Journal of Stored Products Research 108 (2024) 102399

Contents lists available at ScienceDirect

Journal of Stored Products Research

journal homepage: www.elsevier.com/locate/jspr

Research Paper

Integration of dual band radio waves and ensemble-based approach for rice moisture content determination and localisation

Noraini Azmi [a,b], Latifah Munirah Kamarudin [a,b,*], Ahmad Shakaff Ali Yeon [a,c], Ammar Zakaria [a,c], Syed Muhammad Mamduh Syed Zakaria [a,b], Hiromitsu Nishizaki [d], Latifah Mohamed [c], Xiaoyang Mao [d], Mohd Hafiz Fazalul Rahiman [a,c]

[a] *Centre of Excellence for Advanced Sensor Technology (CEASTech), Universiti Malaysia Perlis (UniMAP), Malaysia*
[b] *Faculty of Electronic Engineering & Technology (FKTEN), Universiti Malaysia Perlis (UniMAP), Malaysia*
[c] *Faculty of Electrical Engineering & Technology (FKTE), Universiti Malaysia Perlis (UniMAP), Malaysia*
[d] *Graduate Faculty of Interdisciplinary Research, University of Yamanashi, 4-3-11 Takeda, Kofu, Yamanashi 400-8511, Japan*

ARTICLE INFO

Keywords:
Artificial intelligence
Electromagnetic wave absorption
Ensemble technique
Machine learning

ABSTRACT

Maintaining optimal moisture content in grain storage is critical to ensuring adequate supply throughout the year, but it presents a significant challenge. Current moisture measurement methods often necessitate sophisticated and costly equipment. This paper introduces an approach employing real-time rice moisture content determination and detection of spoilage (specifically wet spots) within a storage facility achieved through the utilisation of radio waves operating at 2.4 GHz and 868 MHz, along with an ensemble-based machine learning algorithm. Experimental samples spanning from 12% to 30% moisture levels were collected, then subjected to pre-processing, and subsequently employed to train the Ensemble-based Rice Moisture Content and Localisation (eRMCL) algorithm. The eRMCL produced an effective prediction of both rice moisture content and the localisation of wet spots within the grain storage unit. The results show that compared to support vector machine, random forest, and machine learning methods, the eRMCL algorithm had the best performance metrics, with an accuracy of 94.8% in predicting the moisture content and location of spoilage in storage. The measurement of moisture content and the identification of wet spots in rice storage using the dual frequency wave approach were found to be more accurate than with a single frequency band. Thus, the dual frequency band is a novel method for the determination of the moisture content of stored rice and the localisation of the spoilage area.

Abbreviations

Abbr.	Details
COTS	Commercial-off-the-shelf
eRMCL	Ensemble-based Rice Moisture Content and Localisation (eRMCL) algorithm
EPC	Electronic product codes
GBT	Gradient boosting trees
HF	High frequency
LF	Low frequency
LOD	Loss of drying
LOS	Line-of-sight
LoRA	Long Range Network
MC	Moisture content
MLP	Multilayer perceptron
NLOS	Non-line-of-sight
PIE	Pulse interval encoding
RF	Random forest
RFID	Radio Frequency Identification
RSSI	Received signal strength indicator
SVM	Support vector machine
TARI	Type A Reference Interval
UHF	Ultra-high Frequency
VNAs	Vector network analysers
VSGs	Vector signal generators

* Corresponding author. Centre of Excellence for Advanced Sensor Technology (CEASTech), Universiti Malaysia Perlis (UniMAP), Malaysia and Faculty of Electronic Engineering & Technology (FKTEN), Universiti Malaysia Perlis (UniMAP), Malaysia.
E-mail address: latifahmunirah@unimap.edu.my (L.M. Kamarudin).

https://doi.org/10.1016/j.jspr.2024.102399
Received 10 June 2024; Received in revised form 2 August 2024; Accepted 18 August 2024
Available online 28 August 2024

1. Introduction

Stored grain is subjected to numerous threats to its quality and quantity, such as germination, mould growth, insect infestation, and animal-related damages (Li et al., 2022; Zhang et al., 2020; Maier and Channaiah, 2010). According to Yang, Wang and Cao (Yang et al., 2018a), moisture content is the most critical factor affecting stored grain. If the moisture content is not properly assessed, inaccurate measurements will lead to additional drying costs and crop losses. When paddy is harvested wetter than necessary, the grain needs to be dried; otherwise, it will remain wet during storage, resulting in spoilage. Excessive drying, on the other hand, results in lower rice grain yields, and if the grain is milled with the wrong moisture content, there will be a loss of weight and subsequently, a decrease in profit. Therefore, accurate moisture content measurement and control are critical to ensure safe and profitable grain storage.

The rapid advancement of wireless technologies and artificial intelligence has demonstrated their potential in addressing various issues (Girmay et al., 2023; Le-Huu and Seo, 2023; Nagai et al., 2021; Yigit and Duysak, 2022). The diverse wireless technologies operating in different frequency bands and communication protocols have facilitated the integration of these technologies. The emitted radio waves of these wireless technologies were leveraged for rice moisture content measurement and wet spot localisation with the assistance of artificial intelligence. Consequently, it is anticipated that the approach introduced in this paper will enhance the accuracy of moisture content measurements and wet spot localisation. Moreover, by utilising commercially available wireless devices, the experimental tests in this study are more easily reproducible by other researchers. Precise moisture content measurement and wet spot localisation can help prevent food spoilage and improve food safety, benefiting the food and agricultural industries.

This study introduces a novel method for density-independent moisture content detection in stored grain using radio-frequency signals that fluctuate with the grain's moisture content. A measurement system was successfully developed in this research using commercially available wireless technologies operating based on established protocols such as Zigbee, Wi-Fi, Long Range (LoRa), and Radio Frequency Identification (RFID), to enable real-time and non-invasive monitoring of grain moisture content. In addition to measuring moisture content, localising wet spots in storage is crucial for reducing rice spoilage during long-term storage. However, to the authors' knowledge, only a few studies have focused on grain moisture localisation. However, existing methods, like those proposed by Asefi et al. (2017) and Gilmore et al. (2017), rely on sophisticated and delicate equipment like vector network analysers (VNAs), which are impractical for rural agricultural settings.

This research aims to utilise the influence of moisture content on the RSSI measured from commercially available wireless technology for the determination and localisation of wet spots in storage. Hence, this research investigates several commercially available wireless technologies and selects wireless technologies that correlate with the moisture content in rice. Dual frequency bands were selected in conjunction with RSSI measurements from two identified wireless technologies in order to improve the detection of rice moisture content and localisation of spoilage in storage. To the best of the authors' knowledge, a few studies have used RSSI measurement from a single frequency band and simple machine learning algorithms, such as a support vector machine and regression analysis, for the measurement of moisture content (Yang et al., 2018b; Almaleeh et al., 2022). However, a study by Le-Huu and Seo (2023) highlighted that a measurement system using dual-frequency can improve the measurement accuracy threefold compared to when a single frequency is used.

The remainder of this paper is organised as follows: Section 2 reviews related works on moisture content measurement methods. Section 3 presents the testbed and experimental design. Section 4 describes the data collection process and dataset preparation. Section 5 details the model development for the determination of moisture content and localisation of wet spots. Section 6 discusses the results, and Section 7 provides the conclusion and future research directions.

2. Related works

Farmers and grain handlers employ various methods to ensure optimal grain conditions post-harvest and during storage, which can range from a minimum of 12 months up to 36 months (Müller et al., 2022; Wang et al., 2022). Moisture measurement methods range from simple practices such as opening storage bins to check (e.g. smell) for mould and fungus, to more sophisticated methods such as point-sensing systems, temperature and moisture sensors, and the dielectric method. Entering storage bins to assess moisture content is time-consuming, dangerous, and often unreliable for detecting poor storage conditions (Freeman et al., 1998). Grain moisture content can be measured using primary (direct) methods, which involve removing water from the grain, or secondary (indirect) methods, which rely on measurements based on the grain's physical or chemical properties, which are influenced by moisture content (Gilmore et al., 2017; Chen, 2003; Hossain et al., 2016; Nath et al., 2017; Asefi et al., 2015). While direct methods, such as the loss of drying (LOD) method, are accurate, they are time- and power-consuming, and unsuitable for online moisture detection (Liu et al., 2015). In contrast, indirect electrical measurements are non-destructive, fast, easy to use, and more practical for online moisture detection. The non-destructive nature of these measurements is preferable because it allows the sample to be used directly without crushing the grain (Nath et al., 2017).

Extensive research on electrical measurement methods has led to various techniques for determining grain moisture content (Gilmore et al., 2017; Toba and Kitagawa, 2011; Vallejos and Grote, 2009; Jusoh et al., 2011). Microwave moisture sensors operate on the principle of the dielectric constant measurement in a composite material, where the permittivity is primarily influenced by the presence of water (Aichholzer et al., 2018). Studies by Moldenhauer et al. (2018) and Berkley (2016) indicate that water attenuates the received signal strength indicator (RSSI), causing a decrease in signal strength with the introduction of water into the container. This phenomenon can be leveraged to detect moisture content in stored grains by analysing the RSSI values of the radio waves transmitted through the grain, thus enabling the prediction of spoilage likelihood. Consequently, this research is driven by the potential use of RSSI values from free-space radio waves, alongside the development of various machine learning models that can be tailored for real-time moisture content in grain storage.

Wee et al. (2009) and Hassan et al. (2013) employed a free-space microwave measurement setup to measure the dielectric properties of agricultural residues, such as paddy husks. Kandala et al. (2013) developed an impedance analyser to measure impedance and phase angle at frequencies 1, 5, and 9 MHz using a parallel plate sensor for in-shell peanut moisture content. Yang et al. (2018b) used Wi-Fi cards in laptops to measure moisture content in a small paddy sample, considering line-of-sight (LOS) and non-line-of-sight (NLOS) propagation in a computer laboratory environment. In another study, Trabelsi et al. (2016) measured the dielectric properties of unshelled peanuts using 5.8 GHz free-space transmission by evaluating signal attenuation and phase shift of a microwave signal propagating through a layer of material situated between two opposing antennas. Therefore, moisture determination and spoilage localisation through radio frequency sensing methods and artificial intelligence are desirable, as they minimise the need for human operators inside the grain storage.

In recent years, there has been extensive research into the use of radio frequency–based methods for measuring moisture content in stored grain. However, to the authors' knowledge, most of these methods utilise only a single frequency band and often depend on expensive and complex laboratory-grade instruments, such as horn/lens antennas, vector signal generators (VSGs), and vector network analysers

(VNAs) (Li et al., 2022; Yigit and Duysak, 2022; Gilmore et al., 2017; Asefi et al., 2015). Fortunately, the availability of affordable commercial-grade wireless technologies has been increasing in the market, along with growing awareness of the importance of artificial intelligence in enhancing existing applications. Therefore, this research aims to investigate and characterise the effect of grain moisture content with radio waves using commercial-off-the-shelf (COTS) transceivers and ensemble-based machine learning algorithms. Additionally, this research introduces a method to determine the moisture content and localise wet spots, enabling real-time measurement and monitoring of rice conditions in storage.

