Dual Radio Frequency Characterization with Machine Learning for Rice Moisture Content Classification and Localization

by

Noraini Azmi
(1740212371)

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

Faculty of Electronic Engineering Technology
UNIVERSITI MALAYSIA PERLIS

2022

ACKNOWLEDGEMENT

First and foremost, I am deeply grateful to Allah for his blessing towards the completion of this thesis despite all the challenges and the outbreak of the Corona Virus (covid-19) pandemic from 2019 until recently. With all hope and prayer, hopefully, the covid-19 will cease to disappear, and we can resume our life as or even better than before.

I would like to offer my gratitude to my supervisor, Assoc. Prof. Dr. Latifah Munirah Kamarudin and co-supervisors for their continuous and relentless support and guidance throughout my study. Their inputs and feedback were invaluable and important in completing this thesis to the best of my ability.

I am also thankful to the reviewers for taking their time to assess my work and provide insightful comments that significantly help improve this thesis. Apart from that, I am also grateful to my parents for their support and advice. Besides that, I also want to thank the Centre of Excellence for Advanced Sensor Technology (CEASTech), Universiti Malaysia Perlis (UniMAP), Malaysia for providing support in terms of equipment and location for my experimental test.

Finally, yet importantly, I would like to thank those whose names are not mentioned that contribute either directly or indirectly towards the successful completion of this thesis. With the help of the following people, I was able to complete this thesis project and I express my gratitude for their help in every way.

Noraini Azmi

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

AI	Artificial Intelligence
DL	Deep Learning
EMC	Equilibrium Moisture Content
EMI	Electromagnetic Imaging
FP	False Positive
FN	False Negative
IoT	Internet of Things
LOD	Loss of Drying
ML	Machine Learning
RF	Radio Frequency
RFID	Radio Frequency Identification
RSSI	Received Signal Strength Indicator
SVM	Support Vector Machine
TP	True Positive
TN	True Negative
VSG	Vector Signal Generator
VNA	Vector Network Analyzer
WSN	Wireless Sensor Network

LIST OF SYMBOLS

°C	Degree Celsius

Pencirian Dua Frekuensi Radio dengan Pembelajaran Mesin bagi Menentukan Kandungan Kelembapan Beras dan Lokasinya

