerging technologies for monitoring plant health.", *Sensors, vol. 8, no. 5,* pp. 3205-3239, 2008.

[278] H. Wang, J. Peng, C. Xie, Y. Bao, and Y. He, "Fruit Quality Evaluation Using Spectroscopy Technology: A Review," *Sensors, vol. 15, no. 1*, pp. 11889-11927, 2015.

[279] M. Tonouchi, "Cutting-edge THz technology," Nature Photonics, vol. 1, pp. 97-105, 2007.

[280] P. Butz, C. Hofmann, and B. Tauscher, "Recent Developments in Noninvasive Techniques for Fresh Fruit and Vegetable Internal Quality Analysis," *Journal of food science, 79(9)*, R131-R141, 2005.

[281] B., Ferguson and X. Zhang, "Materials for terahertz science and technology," *Nature Materials, vol. 1,* pp. 26-33, 2002.

[282] K.G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine Learning in Agriculture: A Review," *Sensors, vol. 18, no. 8,* pp. 2674, 2018.

[283] A. Ren, A. Zahid, D. Fan, X. Yang, M. Imran, A. Alomainy, and Q. Abbasi, “State-of-the-art terahertz sensing for food and water security – a comprehensive review,” *Trends in Food Science and Technology, vol. 85,* pp. 241-251, 2019.

[284] A. Zahid, H. T. Abbas, M. A. Imran, K. Qaraqe, A. Alomainy, D. R. S. Cumming, Q. H. Abbasi, "Characterization and Water Content Estimation Method of Living Plant Leaves Using Terahertz Waves", *Applied Sciences, vol. 9,* no. 14, 2781, 2019.

[285] A. Zahid, H. Heidari, Chong Li, M. A. Imran, A. Alomainy, and Q. H. Abbasi, “Terahertz characterisation of living plant leaves for quality of life assessment applications,” *in Baltic URSI Symposium (URSI), Poznan, 2018,* pp. 117-120.

[286] A. Khalid, D. Cumming, R. Clarke, C. Li, and N. Ridler, “Evaluation of a VNA-based material characterization kit at frequencies from 0.75 THz to 1.1 THz,” *in Proceedings of IEEE 9th UK-Europe-China Workshop on Millimetre Waves and Terahertz Technologies, 2016,* pp. 31-34.

[287] P. Nie, F. Qu, L. Lin, T. Dong, Y. He, Y. Shao, and Y. Zhang, “Detection of water content in rapeseed leaves using terahertz spectroscopy,” *Sensors, vol. 17,* no. 12, pp. 2830, 2017.

[288] X. Yin, B. Ng, and D. Abbott, “Terahertz Imaging for Biomedical Applications: pattern recognition and tomographic reconstruction,” *Springer Science & Business Media, 2012,* pp. 65-90.

[289] H. Li, D. Yuan, Y. Wang, D. Cui, and L. Cao, “Arrhythmia Classification Based on Multi-Domain Feature Extraction for an ECG Recognition System,” *Sensors, vol. 16, no. 10,* pp. 1744, 20 Oct. 2016.

[290] E. Berry, A. J. Fitzgerald, N. N. Zinov’ev, G. C. Walker, S. H.-Vanniasinkam, C. D. Sudworth, R. E. Miles, J. M. Chamberlain, and M. A. Smith, “Optical properties of tissue measured using terahertz-pulsed imaging,” *Proc. SPIE 5030, Medical Imaging 2003: Physics of Medical Imaging, (5 June 2003);* doi: 10.1117/12.479993.

[291] S. H. Von and F. W Zwiers, “Statistical analysis in climate research,” *Cambridge University Press, 2001.* ISBN 978-0-521-01230-0.

[292] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, “Feature reduction and selection for EMG signal classification,” *Expert Systems with Applications, vol. 39,* pp. 7420-7431, 2012.

[293] Javaid, H. Ali, N. Rashid, M. Tiwana, and M. W. Anwar, “Comparative Analysis of EMG Signal Features in Time-domain and Frequency-domain using MYO Gesture Control,” *in Proceedings of the 2018 4th International Conference on Mechatronics and Robotics Engineering, ACM, 2018,* pp. 157-162.