3. Experimental design and procedures

The experimental tests were conducted indoors using COTS wireless technologies, adhering to relevant institutional, national, and international guidelines and regulations. Various commercially available wireless technologies were utilised, including Wi-Fi at 2.4 GHz, Ultra-high Frequency (UHF) RFID at 865 MHz up to 928 MHz band, LoRa at 915 MHz, and Zigbee at 2.4 GHz band. The Waspmote board (Libelium, 2014) was used For communication utilising the Zigbee, the in-house designed Tri-comm Shield (a multi-communication board) enabled Wi-Fi and LoRa communication (Azmi et al., 2018), and the ThingMagic Mercury6e (M6e) embedded UHF RFID Reader Module Developer Kit was used for RFID communication (ThingMagic, 2016).

Zigbee transceivers were connected to a 3 dBi omnidirectional antenna, and the transmit power was set to 0 dBm. For RFID, only one antenna port was used, despite the ThingMagic M6e supporting four antenna ports with a maximum transmission power of 31.5 dBm. RFID communication parameters included a backscatter link frequency (BLF), Type A Reference Interval (TARI), frequency bands for tag-reader, reader-to-tag, tag-to-reader, and tag contention interactions. The tag communicates with the reader through FM0 baseband or Miller modulation of a subcarrier at the transmission data rate, where the reader is responsible for detecting and decoding the tag's response. The BLF selection is based on the modulation type, with options of 250 kHz and 640 kHz. For this study, a BLF of 250 kHz was chosen due to the use of Miller encoding, which requires a BLF equal to 250 kHz, as higher frequencies only support FM0 encoding. TARI, analogous to Morse code's dot (or dash) length, signifies a binary '0′ with a short, high-level pulse followed by an equal-length low pulse. TARI lengths range from 6.25 μs to 25 μs (e.g., 25 μs, 12.5 μs, and 6.25 μs), and a 6.25 μs TARI was chosen to allow weak energy tags ample time to respond to the reader. The link rate dictates the speed of transmitting dots and dashes. The 'M' value, influencing tag-reader communication and symbol repetition, was set to 8. EPC Gen 2-compliant readers employ pulse interval encoding (PIE) for binary data coding. The Q slot number assigned from reader-to-tag involves dynamic Q when the number of tags is unknown, and static Q otherwise. Hence, the Q value was set to 4. RFID commonly operates at low frequency (LF), high frequency (HF), and ultra-high frequency (UHF). Among these, UHF, utilising backscatter and an electromagnetic field, was chosen.

3.1. Rice sample preparation

The rice utilised in this research was procured from local stores with an initial moisture level of approximately 12%. The rice was divided into 5 kg portions and moistened to achieve specific moisture levels of 12%, 14%, 20%, 25%, and 30%, following ASAE Standards 1987 (S352) (Chen, 2003; Grabe, 2015). Moisture content (MC) was calculated using the percentage difference between the wet weight (W_w) and dry weight (W_d) of the material. Eq. (1) represents the general formula used for calculating the moisture content of a material. MC can be expressed in terms of either wet basis (MC_{wb}) or dry basis (MC_{db}), as shown in Eq. (2) and Eq. (3), respectively. Previous studies have used both wet basis (Asefi et al., 2017; Kok et al., 2017; Basati et al., 2018; Balasubramanian, 2011) and dry basis (Zareiforoush et al., 2009; Sacilik et al., 2003; Coşkun et al., 2006) for moisture content calculations.

$$\mathrm{MC} = \frac{\text{Weight of moisture}}{\text{Weight of sample}} \times 100\% \tag{1}$$

$$\mathrm{MC_{wb}} = \frac{(\mathrm{W_w} - \mathrm{W_d})}{\mathrm{W_w}} \times 100\% \tag{2}$$

$$\mathrm{MC_{wb}} = \frac{(\mathrm{W_w} - \mathrm{W_d})}{\mathrm{W_d}} \times 100\% \tag{3}$$

3.2. Radiofrequency signal characterisation

The configuration in Fig. 1 was used for the preliminary test to identify the optimal wireless technology for moisture content determination and spoilage localisation in storage facilities. Fig. 1 (a) illustrates the testbed arrangement for Zigbee, Wi-Fi, and LoRa modules, while Fig. 1 (b) shows the setup for a single passive RFID tag.

As depicted in Fig. 2, the RSSI values from Zigbee and RFID decreased with increasing moisture content, showing high correlation values of −0.974 and −0.948, respectively. In contrast, LoRa and Wi-Fi exhibited lower correlations of −0.753 and −0.885, respectively. The negative sign indicates that the RSSI measured for each wireless technology has a negative relationship with the moisture content. In general, as the moisture content in rice increased, the measured RSSI value decreased. As depicted in Fig. 2, the RSSI values decrease as the moisture content increases, as expected. This phenomenon occurs because water tends to absorb microwave energy (Kok et al., 2017). Among the COTS technologies, only Zigbee and RFID show a high correlation with moisture content. This is because water molecules in rice behave differently when exposed to different frequency ranges, as mentioned by (Kok et al., 2017).

The result in Fig. 2 suggests that the characteristics of each wireless technology also affect the mechanism of water molecules in rice. For instance, Wi-Fi has a high transmit power of about 20 dBm compared to Zigbee, which only transmits power at 1 dBm. Therefore, Zigbee's RSSI value is significantly affected by the moisture in rice (Kraszewski et al., 1998). It can also be concluded that for a small quantity of rice, the Wi-Fi signal was not significantly affected by the moisture content and may be deemed unsuitable for the method to be used in this research. Moreover, the Wi-Fi signal remains unaffected because higher microwave power can cause over-drying, as claimed by (Kok et al., 2017). Despite that, further work needs to be conducted in the future to investigate the characteristics of Wi-Fi for larger quantities of rice.

LoRa, on the other hand, employs a unique modulation scheme that enables it to transmit over a long distance, overcoming obstacles such as buildings and trees. This suggests that changes in moisture content are not capable of affecting the LoRa signal. Considering that LoRa technology was primarily designed for long-distance communication, the small amount of water in rice is negligible. Therefore, this research found that LoRa is not applicable for detecting the moisture content of rice in 30 cm^2 storage. However, further studies are needed to investigate the feasibility of using LoRa technology for measuring rice moisture content in larger industrial-grade silos. Moreover, RFID utilizes a different communication method than Zigbee, Wi-Fi, and LoRa. Instead of point-to-point communication, the passive RFID tag employs the backscatter concept, wherein each tag receives power from the antenna and reflects the received power to the antenna. Despite the antenna transmitting at a high transmit power of approximately 30 dBm, the power reflected by the passive RFID tags to the antenna is attenuated and also affected by the moisture content in rice (Trabelsi et al., 2001).

The results from the characterisation test suggest that two out of four tested wireless technologies, operating based on the established protocols, Zigbee and RFID, yielded a significant negative correlation with the moisture content of rice, with correlation coefficients of −0.974 and

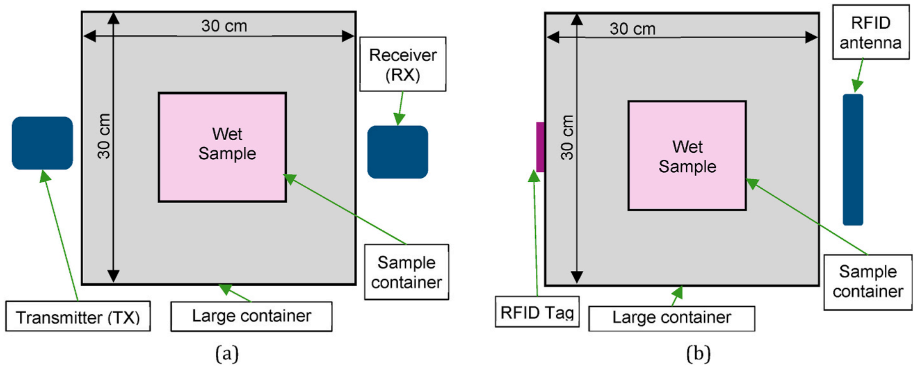

Fig. 1. Testbed for the preliminary test where (a) the testbed for Zigbee, Wi-Fi and LoRa, and (b) the testbed for RFID (single passive tag).

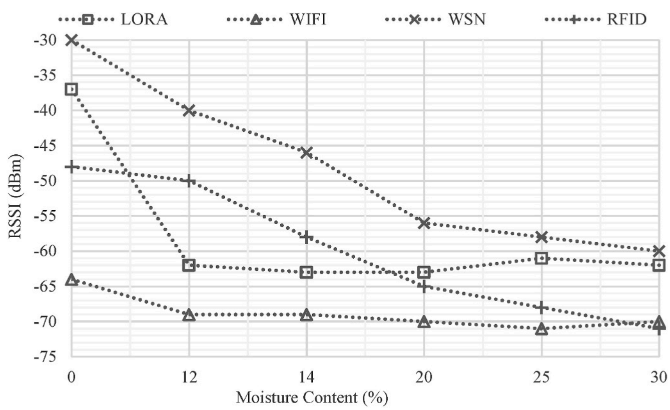

Fig. 2. The correlation between RSSI and Moisture Content.

−0.948, respectively. Since Zigbee and RFID showed a high correlation with moisture content, they were selected for the determination of moisture content and localisation of the rice spoilage wet spot, in this research. As the RSSI from the Zigbee yielded the highest correlation with the moisture content of rice, Zigbee is deemed suitable for the method introduced in this research. Additionally, since passive RFID tags are batteryless, they are better suited for localising wet spots. Hence, the combination of Zigbee and RFID was chosen for this research to increase the accuracy of the method of moisture content determination and localisation of grain wet spots in storage.

Furthermore, considering that radio waves must penetrate the medium (in this instance, rice) and avoid ground reflection, the Fresnel zone equation, as shown in Eq. (4), was employed to ascertain the appropriate node height. Thus, for this measurement, the wireless node was positioned on a pole approximately 15 cm above the table, as shown in Fig. 3. This positioning allowed the signals to traverse the medium in accordance with the principles of the Fresnel zone.

$$\text{Radius}, \boldsymbol{r}\ (\text{cm}) = \left(17.31 \times \sqrt{\frac{\text{Distance, d (km)}}{4 \times \text{Freq}, f\ (\text{GHz})}}\right) \times 100$$

$$\text{Radius}, \boldsymbol{r}\ (\text{cm}) = \left(17.31 \times \sqrt{\frac{0.00070}{4 \times 2.4}}\right) \times 100 \tag{4}$$

$$\text{Radius}, \boldsymbol{r}\ (\text{cm}) \approx 15\ \text{cm}$$

For the determination of the suitable RFID frequency, approximately 30 passive RFID tags were placed around the container, and the number of unique electronic product codes (EPC) detected was recorded and plotted in the bar chart shown in Fig. 4. The investigation revealed that only the 865–869 MHz frequency range could detect all 30 unique EPCs. In contrast, the 902–928 MHz range failed to detect all 30 unique EPCs or tags. Consequently, the 868 MHz channel was selected for the subsequent tests. While RFID operated at 868 MHz bands, Zigbee operated at 2.4 GHz band on channel 26, chosen based on our previous works (Azmi et al., 2014; Azmi, 2016), which indicated low interference to other channels.