ABSTRAK

Penstoran padi adalah sangat penting kerana padi merupakan tanaman bermusim di mana ia hanya dituai dua kali setahun. Kandungan kelembapan adalah faktor yang paling mempengaruhi penyimpanan padi. Terdapat banyak kaedah yang telah dibanggunkan bagi mengukur kandungan kelembapan padi, namun, kebanyakan kaedah yang ada tidak dapat mengukur kandungan kelembapan (MC) padi secara masa-nyata semasa penyimpanan, sampel perlu dimusnahkan semasa proces pengukuran, kawasan padi yang basah di dalam tempat penyimpanan tidak dapat ditentukan, dan lain-lain. Pada masa ini, kaedah berdasarkan radio frekuensi (RF) diperkenalkan sebagai kaedah yang berpotensi bagi menentukan kandungan kelembapan padi. Walau bagaimanapun, kajian sedia ada memerlukan peralatan yang kompleks dan mahal seperti *Vector Signal Generator* (VSG) dan *Vector Network Analyser* (VNA) yang hanya menggunakan satu jalur frekuensi disamping tiada penentuan kawasan padi yang basah. Oleh itu, kajian ini bertujuan membuat pencirian jalur frekuensi RF untuk menentukan kandungan kelembapan dalam masa nyata dalam penyimpanan beras dengan mengunakan algoritma klasifikasi berasaskan pembelajaran mesin. Beberapa jalur frekuensi radio dikenal pasti dan diselidiki untuk penentuan kandungan kelembapan dan penentuan kawasan beras yang basah. Dari ujian pencirian menggunakkan pengukuran RSSI; dua peranti tanpa wayar yang beroperasi berdasarkan protokol yang ditetapkan iaitu IEEE802.15.4/WSN dan Radio Frequency Identification (RFID) menunjukkan bahawa nilai RSSI menurun apabila kandungan kelembapan beras meningkat. Oleh itu, dua teknologi tanpa wayar yang beroperasi pada jalur 868 MHz dan 2.4 GHz dipilih sebagai dua jalur frekuensi dalam sistem pengukuran yang dikembangkan. Terdapat enam kondisi yang diteliti dalam penyelidikan ini iaitu 0% (bekas kosong), 12% (beras berkelembapan asli), dan 5 kg sampel beras yang dikondisikan kepada kandungan lembapan 14%, 20%, 25%, dan 30%. Sampel beras ditetapkan di lokasi yang berbeza di dalam bekas simpanan (bekas besar yang mengandungi 20 kg bijirin). Empat model pembelajaran mesin iaitu *support vector machine*, *random forest*, *gradient boosting tree*, dan *multiple layer perceptron* dikembangkan untuk klasifikasi kandungan kelembapan. Di antara algoritma tersebut, *gradient boosting tree* memberikan ketepatan tertinggi 94.8% untuk satu ciri output. Sementara itu, dengan ketepatan 95.4%, *random forest* adalah algoritma klasifikasi terbaik untuk dua ciri output. Hasilnya, *gradient boosting tree* hanya sesuai untuk menentukan kandungan kelembapan padi. Sebaliknya, *random forest* adalah algoritma yang sesuai untuk menentukan kandungan kelembapan serta penentuan lokasi kawasan padi yang basah di dalam tempat penyimpanan. Kaedah gabungan telah diperkenalkan dalam penyelidikan ini untuk meningkatkan lagi ketepatan pengelasan daripada algoritma pembelajaran yang lemah. Gabungan *random forest* dan *gradient boosting tree* telah menghasilkan ketepatan tertinggi bagi penentuan kandungan lembapan padi dan juga penentuan kawasan beras yang basah dalam simpanan dengan ketepatan sebanyak 99.5%.

Dual Radio Frequency Characterization with Machine Learning for Rice Moisture Content Classification and Localization

ABSTRACT

Storage is crucial because paddy is a seasonal crop where it is only harvested twice a year. Moisture content is a factor that significantly affects the storage of paddy. Numerous methods have been developed to measure the moisture content in paddy, however, most of the existing techniques are unable to measure the moisture content (MC) of the paddy grain in real-time during storage, destroy the samples during the measurements, unable to localize the grain wetspot in bulk density storage, etc. Recently, radio frequency (RF) based methods have been introduced as a potential method to determine the moisture content. However, the existing work requires complex and expensive equipment such as Vector Signal Generator (VSG) and Vector Network Analyser (VNA) that utilised only one frequency band without localization of grain spoilage wetspots in the storage. Hence, this study aims to characterize the RF frequency band for moisture content determination in real-time in bulk density rice grain storage with a machine learning-based classification algorithm. Multiple radio frequency bands were investigated and identified in order to increase the accuracy of the determination of moisture content and the localization of grain spoilage wetspots. From the characterization test in terms of RSSI measurement, two wireless devices operating based on the established protocol IEEE802.15.4/WSN and Radio Frequency Identification (RFID) was found to have a decreasing RSSI value when the moisture content of the rice was increased. Hence, the two wireless technologies operating in 868 MHz and 2.4 GHz bands, respectively were selected as the dual-frequency band in the developed measurement system. There are six test conditions investigated in this research which is 0% (empty container), 12% (normal moisture condition), and 5 kg of sample conditioned to 14%, 20%, 25%, and 30% moisture content. The samples are set at different locations in the storage container (a large container that contains 20 kg of grain) for wetspot localization purposes. Four machine learning models which are support vector machine, random forest, gradient descent tree, and multilayer perceptron are developed for the classification of moisture content. Among the algorithms, the gradient boosting tree provides the highest accuracy of 94.8% for one output feature. Meanwhile, with 95.4% accuracy, random forest is the best classification algorithm for two output features. As the result, the gradient boosting tree is only suitable for predicting the moisture content of rice. On the other hand, random forest is a suitable algorithm for both predicting the moisture content as well as the localization of the location of the rice spoilage wetspot in the storage. An ensemble method was introduced in this research to further improved the classification accuracy of the weak learners. The ensemble of random forest and gradient boosting trees has provided the highest accuracy for the determination of rice moisture content and localisation of the wetspot in the storage with an accuracy of 99.5%.