[294] H. Chen, X. Chen, S. Ma, X. Wu, W. Yang, W. Zhang, and X. Li, “Quantify Glucose Level in Freshly Diabetic’s Blood by Terahertz Time-Domain Spectroscopy,” *Journal of Infrared, Millimeter, and Terahertz Waves, vol 39, no. 4,* pp. 399-408, 2018.

[295] Siuly, X. Yin, S. Hadjiloucas, and Y. Zhang, “Classification of THz pulse signals using two-dimensional cross-correlation feature extraction and non-linear classifiers,” *Computer Methods, and Programs in Biomedicine, vol. 127,* pp. 64-82, April 2016.

[296] E. Berry, et al., “Time frequency analysis in terahertz pulsed imaging,” *In: Bhanu, B. and Pavlidis, I., (eds.) Computer Vision: Beyond the Visible Spectrum. Advances in Pattern Recognition. Springer Verlag,* London, UK , 2005, pp. 290-329.

[297] X. Yin, B. Ng, and D. Abbott, “Application of auto regressive models of wavelet subbands for classifying terahertz pulse measurements,” *Journal of Biological Systems, vol. 15, no. 4*, pp. 551-571, 2007.

[298] R. Haddadi, E. Abdelmounim, M. El Hanine, and A. Belaguid, “Discrete Wavelet Transform based algorithm for recognition of QRS complexes,” *in International Conference on Multimedia Computing and Systems (ICMCS), Marrakech, 2014,* pp. 375-379.

[299] T. Noi, Phan, and M. Kappas, “Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery,” *Sensors, vol. 18, no. 1, pp. 18, 2018.*

[300] L. Hu, M. Huang, S. Ke, and C. Tsai, “The distance function effect on k-nearest neighbor classification for medical datasets,” *Springerplus, vol. 5, no. 1,* pp. 1304, 2016.

[301] S. Z. Gürbüz, B. Erol, B. Çağlıyan, and B. Tekeli, “Operational assessment and adaptive selection of micro-Doppler features,” *IET Radar, Sonar & Navigation, vol. 9, no, 9,* pp. 1196-1204, 2015.

[302] J. Pohjalainen, O. Räsänen, and S. Kadioglu, “Feature Selection Methods and Their Combinations in High-Dimensional Classification of Speaker Likability,” *Intelligibility and Personality Traits, Computer Speech and Language, vol. 29, no. 1,* pp. 145-171, 2015

[303] R. J. Urbanowicz, et al., “Benchmarking Relief-Based Feature Selection Methods for Bioinformatics Data Mining,” *Journal of Biomedical Informatics,* pp. 168-188, 2018.

[304] Estefanıa Perez-Castano, et al., “Comparison of different analytical classification scenarios: application for the geographical origin of edible palm oil by sterolic (NP) HPLC fingerprinting,” *Analytical Methods, vol. 7,* pp. 4192-4201, 2015.

[305] Alvarez, P., Chan, C., Elimelech, M., Halas, N. & Villagrán, D., “Emerging opportunities for nanotechnology to enhance water security”, *Nat. Nanotechnol. 13*, 634–641. (2018).

[306] Zulkifli, S. N., Rahim, H. A. & Lau, W.-J., “Detection of contaminants in water supply: A review on state-of-the-art monitoring technologies and their applications”, *Sensors Actuators B: Chem. 255, 2657* – 2689. (2018).

[307] Shalit, A., Ahmed, S., Savolainen, J. & Hamm, P., “Terahertz echoes reveal the in homogeneity of aqueous salt solutions”, *Nat. Chem. 9*, 273–278, (2017).

[308] Najah Ahmed, A. et al. “Machine learning methods for better water quality prediction”,(2019). *J. Hydrol*, 578. (2019).

[309] Son, J.-H. “Terahertz electromagnetic interactions with biological matter and their applications”, *J. Appl. Phys. 105*, 102033, (2009).

[310] Song, C. et al., “Terahertz and infrared characteristic absorption spectra of aqueous glucose and fructose solutions”, *Sci. Reports 8*, 2–9. (2018).

[311] Liakos, K. G., Busato, P., Moshou, D., Pearson, S. & Bochtis, D., “Machine learning in agriculture: A review.”, *Sensors 18*, (2018).