Based on the preliminary experimental test conducted in Section 3.2, the 868 MHz band was selected as the communication frequency between passive RFID tags and the reader. Unlike Zigbee, Wi-Fi, and LoRa, RFID has a wide range of frequency bands. The parameters used for Zigbee, Wi-Fi, and LoRa are provided in Table 1.

3.3. Rice sample preparation

The experimentation encompassed five distinct moisture conditions: 12%, 14%, 20%, 25%, and 30%. A control condition with an empty container was also included designated as container ‘0’ (empty

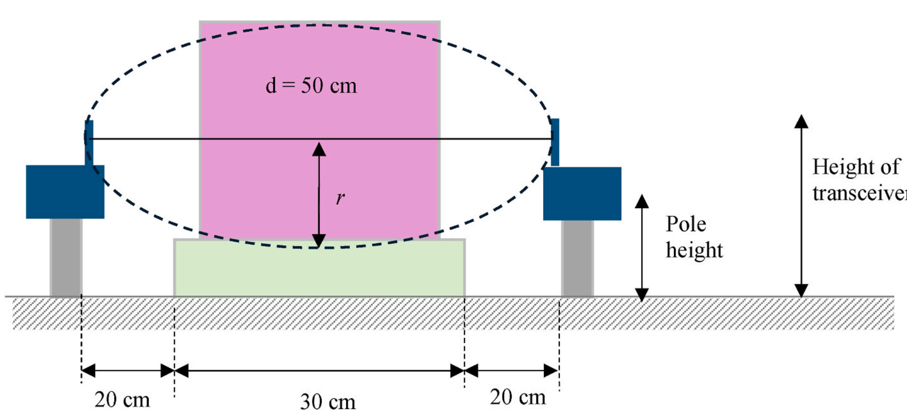

Fig. 3. Pole’s height based on Fresnel Zone Formula.

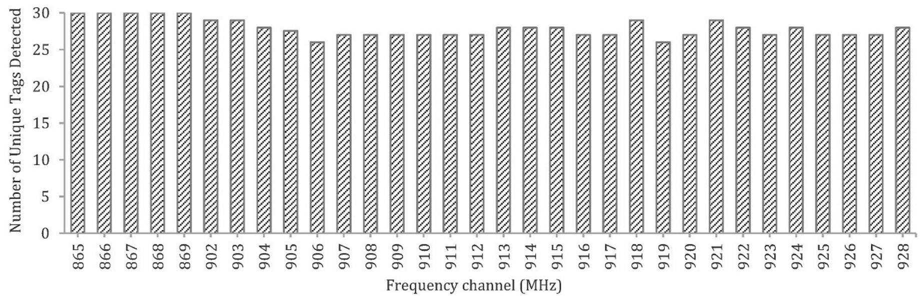

Fig. 4. The number of unique tags detected per frequency channel.

Table 1
Parameters for each wireless technology.

Parameter	Zigbee	Wi-Fi	LoRa
Physical layer	IEEE802.15.4	IEEE802.11 b/g/n Wi-Fi	LoRa
Application layer	Zigbee	HTTP/FTP	Application
Module	XBee 802.15.4	ESP-12F	sx1272
Frequency band	2.4 GHz	2.4 GHz	915 MHz
Transmit power	0 dBm (1 mW)	20 dBm (100 mW)	14 dBm
Channel	Channel 26	Channel 6	Channel 17
Antenna	3 dBi	3 dBi	3 dBi
Antenna polarization	Omni-directional	Omni-directional	Omni-directional

container). Ambient air temperature and relative humidity were monitored throughout, with trials conducted at room temperature (24–37.8 °C) and relative humidity between 70.1% and 95.3%. Fig. 5 shows a block diagram of the developed dual-band system for the measurement of rice moisture content and the localisation of wet spots in storage. Fig. 6 (a) illustrates the actual setup of the system. The RFID system included 30 passive RFID tags, strategically positioned on and inside the container. Each side of the container had six tags arranged in a 3 by 2 configuration, with an additional six tags on the cardboard within the container. Fig. 6 illustrates the arrangement and orientation of these tags.

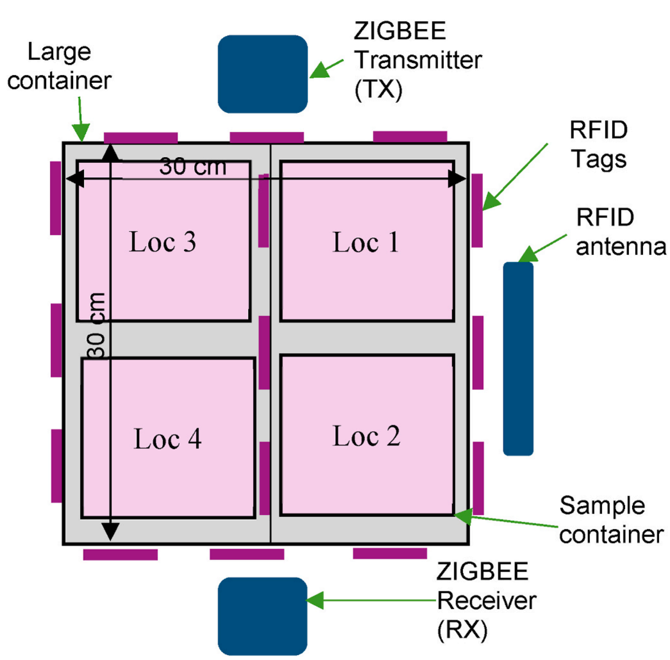

Fig. 5. The developed system for the determination of rice moisture content and localisation of wet spots.

The study utilised UHF RFID Alien 9662 H3 Wet Inlay EPC Rewritable tags. Tags were labelled to indicate their positions: Side A (front), Side B (left side), Side C (back facing the RFID antenna), Side D (right side), and Side IN (inside the container). For instance, tags on Side A were labelled A11, A12, A13, A21, A22, and A23. Tags were spaced 10 cm horizontally and 14 cm vertically, as shown in Fig. 6 (b), ensuring comprehensive coverage of the container. The reader is affixed to a 6 dBi directional antenna, enhancing the efficiency of the RFID system. The reader was serially connected to a personal computer for data collection. RSSI values from the tags were obtained through Universal Reader Assistant (URA) software and PuTTY. The system measured the RSSI value of radio waves transmitted through the rice medium, with data collected via serial communication from each wireless technology to a laptop. For validation, the moisture content of each sample was measured using a commercial portable moisture meter. The collected data were used to develop a model for determining moisture content and the localisation of spoilage in storage.

The flowchart in Fig. 7 illustrates the experimental process in this study. Rice samples were conditioned to six moisture content levels (MC0, MC12, MC14, MC20, MC25, and MC30) and placed in four locations (Loc1, Loc2, Loc3, and Loc4) within the large container, as shown in Fig. 5. The output data, labelled as combinations of moisture content and location (e.g. MC0_Loc1, MC0_Loc2), were classified into 24 groups (six moisture levels × four locations). The combination of input features and output labels was used in four machine learning algorithms to determine the most accurate model for the determination of moisture content and localisation of wet spots.

4. Data collection and preparation

Initial data collection and analysis were conducted to identify the optimal combination of commercially available wireless technologies. Among the four wireless technologies considered, Zigbee and RFID were selected for the determination of the moisture content and localisation of wet spots within storage. RSSI data were gathered from RFID, Wi-Fi, Zigbee, and LoRa technologies for five moisture levels (12%, 14%, 20%, 25%, and 30%) and a control condition (labelled as MC0). Data were collected from four sample positions (Loc 1, Loc 2, Loc 3, and Loc 4) as shown in Fig. 5. Each test lasted about 10 min, with moisture content measurements repeated thrice to reduce random error. This dataset was subsequently employed for training purposes, while a fresh grain sample was prepared for each test, with new data collected and employed as the test dataset for evaluation.

In this research, the two independent variables are moisture content and the locations of the sample, while the dependent variables include temperature, humidity, RSSI value from Zigbee, and 30 RSSI values from passive RFID tags, totalling 33 dependent variables. In this study, the independent variables are also referred to as labels, while the dependent variables are known as input parameters or features.

The entire dataset comprised 5376 observations, divided into training (75%, $n = 4032$) and test (25%, $n = 1344$) subsets. The training dataset included 64 observations per sample position, repeated three

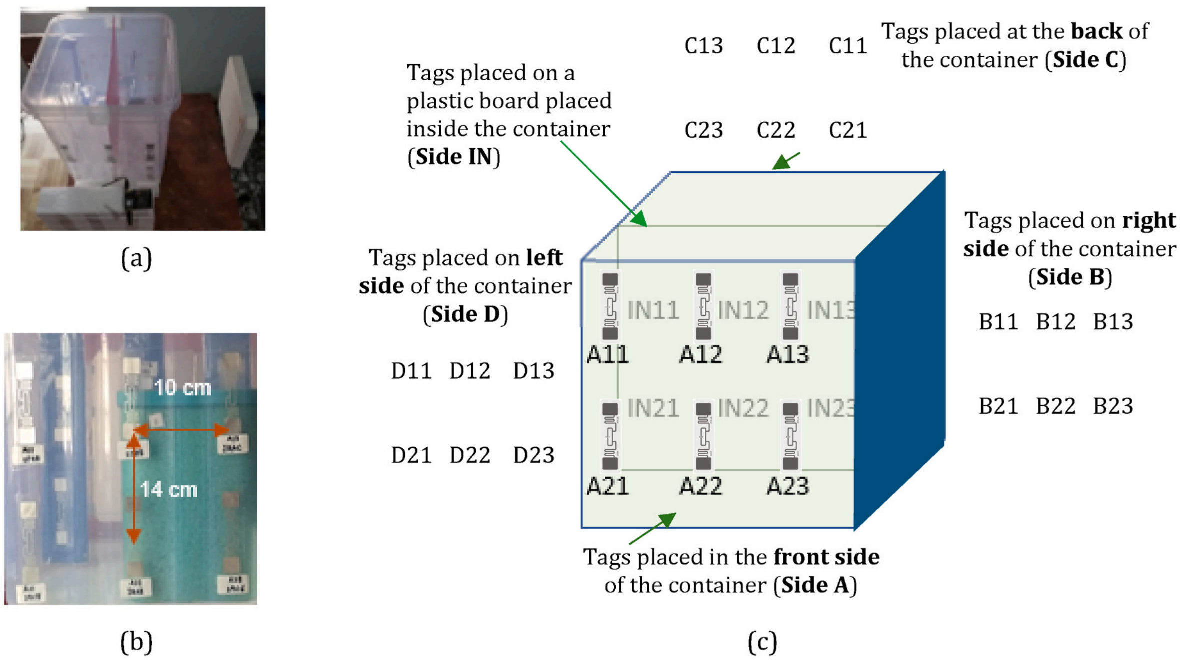

Fig. 6. (a) Actual setup of the experiment, (b) separation distance between RFID tags on the same side, and (c) RFID tags position of each RFID tag attached to the container.