CHAPTER 1 : INTRODUCTION

1.1 Overview

Rice (produced after paddy was processed) is one of the major staple food sources, especially in Asian countries. Figure 1.1 highlights the global grain consumption (Enghiad et al., 2017). Paddy is a seasonal crop where the number of plantation and harvest seasons differ for different countries worldwide (Laborte et al., 2017). The paddy is harvested twice a year in Malaysia, thus, some proportion of the harvested paddy and rice must be stored for gradual consumption until the next harvest, and seeds must be held for the next season's crop. From here onwards, the terms paddy and rice will be used interchangeably to refer to either paddy or rice, otherwise, the term grain will be used to refer to both paddy and rice.

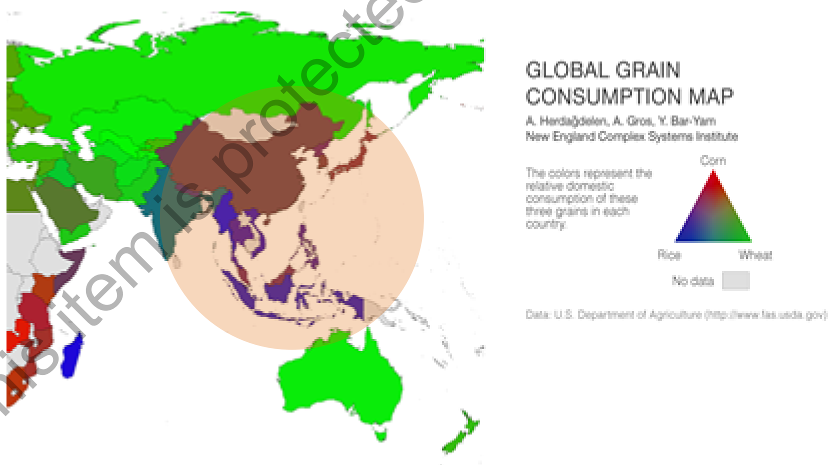

Figure 1.1 Global grain consumption.
Source: (Enghiad et al., 2017). Global grain consumption throughout countries around the world.

Due to the extreme importance of these crops to humanity's food supply, safe storage is critical (Gilmore et al., 2017). Storage duration can vary from several weeks to many years. Stored grain represents significant economic and nutritional value to humankind, but despite this value, storage losses are common. The estimates of storage losses vary from 50% to 60% and these losses lead to economic costs (Kumar and Kalita, 2017).

There are numerous threats to the quality and quantity of stored grain: the seeds may germinate while in storage, moulds and fungi may grow, insect (e.g. *Sitophilus zeamais* or commonly known as rice weevil) infestations may occur (Zhang et al., 2020), and grain got eaten by rodents and other animals (Maier et al., 2006; Maier and Channaiah, 2010). Due to the tropical weather in Malaysia, maintaining the optimal condition for the storage is quite challenging. Even though various factors greatly affect the safe storage of grain, according to Yang et al. (2018), the moisture content is the most critical factor that affects the quality of the grain compared to the other factors. The moisture content of grains is one of the important parameters for grain quality control, especially during harvesting, milling, and storage. The moisture content of harvested paddy is usually high (19–25%) and thus needs to be dried to 14% or less for safe storage. Grain wastage often occurs due to improper storage conditions where high moisture content promotes the growth of mould and insect infestation whereas very dry grains are brittle and susceptible to breakage. If the moisture content is not tested properly, the inaccurate result will lead to extra drying costs and harvesting loss, or if paddy is harvested wetter than necessary, then it will be too wet in storage, and it will spoil. However, excessive drying leads to a reduction in head rice recovery if the rice is milled at the wrong moisture content level, there will be weight loss and thus a drop in profit.