[312] Zahid, A. et al, “Characterization and water content estimation method of living plant leaves using terahertz waves”, *Appl. Sci.* 9, (2019).

[313] Zahid, A. et al, “Machine learning driven non-invasive approach of water content estimation in living plant leaves using terahertz waves”. *Plant Methods 15*, 1–13, (2019).

[314] Woodward, R. M. et al, “Terahertz pulse imaging in reflection geometry of human skin cancer and skin tissue”, *Phys. Medicine Biol. 47*, 3853–3863, (2002).

[315] Yin, L., Chen, F., Zhang, Q. & Ma, X. Arrhythmia, “Classification based on multi-domain feature extraction”, *J. Physics: Conf. Ser. 1237, 022062*, (2019).

[316] Chen, H. et al, “Quantify Glucose Level in Freshly Diabetic’s Blood by Terahertz Time-Domain Spectroscopy” *J. Infrared, Millimeter, Terahertz Waves 39, 399–408*,(2018).

[317] Siuly, Yin, X., Hadjiloucas, S. & Zhang, Y, “Classification of thz pulse signals using two-dimensional cross-correlation feature extraction and non-linear classifiers”, *Comput. Methods Programs Biomed. 127, 64 – 82*, (2016).

[318] Dutta, S., Chatterjee, A. & Munshi, S, “Correlation technique and least square support vector machine combine for frequency domain based ecg beat classification”, *Med. engineering & physics 32, 1161—1169*, (2010).

[319] Thanh Noi, P. & Kappas, M, “Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using sentinel-2 imagery”, *Sensors 18,* (2018).

[320] Hao, J. & Ho, T. K, "Machine learning made easy: A review of scikit-learn package in python programming language", *J. Educ. Behav. Stat. 44, 348–361*,(2019).

[321] Saçlı, B. et al, "Microwave dielectric property based classification of renal calculi: Application of a knn algorithm", *Comput. Biol Medicine 112, 103366,* (2019).

[322] Hu, L. Y., Huang, M. W., Ke, S. W. & Tsai, C. F., "The distance function effect on k-nearest neighbor classification for medical datasets", *Springer Plus 5,* (2016).

[323] Li, K., Gu, Y., Zhang, P., An, W. & Li, W., "Research on knn algorithm in malicious pdf files classification under adversarial environment", *In Proceedings of the 2019 4th International Conference on Big Data and Computing, ICBDC 2019, 156–159*, (2019).

[324] Shaikhina, T. et al., "Decision tree and random forest models for outcome prediction in antibody incompatible kidney transplantation", *Biomed. Signal Process. Control. 52, 456 – 462,* (2019).

[325] Pohjalainen, J., Räsänen, O. & Kadioglu, S., "Feature selection methods and their combinations in high-dimensional classification of speaker likability, intelligibility and personality traits", *Comput. Speech & Lang. 29, 145 – 171,* (2015).

[326] Urbanowicz, R. J., Olson, R. S., Schmitt, P., Meeker, M. & Moore, J. H., "Benchmarking relief-based feature selection methods for bioinformatics data mining", *J. Biomed. Informatics 85, 168 – 188*, (2018).

[327] Gibson, R. M., Amira, A., Ramzan, N., de-la Higuera, P. C. & Pervez, Z., "Multiple comparator classifier framework for accelerometer-based fall detection and diagnostic", *Appl. Soft Comput. 39, 94 – 103*, (2016).

[328] Saha, S. C. et al., "Terahertz frequency-domain spectroscopy method for vector characterization of liquid using an artificial dielectric", *IEEE Transactions on Terahertz Sci. Technol. 2, 113–122,* (2012).

[329] Saha, S. C., "Application of terahertz spectroscopy to the characterization of biological samples using birefringence silicon grating", *J. Biomed. Opt. 17, 067006*, (2012).

[330] Ren, A. et al., "Machine learning driven approach towards the quality assessment of fresh fruits using non-invasive sensing", *IEEE Sensors J.* 20, 2075–2083 (2020).

[331] Ren, A. et al., "Terahertz sensing for fruit spoilage monitoring", *In 2019 Second International Workshop on Mobile Terahertz Systems (IWMTS), 1–4* (2019).