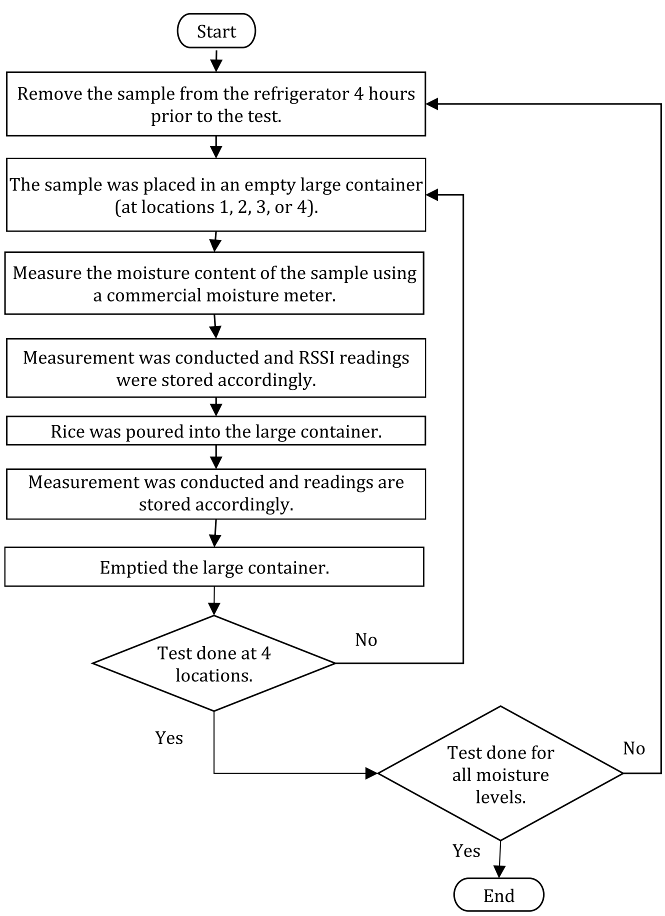

Fig. 7. Flowchart of the experimental test.

times, resulting in 192 observations per moisture condition at each location. Additionally, the test dataset comprised 64 readings per test. The dataset D obtained from the experimental test consists of n vectors of x_i, each associated with a value y_i indicating class membership (+1 or −1). The formal definition of the dataset used in this research is provided in Eq. (5), with x_i being a p-dimensional vector if it has p dimensions. The RSSI raw data from all tests were first pre-processed to remove null data and other misread information from the serial interface. The dataset was then prepared for training, testing, and validation of the model.

$$D = \left\{ (\mathbf{x}_i, y_i) \middle| \mathbf{x}_i \in \mathbb{R}^2, y_i \in \{-1, 1\} \right\}_{i=1}^{n} \tag{5}$$

5. Model development

Artificial intelligence elements, namely data science, big data, machine learning, and deep learning have gained popularity for classifying and predicting various phenomena (Liakos et al., 2018). The integration of artificial intelligence in modelling for rice moisture content prediction and spoilage localisation is advantageous due to its reliable predictions, minimal input data requirements (e.g., RSSI, temperature, humidity), and the lack of explicit knowledge of radio frequency parameters. This research employed four machine learning classification algorithms: random forest (RF), gradient boosting trees (GBT), support vector machine (SVM), and multilayer perceptron (MLP). Based on the performance metrics of these four machine learning algorithms, an ensemble-based approach incorporating data from the selected wireless technologies was incorporated into the algorithm to improve detection accuracy. Henceforth, the ensemble-based approach is named Ensemble-based Rice Moisture Content and Localisation, or eRMCL for short.

Each machine learning model was assessed across various input feature configurations. The configurations were defined as follows: 1 input feature (RSSI from RFID), 2 input features (RSSI from RFID and Zigbee), and 4 input features (RSSI from RFID, RSSI from Zigbee, temperature, and humidity). Notably, these configurations were designed to predict either the moisture content alone or both the moisture content and the sample location of the wet spot within the container. The number of input features and output labels is listed in Tables 2 and 3, respectively.

Each model was trained with different input and output

Table 2
Number of input features tested.

Number of input features	Input features
1	RSSI of RFID
2	RSSI of RFID, RSSI of ZIGBEE
4	RSSI of RFID, RSSI of ZIGBEE, temperature, humidity

Table 3
Number of output labels/class.

Number of output Labels/class	Output Labels/class
1	Moisture content
2	Moisture content, sample location

configurations. Initially, only RSSI values from RFID were used as input features to predict moisture levels. Subsequently, the model was trained with two input features (RSSI values from RFID and Zigbee) and one output label (moisture levels). Then, the model was trained with four input features (RSSI values from RFID, RSSI values from Zigbee, temperature, and humidity) and one output label (moisture levels). Similar configurations were used to train the model to predict both moisture levels and sample locations. Finally, different hyperparameters were tested to determine the optimal features for predicting rice moisture content and localising wet spots in storage.

5.1. Support vector machine

The functions for linear SVM, polynomial kernel SVM, Gaussian kernel SVM, and sigmoid kernel SVM are given by Eq. (6), Eq. (7), Eq. (8), and Eq. (9), respectively. In these equations, *K(x, y)* represents the kernel function value between vectors *x* and *y*, *c*, *d*, γ, and α are adjustable parameters specific to each kernel.

$$\text{The Linear Kernel SVM}: K(x,y) = \widehat{x}T * y \tag{6}$$

$$\text{The Polynomial Kernel SVM}: K(x,y) = (\widehat{x}T * y + \widehat{c)}d \tag{7}$$

$$\text{The Gaussian Kernel SVM}: K(x,y) = exp(-\gamma * \|x - y|\widehat{|}2) \tag{8}$$

$$\text{The Sigmoid Kernel SVM}: K(x,y) = tanh(\alpha * \widehat{x}T * y + c) \tag{9}$$

The introduction of kernels in SVM improves processing time by simplifying the calculations in high-dimensional spaces. Additionally, each SVM kernel type was used to determine the optimal values of the regularisation parameters *C* and gamma. The parameter *C* adjusts the margin between the separating hyperplanes, starting with an initial value of 1, which is then decreased accordingly. The gamma setting was also evaluated where lower gamma values ensure that points far from the hyperplane are considered in its tuning.

5.2. Random forest

Random forest is an ensemble technique that combines the results of multiple decision trees through majority voting to produce the final output. Initially, the RF model was trained under default condition provided by the Scikit-learn Library, which includes a max_depth of None, meaning that the tree will grow without limit as to obtain a better insight into the full predictive power of the RF model on the dataset. To enhance prediction accuracy, several hyperparameters were tuned: the number of trees before averaging the predictions (n_estimators), the maximum number of features considered for splitting a node (max_features), the loss function to measure the quality of the split (criterion), and the minimum number of leaves required to split an internal node (mini_sample_leaf). Entropy was used to measure the quality of splits, with data splits minimizing entropy until each node reached a minimum node size or a maximum tree depth. This process involved selecting a random subset of available variables for each node split. The estimate for the dependent variable is given by Eq. (10), where the entropy of a dataset is defined as the sum of the probability of each class multiplied by its logarithm. For binary classification, entropy ranges from 0 to 1, where *p* represents the dataset, *N* is the number of classes, and p_i is the frequency of class *i*.

$$Entropy\ (p) = -\sum_{i=1}^{N} p_i \log_2 p_i \tag{10}$$

5.3. Gradient boosting trees

Gradient boosting trees (GBT) combine predictions from multiple decision trees to generate the final predictions by using a boosting technique where subsequent predictors learn from the errors of the previous predictors. GBT employs a loss function minimised via gradient descent; hence, the name gradient boosting. Among the hyperparameters that need to be tuned for GBT are the number of estimators, learning rate, subsample, and maximum depth. The number of estimators must be carefully managed to avoid overfitting, while the learning rate determines the weight of each tree on the final prediction. These hyperparameters were carefully tuned to optimize the model.

5.4. Multilayer perceptron

MLP is a fundamental neural network model. Hyperparameters for MLP include the activation function, optimiser, and learning rate. The MLP was evaluated using various activation functions and optimizers. Its architecture consists of an input layer, one or more hidden layers, and an output layer. The output of the network, *y(t)* at output layer *m*, is given by Eq. (11):

$$y_t(t) = h\left(\sum_{j=1}^{n_{m-1}} w_{j1}^{m} v_j^{m-1}(t)\right); 1 \le l \le n_m \tag{11}$$

where,

n_k: number of nodes in the kth layer
n_m: number of nodes in the output layer
w: weights
h (•): activation function

5.5. eRMCL algorithm

The performance of each model was evaluated using the test data. An ensemble-based algorithm was introduced to enhance the prediction capability of the weak learners, thereby increasing prediction accuracy. Performance metrics such as accuracy, precision, recall, and F1-score were recorded for each model. The SVM, RF, GBT, and MLP models were considered weak learners, each running on the same dataset. The prediction output from each weak learner was then combined, and the prediction that received the maximum votes was selected as the final prediction output. Fig. 8 illustrates the flowchart of the Ensemble-based Rice Moisture Content and Localisation (eRMCL) algorithm.

The eRMCL algorithm, developed for moisture content determination and localisation of wet spots, is based on the concept of the random forest and gradient-boosting tree. The accuracy of each weak learner was compared, and the one with the highest accuracy was used to predict the moisture content and location of spoilage in the storage. Eq. (12) presents the basic functions of the eRMCL algorithm, with the expanded form, including the accuracy of the four weak learners, shown in Eq. (13). While the number of weak learner models could be increased, this would require more resources and computation time.