Moisture content in the grain is also affected by the weather where moisture content becomes high during the rainy season, otherwise too low in the summer or hot season. After harvesting, grains are typically stored in silos (cylindrical storage). Since silos are quite large (e.g., a diameter and height range from 4.6 m to 18.3 m and 4.6 m to 28.7 m, respectively), the moisture content may not be uniformly distributed in the silo. The moisture content measurements are not carried out frequently enough to observe the dynamics of moisture content changes within the bin continuously. Overdrying usually occurs in the bottom layer of the silo during the drying process. Balancing between high moisture content and excessive drying is not easy due to the variable nature of grains within bulk and the inherent difficulties of measuring grain moisture accurately (Hossain et al., 2016). Moreover, the tropical weather in Malaysia also contributes to the challenge of maintaining the optimal condition for safe storage.

Due to the tropical weather in Malaysia, maintaining the optimal condition for the storage is quite challenging. Even though various factors greatly affect the safe storage of grain. However, according to Yang et al. (2018a), the moisture content is the most critical factor that determines the quality of the grain compared to the other factors. If the moisture content is not tested properly, the inaccurate result will lead to extra drying costs and harvesting loss, or if paddy is harvested wetter than necessary, the grain will be too wet in storage, and it will spoil. Sprouting, infestations of insects, and mould growth as shown in Figure 1.2 create heat, carbon dioxide and moisture through biochemical processes which further increase the moisture content and temperature (Kaushik and Singhai, 2019; Zhang et al., 2020). Increases in temperature and/or moisture both raise the complex permittivity of the grain (Gilmore et al., 2017). On the other hand, excessive drying leads to a reduction in head rice recovery if the grain is milled at the wrong moisture content level, there will be weight loss, induce cracks in the grain leading to breakage during processing and thus a drop in profit (Frías et al., 2002). From the literature, moisture is a factor that significantly affects the storage of paddy or rice. Hence, numerous researches have been conducted to investigate the effect of moisture on the physical and chemical properties of rice (Visvanathan et al., 1996; Zareiforoush et al., 2009).