$$\mathrm{RMCL} = \max(Acc_{SVM}, Acc_{RF}, Acc_{GBT}, Acc_{MLP}) \tag{12}$$

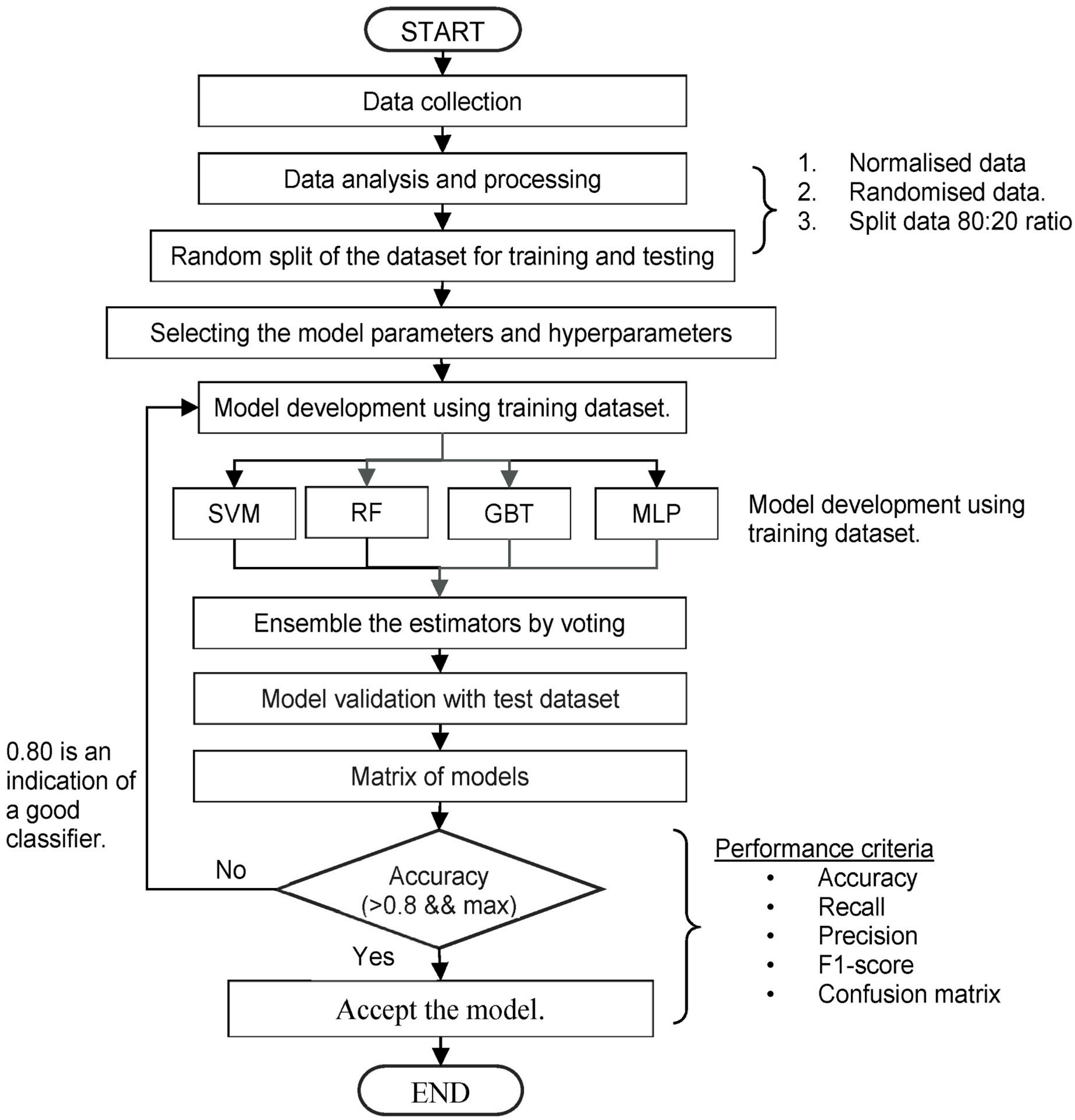

Fig. 8. eRMCL algorithm for moisture content determination and wet spot localisation.

where,

Acc_{SVM}: Prediction accuracy of SVM.

Acc_{RF}: Prediction accuracy of RF.

Acc_{GBT}: Prediction accuracy of GBT.

Acc_{ML}: Prediction accuracy of MLP.

Let $Acc_{SVM}=a$, $Acc_{RF}=b$, $Acc_{GBT}=c$, $Acc_{MLP}=d$

$$\mathrm{eRMCL} = \max(a, \max(b, \max(c, d)))$$

$$\mathrm{eRMCL} = \frac{a + \left(\frac{b + \left(\frac{c+d}{2} - \frac{|c-d|}{2}\right)}{2} + \frac{\left|b - \left(\frac{c+d}{2} + \frac{|c-d|}{2}\right)\right|}{2}\right)}{2} + \frac{a - \left(\frac{b + \left(\frac{c+d}{2} - \frac{|c-d|}{2}\right)}{2} + \frac{\left|b - \left(\frac{c+d}{2} + \frac{|c-d|}{2}\right)\right|}{2}\right)}{2} \tag{13}$$

5.6. Model evaluation

Common evaluation methods for classification models include accuracy, precision, recall, and F-measure (also known as the F1-score). These metrics are obtained from the confusion matrix, which measures classifier performance on real data. As shown in Fig. 9, the confusion matrix provides information on actual versus predicted classifications, indicating how often the classifier makes correct or incorrect predictions. True positives (TP) are data points classified as positive by the model and are the correct class, while false negatives (FN) are data points the model identifies as negative (identified as other class) but are actually in that class (incorrect classification).

Precision, also known as positive predictive value, quantifies the number of correct positive predictions made by the model. A Type I Error, represented by α, occurs when a true positive is incorrectly classified as negative. Precision is calculated using Eq. (14), where *TP* is true positive, *TN* is true negative, *FP* is false positive, and *FN* is false negative.

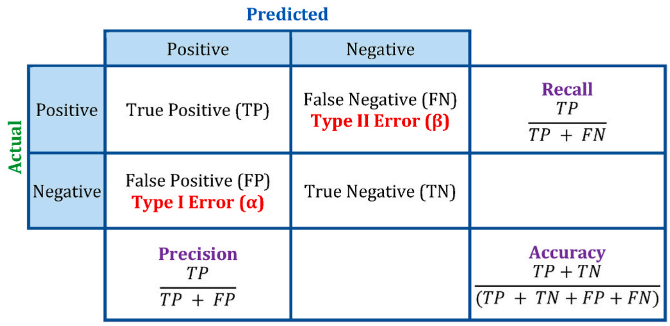

Fig. 9. Measures obtainable from the confusion matrix.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{14}$$

Recall, also known as sensitivity, is a metric that tells the number of correct positive predictions made out of all possible positive predictions. It indicates the proportion of missed positive predictions, with Type II Error represented by β that occurs when the wrong value/object is mistaken as true. The formula to calculate the recall is provided in Eq. (15).

$$\text{Recall or Sensitivity} = \frac{TP}{TP + FN} \tag{15}$$

Accuracy is the ratio of the number of items that have been correctly predicted/classified to the total number of times the items have been predicted/classified. It reflects the proportions of correct predictions, with a best accuracy is 1.0 and a worst value of 0.0. Accuracy is calculated as the number of correct predictions divided by the total number of predictions, as shown in Eq. (16).

$$\text{Accuracy} = \frac{TP + TN}{(TP + TN + FP + FN)} \tag{16}$$

The F1-score is the harmonic mean of recall and precision. When trying to optimize recall, the algorithm might end up predicting outputs that belong to the positive class but might also predict too many false positives, consequently leading to low precision. On the other hand, when optimising precision, the algorithm might end up predicting very few positive results (those with the highest probability of being positive) and the recall would have an exceptionally low value. Consequently, the F1 score aids in finding a compromise between recall and precision. The formula to calculate the F1 score is provided in Eq. (17).

$$\text{F1 score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{17}$$

6. Result and discussion

Given the consistent testbed and environmental parameters throughout the conducted tests, it was postulated that the location of each tag and environment does not affect the RSSI value. Tags located on Side C were unaffected by the changes in moisture content in rice or the sample location because they directly aligned to face the antenna; hence, no obstruction between the antenna and the tags on Side C. Conversely, tags on Sides A, B, D, and IN experienced fluctuations in RSSI values in response to variations in the location and moisture levels of rice. The reduction in RSSI values correlated with the increase in moisture content. For instance, tag B21 was affected when the sample was positioned at location 4, while tags D12, D13, D22, and D23 were notably influenced when the sample was located at location 1. Notably, tags on Side A and Side IN were mostly affected by the location of the sample. Consequently, the RSSI value for each RFID tag can help determine the moisture content and identify the location of the wet spots. Nevertheless, the implementation of a machine learning algorithm is imperative for automating the determination of moisture content and the precise localisation of wet spots. This approach enhances the efficiency and accuracy of the system in practical applications.

As previously mentioned, the number of input features in Tables 4–12 (the numerical values of 1, 2, and 4) represents the number of input features fed to the machine learning, where 1 represents the values of RSSI from RFID as the input feature, 2 represents two input features consisting of the RSSI from RFID and the RSSI from Zigbee and 4 represents four input features comprising RSSI from RFID, RSSI from Zigbee, temperature, and humidity. Based on Table 4, the Gaussian kernel SVM gives the highest accuracy among other SVM algorithms with an accuracy of 78% for one output feature. Likewise, the Gaussian kernel SVM also yielded a high accuracy of about 81.5% as shown in Table 5 when two output features were used. Similarly, the linear SVM also provided high accuracy of 72.5% and 81.5% for one output feature and two output features, respectively. The highest accuracy for each type of SVM algorithm is highlighted in bold font. Linear SVM and Gaussian kernel SVM have higher accuracy compared to other types of SVM kernels. In related research, Bains and Kalsi (2019) achieved a high prediction accuracy of about 74.58%, 91.53%, and 95.34% when using SVM, KNN, and Naïve Bayes, respectively, for wheat production prediction. However, it is unclear whether moisture content is included in the list of features. Besides that, results in Tables 4 and 5 suggest that the temperature and humidity did not help to increase the accuracy of the algorithm because accuracy decreases when the temperature and humidity are included as input features.

Similar to SVM, the RF was also tested with different configurations of input and output features. The highest accuracy yield by the RF algorithm for one output feature (moisture content) is about 89.9% as shown in Table 6. For two output features (moisture content and sample location), the accuracy of the RF algorithm increased to 95.4%, as highlighted in Table 7. Based on these results, the RF performs better than SVM for the classification of moisture content and the localisation of the grain spoilage wet spot. The result in Table 7 shows that RF is a good model for the measurement of moisture content and localisation of wet spots. This result is supported by de Oliveira Carneiro et al. (de Oliveira Carneiro et al., 2023) which considered RF as one of the best models for predicting the physicochemical quality in whole and defective rice grains for different moisture contents.

The GBT was also tested with different configurations of input and output features. Based on Tables 8 and 11, the GBT presents a high accuracy of about 94.8% for one output feature. However, the accuracy of gradient boosting dropped significantly to 69.7% for the two output features. This result indicates that the gradient boosting model is excellent for predicting rice moisture content, but it is not suitable for predicting the location of the wet samples. The GBT results show a stark contrast to the RF results.

MLP is one of the algorithms used in neural networks. Several types of activation functions and optimisers are available for MLPs. Different activation functions and optimisers yield different accuracies.

Table 4
Performance of SVM algorithm for one output feature.

Algorithm	Input features	Accuracy	Precision	Recall	F1-score
Linear SVM	1	0.682	0.72	0.74	0.72
	2	**0.725**	0.76	0.79	0.76
	4	0.611	0.66	0.78	0.65
SVM (polynomial kernel)	1	**0.652**	0.70	0.80	0.70
	2	0.643	0.69	0.82	0.69
	4	0.481	0.55	0.65	0.52
SVM (Gaussian kernel)	1	0.722	0.76	0.79	0.76
	2	**0.780**	0.81	0.83	0.80
	4	0.752	0.78	0.86	0.79
SVM (sigmoid kernel)	1	0.548	0.48	0.49	0.45
	2	**0.600**	0.52	0.61	0.49
	4	0.418	0.37	0.51	0.36

Table 5
Performance of SVM algorithm for two output features.

Algorithm	Input features	Accuracy	Precision	Recall	F1-score
Linear SVM	1	0.785	0.78	0.69	0.72
	2	**0.815**	0.81	0.76	0.76
	4	0.651	0.65	0.51	0.55
SVM (polynomial kernel)	1	0.571	0.57	0.50	0.51
	2	0.522	0.52	0.45	0.47
	4	0.330	0.33	0.22	0.25
SVM (Gaussian kernel)	1	**0.815**	0.81	0.75	0.77
	2	0.813	0.81	0.75	0.76
	4	0.548	0.55	0.47	0.49
SVM (sigmoid kernel)	1	0.663	0.66	0.57	0.58
	2	0.667	0.67	0.56	0.58
	4	0.524	0.52	0.36	0.41

Table 6
Performance of RF algorithm for one output feature.

Algorithm	Input feature	Accuracy	Precision	Recall	F1-score
RF	1	0.809	0.83	0.86	0.83
RF	2	**0.899**	0.91	0.91	0.91
RF	4	0.775	0.80	0.87	0.79

Table 7
Performance of RF algorithm for two output features.

Algorithm	Input feature	Accuracy	Precision	Recall	F1-score
RF	1	0.935	0.94	0.96	0.92
RF	2	**0.954**	0.95	0.98	0.94
RF	4	0.899	0.90	0.90	0.93

Therefore, the MLP algorithm was run several times using a different combination of activation functions and optimisers. The MLP performance in Table 10 indicates that the activation function 'logistic' and optimiser 'adam' yielded the highest accuracy of 75.6% for one output feature. On the other hand, for two output features, the MLP with the 'relu' activation function and the 'sgd' optimiser yielded the highest accuracy of 75.6% as depicted in Table 11.

In summary, the Gaussian kernel SVM gives the highest prediction accuracy among other SVM algorithms, with an accuracy of 78% for one output feature. Likewise, the Gaussian kernel SVM also yielded a high accuracy of about 81.5% when two output features were used. Similarly, the linear SVM also provided a high prediction accuracy of 72.5% and 81.5% for one output feature and two output features, respectively. From the evaluation and comparison between each algorithm, the GBT provides the highest accuracy of 94.8% for one output feature. Meanwhile, with 95.4% accuracy, the RF was the best algorithm for the two output features. As a result, the GBT is suitable for predicting the moisture content of rice. On the other hand, RF is a suitable algorithm for predicting moisture content while localising the location of the rice wet spot in storage. The performance matrix of each model in Table 4 through Table 11 shows low prediction accuracy. Based on these findings, the algorithm needs to be improvised so that it can commonly determine both the moisture content of rice and localise the location of the wet spot. Hence, the eRMCL method was introduced to improve the accuracy of the determination and localisation algorithms.

Table 8
Performance of GBT algorithm for one output feature.

Algorithm	Input features	Accuracy	Precision	Recall	F1-score
GBT	1	0.864	0.88	0.92	0.88
	2	**0.948**	0.95	0.96	0.95
	4	0.789	0.82	0.90	0.79

Table 9
Performance of GBT algorithm for two output features.

Algorithm	Input features	Accuracy	Precision	Recall	F1-score
GBT	1	**0.697**	0.70	0.60	0.63
	2	0.697	0.70	0.60	0.63
	4	0.667	0.67	0.62	0.63

Table 10
Performance of MLP algorithm for one output feature.

Activation function	Optimiser	Input feature	Accuracy	Precision	Recall	F1-score
relu	adam	1	0.699	0.74	0.75	0.74
		2	0.579	0.63	0.69	0.64
		4	0.669	0.71	0.67	0.67
	lbfgs	1	0.603	0.53	0.54	0.52
		2	0.701	0.61	0.61	0.60
		4	0.596	0.65	0.61	0.59
	sgd	1	0.754	0.79	0.83	0.77
		2	0.658	0.70	0.75	0.70
		4	0.688	0.73	0.71	0.68
logistic	adam	1	0.717	0.75	0.77	0.76
		2	**0.774**	0.80	0.83	0.80
		4	0.653	0.70	0.82	0.69
	lbfgs	1	0.539	0.60	0.66	0.60
		2	0.533	0.59	0.64	0.59
		4	0.513	0.57	0.68	0.60
	sgd	1	0.190	0.17	0.03	0.05
		2	0.190	0.17	0.03	0.05
		4	0.190	0.17	0.03	0.05
tanh	adam	1	0.592	0.64	0.67	0.65
		2	0.570	0.62	0.67	0.63
		4	0.576	0.63	0.77	0.64
	lbfgs	1	0.481	0.55	0.61	0.51
		2	0.658	0.70	0.66	0.68
		4	0.621	0.67	0.76	0.67
	sgd	1	0.554	0.61	0.62	0.60
		2	0.586	0.64	0.69	0.65
		4	0.690	0.73	0.79	0.74
identity	adam	1	0.597	0.64	0.68	0.64
		2	0.608	0.64	0.73	0.66
		4	0.579	0.63	0.74	0.62
	lbfgs	1	0.684	0.72	0.75	0.73
		2	0.688	0.73	0.72	0.72
		4	0.469	0.41	0.62	0.39
	sgd	1	0.592	0.64	0.68	0.64
		2	0.575	0.63	0.68	0.63
		4	0.616	0.66	0.77	0.67

The eRMCL works by comparing the accuracy of each model and the output class/label produced from the model that yielded the highest accuracy will be selected as the predicted outcome. However, running through four models will increase the prediction time, hence, instead of going through all four models, the eRMCL will only use a combination of two models. The best combination is provided in Tables 12 and 13.

The combination of RF and GBT yielded the highest accuracy, with an accuracy greater than or equal to 0.970. Similarly, the same combination also yielded the highest accuracy when the algorithm was used to determine the moisture content and location of the wet spot in storage. This outcome was possible due to the basic principle of RF and GBT, which operate based on decision probability and voting. On the other hand, the combination of SVM and MLP yielded lower accuracy compared to the RF and GBT because SVM operates particularly through boundary creation, known as hyperplane, between each group. Since the collected data was not linearly separable between each moisture level and location, it led to low accuracy for both SVM and MLP algorithms. When the eRMCL method was used to determine moisture content only, the highest accuracy was yielded with the combination of RF and GBT. Similarly, the same combination also yielded the highest accuracy when the algorithm was used to determine the moisture content and location of the wet spot in storage. According to Al Azies et al. (Al Azies et al., 2019) and Yang et al. (Der Yang et al., 2021), different machine learning triumphs over others depending on the test conditions. Among the

Table 11
Performance of GBT algorithm for one output feature.

Activation function	Optimiser	Input feature	Accuracy	Precision	Recall	F1-score
relu	adam	1	0.604	0.60	0.50	0.53
		2	0.635	0.64	0.58	0.59
		4	0.670	0.67	0.61	0.60
	lbfgs	1	0.543	0.54	0.45	0.46
		2	0.557	0.56	0.51	0.49
		4	0.596	0.60	0.58	0.55
	sgd	1	0.696	0.70	0.58	0.62
		2	0.640	0.64	0.51	0.56
		4	**0.756**	0.76	0.65	0.68
logistic	adam	1	0.593	0.59	0.59	0.56
		2	0.506	0.51	0.49	0.42
		4	0.370	0.37	0.29	0.27
	lbfgs	1	0.050	0.05	0.01	0.02
		2	0.080	0.08	0.02	0.03
		4	0.048	0.05	0.00	0.01
	sgd	1	0.048	0.05	0.00	0.00
		2	0.095	0.10	0.01	0.02
		4	0.048	0.05	0.00	0.00
tanh	adam	1	0.613	0.61	0.52	0.55
		2	0.522	0.52	0.50	0.49
		4	0.478	0.48	0.36	0.39
	lbfgs	1	0.542	0.54	0.39	0.44
		2	0.457	0.46	0.43	0.40
		4	0.508	0.51	0.45	0.43
	sgd	1	0.624	0.62	0.53	0.54
		2	0.620	0.62	0.52	0.55
		4	0.539	0.54	0.46	0.47
identity	adam	1	0.493	0.49	0.40	0.38
		2	0.705	0.70	0.59	0.63
		4	0.477	0.48	0.45	0.43
	lbfgs	1	0.605	0.60	0.54	0.52
		2	0.606	0.61	0.54	0.53
		4	0.509	0.51	0.39	0.43
	sgd	1	0.536	0.54	0.46	0.43
		2	0.664	0.66	0.58	0.58
		4	0.467	0.47	0.40	0.38

Table 12
Moisture content determination using the eRMCL.

		SVM (Gaussian kernel)	RF	GBT	MLP
SVM (gaussian kernel)	Accuracy	0.780	0.857	0.856	0.855
	Precision	0.810	0.870	0.920	0.870
	Recall	0.830	0.860	0.860	0.850
	F1-score	0.800	0.860	0.870	0.860
RF	Accuracy	0.857	0.899	0.971	0.898
	Precision	0.870	0.910	0.980	0.900
	Recall	0.860	0.910	0.970	0.900
	F1-score	0.860	0.910	0.970	0.900
GBT	Accuracy	0.856	0.971	0.948	0.923
	Precision	0.920	0.980	0.950	0.940
	Recall	0.860	0.970	0.960	0.920
	F1-score	0.870	0.970	0.950	0.920
MLP	Accuracy	0.855	0.898	0.923	0.774
	Precision	0.870	0.900	0.940	0.800
	Recall	0.850	0.900	0.920	0.830
	F1-score	0.860	0.900	0.920	0.800

selected algorithms, RF serves as the best rice moisture content determination and spoilage wet spot localisation. However, if the requirement only needs the determination of moisture content, the gradient-boosting trees is preferable.

The heatmap in Figs. 10 and 11 shows the number of correct and incorrect predictions when the test dataset was used to validate the eRMCL algorithm. Based on Fig. 10, about 39 data were incorrectly classified into different moisture content levels. The other 38 had an actual moisture content of 14%, but they were incorrectly classified as 20%. Meanwhile, 1 out of 39 data had a moisture content of 25%, but it

Table 13
Moisture content and localisation of wet spots using the eRMCL.