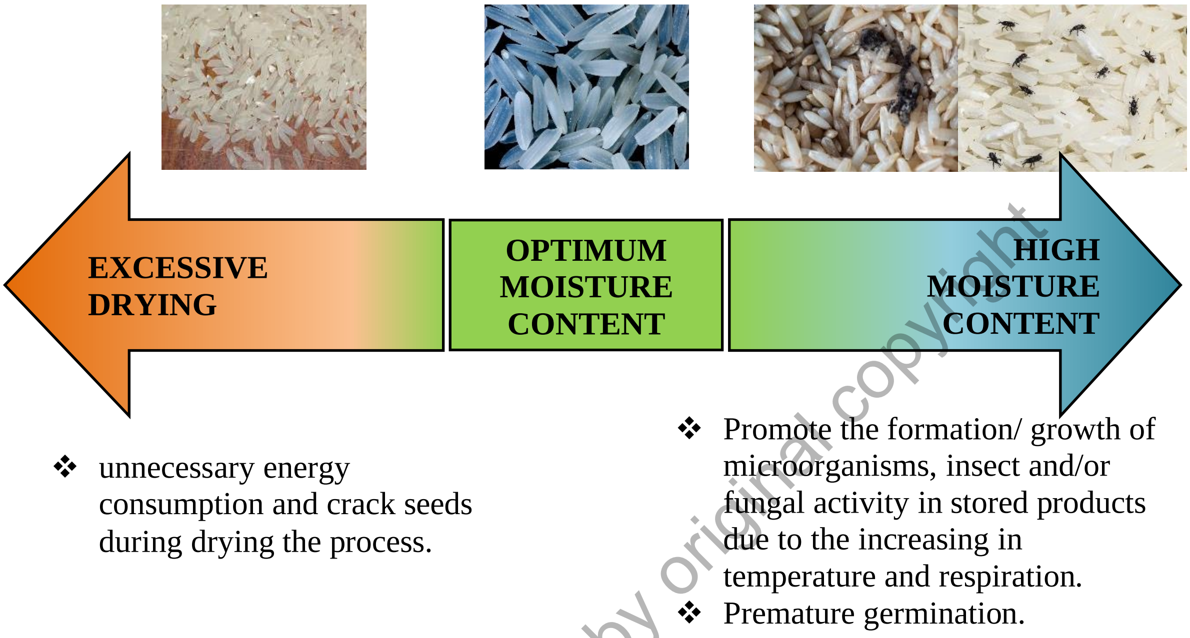

Figure 1.2 Moisture content.

Findings from Moldenhauer et al. (2018) and Berkley (2016) suggest that water attenuates the Received Signal Strength Indicator (RSSI) where the average RSSI attenuation for an empty container and when the container is filled with water is approximately 90 dBm and 69 dBm, respectively. The possibility of using the RSSI of the free space radio wave and the introduction of various applications that eases the development of a model using machine learning is the source of the motivation for this research.

Various methods are available for moisture content measurement. However, most of the existing methods are unable to determine moisture content (MC) in bulk density of rice grain storage, unable to determine wetspot, are expensive, complex and do not provide continual automated monitoring. Continuous and real-time monitoring of moisture content and the localisation of grain spoilage wetspot is important to effectively manage and market the grain. Accurate measurement of grain moisture and timely drying could greatly reduce damage to grain caused by insect and mould activity during subsequent handling, storage, and processing. Therefore, continuous monitoring is critical, especially in tropical climates where the weather varies throughout the year. The temperature sensor is unable to detect grain spoilage wetspot due to the temperature sensor's limited sensing range. Existing methods such as moisture metres used only samples (a portion from the whole grain). Hence, those moisture content detection methods are destructive and dependent on the bulk density. Another method for moisture content determination using oven drying is also time-consuming and makes the production process difficult to control despite its high accuracy measurement.

1.2 Problem Statement

Recently, RF-based methods for moisture content measurement are widely researched. However, the proposed methods in the literature review only utilised a single frequency band for moisture content detection, using laboratory-grade instrumentations that are expensive and complex such as horn/lens antennas, Vector Signal Generator (VSG) and Vector Network Analyzer (VNA). Research work by Almaleeh et al. (2022) proposed measurement of rice moisture content using WiFi operating in the 2.4 GHz frequency band. In another related research work, the system developed by Mohd Ramli et al. (2021) also used the 2.4 GHz frequency band but WSN was used in place of WiFi. In a different study, a Gunn diode installed in the waveguide cavity (WR90) was used to generate a continuous RF signal with a frequency of 10.5 GHz that was then transmitted through a horn antenna (Li et al., 2021). These studies suggest the operation of the measurement system using different frequency bands and different wireless protocols. In different research work, Jain, Kumar Mishra, and Vikas Thakare (2019) propose a dual-frequency of 2.2 GHz and 4.6 GHz to detect moisture content in rice grain using a rectangular microstrip patch antenna. The finding from Jain et al. (2019) motivates the use of multiple wireless transmissions from different wireless RF devices for the measurement of rice moisture content.

To the best of the author's knowledge, there is yet a study done utilising different RF frequencies from different wireless devices to detect grain moisture content in bulk during storage and localization of wetspot for spoilage prediction. Therefore, this research works aims to investigate which existing commercial wireless device can be used for the determination of moisture content and localization of rice wetspot in storage. Hence, this research aims to characterize the widely used off-the-shelf wireless devices operating based on the standard protocol and frequency bands such as 2.4 GHz, 868 MHz, and 915 MHz bands. The wireless devices used operate based on the established protocol such as IEEE80.15.4/Zigbee, IEEE802.11/WiFi, LoRa, and RFID which is low cost and able to monitor the MC of grain in real-time and in a non-invasive way. The relationship between RF frequency, attenuation and moisture content will be characterized and studied. This research aims to propose a new method and developed a system of density-independent moisture content detection utilizing a radio frequency signal that fluctuates based on the moisture content of the grain.

Through the characterization tests, two wireless technology were found applicable for the determination of moisture content and localization of rice wetspots in storage which are the RFID and WSN. Thus, this research work proposed a novel measurement system using two different wireless technologies which are RFID and WSN that were operating in 915 MHz and 2.4 GHz frequency bands, respectively. Hence, the term dual-frequency band was introduced to represent the operating frequency for both wireless technologies. This research developed a novel method of moisture content determination of grain (focusing on rice) and localisation of rice spoilage wetspot using the combination of dual frequencies which is 2.4 GHz and 915 MHz with machine learning techniques for continuous non-destructive monitoring.

1.3 Research Objectives

There are three main objectives that want to be achieved in this research work are as follows:

1) To perform radio frequency characterization in terms of the measured RSSI values from the wireless technologies operating based on the established protocols such as IEEE802.15.4/Zigbee, IEEE802.11/WiFi, LoRa, and UHF RFID, and select suitable wireless technologies for the determination of grain moisture content and localized the location of the grain spoilage wetspot.

2) To design an experimental testbed and perform a series of experimental tests for the determination of grain moisture content and localize the location of the grain spoilage wetspot using the selected wireless technologies from Objective 1.

3) To evaluate the performance of the developed system and algorithm for moisture content determination and localization of grain wetspot using the

dataset collected from Objective 2 and presents the best model for the moisture content classification and localization of grain spoilage wetspot.

1.4 Research Scope

The determination of moisture content and localization of grain spoilage wetspots using wireless technologies operating based on the established protocol has not been extensively studied. Hence, there is a wide area of research available to be further investigated. However, due to time limits and resources, thus, this research work is limited to the following scope:

1) The characterization of RF on moisture content is limited to wireless technologies operating based on the established protocols such as IEEE802.15.4/Zigbee, IEEE802.11/WiFi, LoRa, and UHF RFID.

2) Even though there is a wide range of agricultural products, this research is interested in rice because it is the major staple food, especially in the Asian continent. The grain used is polished rice from the local brand which is *Rambutan* Brand. This brand is chosen because it is one of the popular brands in Malaysia and is available at most of the local stores in Malaysia.

3) The determination of grain moisture content and localization of grain spoilage wetspots is limited to IEEE802.15.4/Zigbee and UHF RFID.

4) Limitation of the environment for sensing techniques where data collection via experimental tests were performed in a laboratory. Thus, a well-designed silo/warehouse can be developed and investigated by future research.

1.5 Research Questions

In order to achieve the research objectives and solve the problem raised in this research, this research is conducted to answer the following questions:

1) Is there any significant impact of different frequencies on the moisture content in the polished rice?

2) Is the classification and localization algorithm able to help determine the moisture content of grain and the location of the grain wetspot in storage?

3) Which machine learning algorithm is suitable for moisture content determination and localisation of grain wetspot in storage?

1.6 Thesis Outline

This thesis is organized into five chapters where Chapter 1 covers the introduction, problem statement, research objectives, research scope, and research question. Chapter 2 presents the literature review of previous and related works. Chapter 3 describes the methodology of the research while Chapter 4 presents the result and discussion of the findings, and lastly, Chapter 5 presents the conclusion of this study. Figure 1.3 shows a summary of the thesis outline.