		SVM (Gaussian kernel)	RF	GBT	MLP
SVM (linear kernel)	Accuracy	0.815	0.909	0.747	0.547
	Precision	0.810	0.990	0.930	0.710
	Recall	0.760	0.910	0.750	0.550
	F1-score	0.760	0.930	0.800	0.600
RF	Accuracy	0.909	0.954	0.995	0.755
	Precision	0.990	0.950	1.000	0.980
	Recall	0.910	0.980	0.990	0.760
	F1-score	0.930	0.940	0.990	0.820
GBT	Accuracy	0.747	0.995	0.697	0.607
	Precision	0.930	1.000	0.700	0.760
	Recall	0.750	0.990	0.600	0.610
	F1-score	0.800	0.990	0.630	0.640
MLP	Accuracy	0.547	0.755	0.607	0.756
	Precision	0.710	0.980	0.760	0.760
	Recall	0.550	0.760	0.610	0.650
	F1-score	0.600	0.820	0.640	0.680

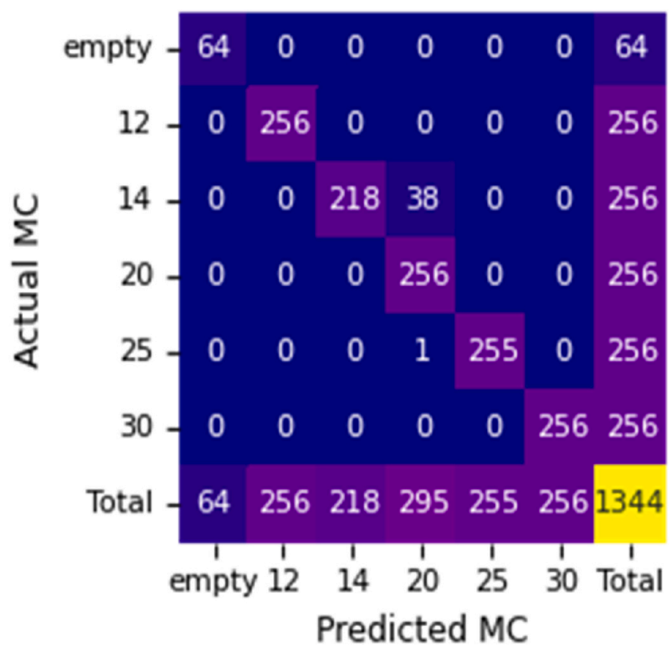

Fig. 10. Heatmap of a confusion matrix for moisture content determination using eRMCL.

was incorrectly classified as 20%. The incorrect classification was due to the RSSI pattern for the 39 data being more similar to 20% rather than their actual values. This incorrect classification can be further reduced by collecting more training datasets in the future. On the other hand, the heatmap in Fig. 11 shows that there are only 7 data that were incorrectly classified when the algorithm was used to determine the moisture content and the localisation of the rice wet spot. The 7 data have a moisture content of 25% and are at location 1 in storage, but they were incorrectly classified as 20% and are at location 1. This suggests that the eRMCL algorithm could find the wet spot perfectly even though there are some misclassifications of the moisture content of rice. However, a test on a larger quantity of rice needs to be conducted in the future to ensure that the algorithm applicable to any amount of rice stored.

The RSSI values from RFID tags around the storage were used to predict the affected wet rice. Based on the literature review, it is expected that signals, and hence RSSI, propagating through samples with higher moisture content, will experience higher levels of attenuation compared to those from samples with lower moisture content (Röbesaat et al., 2017). Due to absorption, the RSSI values for signals that have

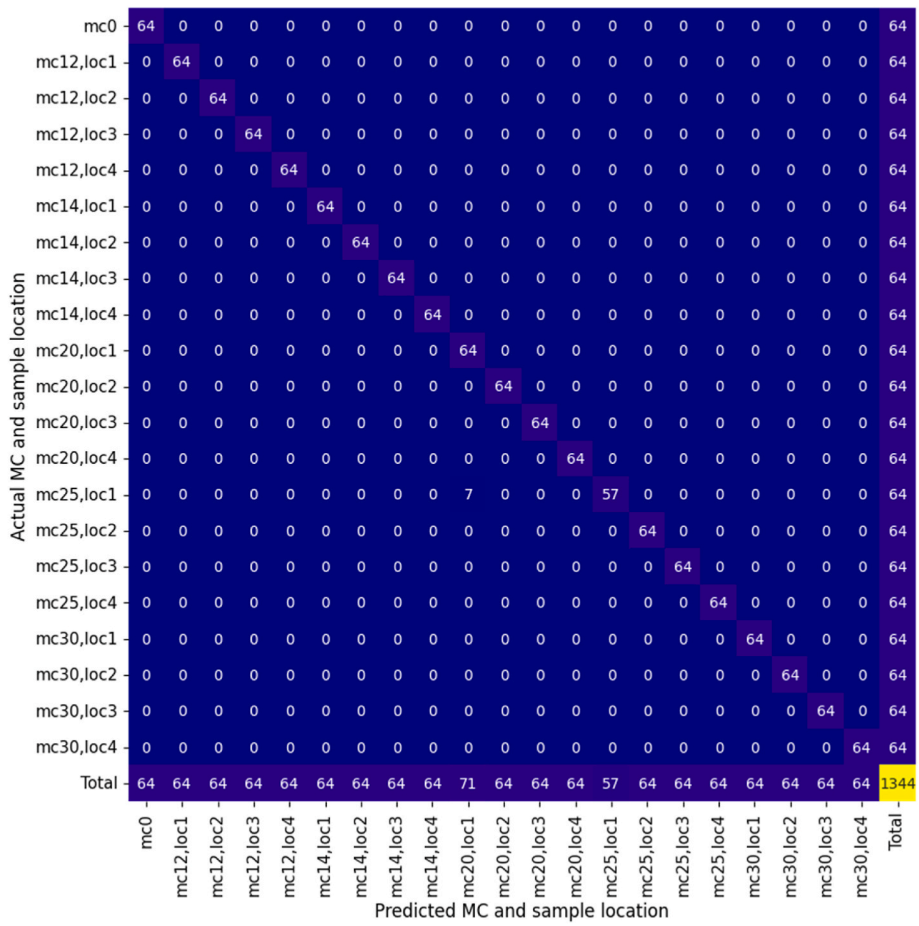

Fig. 11. Heatmap of a confusion matrix for moisture content determination and localisation of wet spot using eRMCL.

propagated through samples with high moisture content decreased, and this can be used to determine the location of the sample within the container.

This research introduces a method to determine the moisture content in rice and the localisation of the rice wet spot (wet area of rice) using machine learning. The method utilizes the measured RSSI value obtained from commercially available wireless technologies, namely Wi-Fi, LoRa, Zigbee and RFID technology, through a series of experimental tests. The characterization test was conducted to identify suitable commercially off-the-shelf wireless devices that can be used to measure the moisture content while localising the location of the wet spot simultaneously. The wireless devices or technologies identified from the characterization test are Zigbee and RFID. Both were then used to collect data for different moisture content levels and locations of wet spots in the storage. A dataset was formed and fed to four types of supervised machine learning: Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting Trees (GBT), and Multilayer Perceptron (MLP). Finally, based on the performance metric from the algorithms, the eRMCL algorithm inspired by the ensemble method was introduced to improve the algorithm for the determination and localisation of rice moisture content.

The finding in this study is also supported by the findings of Makky et al. (2019), who used Partial Least Squares (PLS) regression to determine the moisture content of two types of rice cultivars and found that the PLS regression yields R^2 of 0.6–0.8 depending on the physical properties of the rice. Similarly, each weak learner model in this research also yields an accuracy above 0.6. However, more research is needed to determine whether the eRMCL may be applied to other varieties of grain, such as glutinous rice, brown rice, and long rice, which have different physical characteristics. While a lot of studies have focused on various aspects of the topic or subject area, none of them addressed this particular research idea of localising the wet spot in storage using radio waves.

7. Conclusion

This study introduces a novel approach for detecting moisture content and localizing wet spots in rice storage by using RSSI values from RFID and Zigbee technologies. Among the commercially available off-the-shelf wireless technologies, these off-the-shelf wireless technologies demonstrate a clear correlation with increasing moisture content. The Random Forest (RF) algorithm achieved 95.4% accuracy for dual outputs, while Gradient Boosting Trees (GBT) reached 94.8% for moisture prediction alone. The ensemble eRMCL approach, combining RF and GBT, yielded the highest accuracy, exceeding 97%. In contrast, SVM and MLP showed moderate performance, with the Gaussian SVM achieving up to 81.5% accuracy. Notably, the results suggest that adding temperature and humidity features decreases model accuracy. These findings align with prior studies and demonstrate improved grain monitoring efficiency, though future research should explore broader applications across different grain types.

To the best of the authors' knowledge, there is yet a study done to investigate the use of a dual-frequency band using commercially available technology and an ensemble algorithm to improve the detection of grain moisture content in bulk density during storage and the localisation of wet spots for spoilage prediction. Hence, this research can be further expanded to cover other environmental conditions and there are wide areas of research that can be improved in the future. This research was conducted in a plastic and square container to eliminate uncertainty due to the reflection of radio waves through a metal container. Thus, future work could investigate the effect of using different materials for the container (metal or plastic), the shape of the container, the size of the container, increasing the number of samples, and also increasing the amount of rice in the large storage. In addition, research can venture into a more complex environment, such as the implementation of this technique in an actual industrial environment. The research is scalable for industrial applications, as the number of passive RFID tags can be increased to cover a larger container size. Future studies should investigate Wi-Fi for larger rice quantities and LoRa for industrial-grade silos.

CRediT authorship contribution statement

Noraini Azmi: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Latifah Munirah Kamarudin:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization. **Ahmad Shakaff Ali Yeon:** Writing – review & editing. **Ammar Zakaria:** Writing – review & editing, Funding acquisition. **Syed Muhammad Mamduh Syed Zakaria:** Writing – review & editing, Funding acquisition. **Hiromitsu Nishizaki:** Writing – review & editing, Supervision. **Latifah Mohamed:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Xiaoyang Mao:** Writing – review & editing. **Mohd Hafiz Fazalul Rahiman:** Writing – review & editing, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

LATIFAH MUNIRAH BINTI KAMARUDIN reports financial support was provided by M1inistry of Higher Education (MOHE) Malaysia. LATIFAH MUNIRAH BINTI KAMARUDIN reports financial support was provided by Research Management and Innovation Centre Universiti Malaysia Perlis. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

This research work was funded by the Ministry of Higher Education (MOHE) Malaysia under grant Transdisciplinary Research Grant Scheme (Grant No.: TRGS/1/2018/UNIMAP/02/4/1) titled "Determination and Characterization of Radio Frequency Signal for Moisture Sensing in Rice Grain Silos". The authors want to thank Universiti Malaysia Perlis (UniMAP) for providing postdoctoral research grants 9001-00754 that helped the realisation of this paper.

References

Aichholzer, A., Schuberth, C., Mayer, H., Arthaber, H., 2018. Microwave testing of moist and oven-dry wood to evaluate grain angle, density, moisture content and the dielectric constant of spruce from 8 GHz to 12 GHz. Eur. J. Wood Wood Prod. 76 (1), 89–103. https://doi.org/10.1007/s00107-017-1203-x.

Al Azies, H., Trishnanti, D., Mustikawati P H, E., 2019. Comparison of kernel support vector machine (SVM) in classification of human development index (HDI). IPTEK J. Proc. Ser. (6), 53. https://doi.org/10.12962/j23546026.y2019i6.6339.

Almaleeh, A.A., Zakaria, A., Kamarudin, L.M., Rahiman, M.H.F., Ndzi, D.L., Ismail, I., 2022. Inline 3D volumetric measurement of moisture content in rice using regression-based ML of RF tomographic imaging. Sensors 22 (1), 405. https://doi.org/10.3390/s22010405.

Asefi, M., Jeffrey, I., LoVetri, J., Gilmore, C., Card, P., Paliwal, J., 2015. Grain bin monitoring via electromagnetic imaging. Comput. Electron. Agric. 119, 133–141. https://doi.org/10.1016/j.compag.2015.10.016.

Asefi, M., Gilmore, C., Jeffrey, I., LoVetri, J., Paliwal, J., 2017. Detection and continuous monitoring of localised high-moisture regions in a full-scale grain storage bin using electromagnetic imaging. Biosyst. Eng. 163 (2017), 37–49. https://doi.org/10.1016/j.biosystemseng.2017.08.015.

Azmi, N., 2016. Performance analysis of IEEE802.15.4 wireless sensor network to mitigate IEEE802.11 wireless local area network interference[Unpublished bachelor's thesis]. Universiti Malaysia Perlis, p. 142.

Azmi, N., et al., 2014. Interference issues and mitigation method in WSN 2.4GHz ISM band: a survey. In: 2014 2nd International Conference on Electronic Design, ICED 2014, pp. 403–408. https://doi.org/10.1109/ICED.2014.7015839. Penang, Malaysia.

Azmi, N., et al., 2018. Design and development of multi-transceiver lorafi board consisting LoRa and esp8266-wifi communication module. In: IOP Conference Series: Materials Science and Engineering. https://doi.org/10.1088/1757-899X/318/1/012051.

Bains, S.K., Kalsi, S., 2019. Analysis of wheat production using Naïve Bayes classifier. Int. J. Comput. Appl. 178 (14), 38–41. https://doi.org/10.5120/ijca2019918908.

Balasubramanian, D., 2011. Physical properties of raw cashew nut. J. Agric. Eng. Res. 78 (3), 291–297. https://doi.org/10.1006/jaer.2000.0603.

Basati, Z., Rasekh, M., Abbaspour-Gilandeh, Y., 2018. Using different classification models in wheat grading utilizing visual features. Int. Agrophys. 32 (2), 225–235. https://doi.org/10.1515/intag-2017-0008.

Berkley, R., 2016. Measuring Soil Moisture with Radio-Waves. The Ohio State University. https://doi.org/10.3929/ethz-b-000238666.

Chen, C., 2003. Evaluation of air oven moisture content determination methods for rough rice. Biosyst. Eng. 86 (4), 447–457. https://doi.org/10.1016/j.biosystemseng.2003.08.010.

Coşkun, M.B., Yalçin, I., Özarslan, C., 2006. Physical properties of sweet corn seed (Zea mays saccharata Sturt.). J. Food Eng. 74, 523–528. https://doi.org/10.1016/j.jfoodeng.2005.03.039.

de Oliveira Carneiro, L., et al., 2023. Characterizing and predicting the quality of milled rice grains using machine learning models. AgriEngineering 5 (3), 1196–1215. https://doi.org/10.3390/agriengineering5030076.

Der Yang, M., et al., 2021. Assessment of grain harvest moisture content using machine learning on smartphone images for optimal harvest timing. Sensors 21 (17). https://doi.org/10.3390/s21175875.

Freeman, S.A., Kelley, K.W., Maier, D.E., Field, W.E., 1998. Review of entrapments in bulk agricultural materials at commercial grain facilities. J. Saf. Res. 29 (2), 123–134. https://doi.org/10.1016/S0022-4375(98)00008-5.

Gilmore, C., Asefi, M., Paliwal, J., LoVetri, J., 2017. Industrial scale electromagnetic grain bin monitoring. Comput. Electron. Agric. 136, 210–220. https://doi.org/10.1016/j.compag.2017.03.005.

Girmay, M., Maglogiannis, V., Naudts, D., Aslam, M., Shahid, A., Moerman, I., 2023. Technology recognition and traffic characterization for wireless technologies in ITS band. Veh. Commun. 39, 100563 https://doi.org/10.1016/j.vehcom.2022.100563.

Grabe, D.F., 2015. Measurement of seed moisture. In: Seed Moisture, vol. 14, pp. 69–92. https://doi.org/10.2135/cssaspecpub14.c5, 8448.

Hassan, N., et al., 2013. Rice husk truncated pyramidal microwave absorber using quadruple p-spiral split ring resonator (QPS-SRR). Aust. J. Basic Appl. Sci. 7 (3), 56–63.

Hossain, M.A.A.A., Awal, M.A.A.A., Ali, M.R.R., Alam, M.M.M.M., 2016. Use of moisture meter on the post-harvest loss reduction of rice. Progress. Agric. 27 (4), 511–516 [Online]. Available. https://www.k-state.edu/phl/what-we-do/Moisture. (Accessed 17 January 2019).

Jusoh, M., Abbas, Z., Lee, K., You, K., Norimi, A., 2011. Determination of moisture content in mortar at near relaxation frequency 17 GHz. Meas. Sci. Rev. 11 (6), 203–206. https://doi.org/10.2478/v10048-011-0031-0.

Kandala, C.V.K.V.K., Butts, C.L.L., Lamb, M.C.C., 2013. Moisture content determination for in-shell peanuts with a low-cost impedance analyzer and capacitor sensor. Trans. ASABE (Am. Soc. Agric. Biol. Eng.) 51 (4), 1377–1381. https://doi.org/10.13031/2013.25221.

Kok, Y.Y., Li, L.Y., Chen, S.Y., Hou, K.M., Chia, Y.L., 2017. Physical and chemical characterization of rice using microwave and laboratory methods. In: Rice. InTech, pp. 81–99. https://doi.org/10.5772/66001.

Kraszewski, A.W.W., Trabelsi, S., Nelson, S.O.O., 1998. Comparison of density-independent expressions for moisture content determination in wheat at microwave frequencies. J. Agric. Eng. Res. 71 (3), 227–237. https://doi.org/10.1006/jaer.1998.0320.

Le-Huu, H., Seo, C., 2023. Dual-band transmitter for wireless power transfer to biomedical implant applications. IEEE Antenn. Wireless Propag. Lett. 22 (6), 1461–1465. https://doi.org/10.1109/LAWP.2023.3246163.

Li, C., Zhao, C., Ren, Y., He, X., Yu, X., Song, Q., 2022. Microwave traveling-standing wave method for density-independent detection of grain moisture content. Measurement 198 (April), 111373. https://doi.org/10.1016/j.measurement.2022.111373.

Liakos, K.G., Busato, P., Moshou, D., Pearson, S., Bochtis, D., 2018. Machine learning in agriculture: a review. Sensors 18 (8), 1–29. https://doi.org/10.3390/s18082674.

Libelium, 2014. Waspmote Quickstart Guide, p. 19.

Liu, Z., Wu, Z., Zhang, Z., Wu, W., Li, H., 2015. Research on online moisture detector in grain drying process based on V/F conversion. Math. Probl Eng. 10 https://doi.org/10.1155/2015/565764.

Maier, D.E., Channaiah, L.H., 2010. Monitoring carbon dioxide concentration for early detection of spoilage in stored grain. In: 10th International Working Conference on Stored Product Protection Monitoring, pp. 505–509. https://doi.org/10.5073/jka.2010.425.332.

Makky, M., Santosa, Putri, R.E., Nakano, K., 2019. Determination of moisture content in rice using non-destructive short-wave near infrared spectroscopy. In: Proceedings of the 2nd International Conference on Biosciences and Medical Engineering

(ICBME2019), pp. 1–8. https://doi.org/10.1063/1.5125518 (Australian Academy of Science), the I. I. Rabi Prize (APS), the European Frequency and Time Forum Award, the Carl Zeiss Research Award, the William F. Meggers Award and the Adolph Lomb Medal (OSA).
Moldenhauer, L., Helmerich, R., Köppe, E., Wittmann, J., 2018. Experimental modeling approach for determining the moisture damping exponent of a bluetooth low energy signal in moist building material. Mater. Today Proc. 5 (13), 26609–26614. https://doi.org/10.1016/j.matpr.2018.08.124.
Müller, A., et al., 2022. Rice drying, storage and processing: effects of post-harvest operations on grain quality. Rice Sci. 29 (1), 16–30. https://doi.org/10.1016/j.rsci.2021.12.002.
Nagai, Y., Guo, J., Orlik, P., Sumi, T., Rolfe, B.A., Mineno, H., 2021. Sub-1 GHz frequency band wireless coexistence for the internet of things. IEEE Access 9, 119648–119665. https://doi.org/10.1109/ACCESS.2021.3107144.
Nath, D.K., et al., 2017. Non-destructive methods for the measurement of moisture contents - a review. Sens. Rev. 37 (1), 71–77. https://doi.org/10.1108/SR-01-2016-0032.
Röbesaat, J., Zhang, P., Abdelaal, M., Theel, O., 2017. An improved BLE indoor localization with Kalman-based fusion: an experimental study. Sensors 17 (5), 1–26. https://doi.org/10.3390/s17050951.
Sacilik, K., Öztürk, R., Keskin, R., 2003. Some physical properties of hemp seed. Biosyst. Eng. 86 (2), 191–198. https://doi.org/10.1016/S1537-5110(03)00130-2.
ThingMagic, 2016. Universal Reader Assistant User Guide.
Toba, T., Kitagawa, A., 2011. Wireless moisture sensor using a microstrip antenna. J. Sens. 2011 https://doi.org/10.1155/2011/827969.
Trabelsi, S., Kraszewski, A.W., Nelson, S.O., 2001. New calibration technique for microwave moisture sensors. IEEE Trans. Instrum. Meas. 50 (4), 877–881. https://doi.org/10.1201/b18301-6.
Trabelsi, S., Lewis, M.A., Nelson, S.O., 2016. Microwave moisture meter for in-shell peanut kernels. Food Control 66, 283–290. https://doi.org/10.1016/j.foodcont.2016.02.016.
Vallejos, C., Grote, W., 2009. Wood moisture content measurement at 2.45 GHz. In: SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC 2009). IEEE, pp. 221–225. https://doi.org/10.1109/IMOC.2009.5427593.
Wang, H., et al., 2022. Impact of long-term storage on multi-scale structures and physicochemical properties of starch isolated from rice grains. Food Hydrocolloids 124 (PA), 107255. https://doi.org/10.1016/j.foodhyd.2021.107255.
Wee, F.H., Soh, P.J., Suhaizal, A.H.M., Nornikman, H., Ezanuddin, A.A.M., 2009. Free space measurement technique on dielectric properties of agricultural residues at microwave frequencies. In: International Microwave & Optoelectronics Conference (IMOC 2009), pp. 183–187. https://doi.org/10.1109/IMOC.2009.5427603.
Yang, W., Wang, X., Cao, S., Wang, H., Mao, S., 2018a. Multi-class wheat moisture detection with 5GHz Wi-Fi: a deep LSTM approach. In: Proceedings - International Conference on Computer Communications and Networks, ICCCN. IEEE. https://doi.org/10.1109/ICCCN.2018.8487423.
Yang, W., Wang, X., Song, A., Mao, S., 2018b. Wi-Wheat: contact-free wheat moisture detection with commodity WiFi. In: IEEE International Conference on Communications. IEEE. https://doi.org/10.1109/ICC.2018.8423034.
Yigit, E., Duysak, H., 2022. Determination of flowing grain moisture contents by machine learning algorithms using free space measurement data. IEEE Trans. Instrum. Meas. 71, 1–8. https://doi.org/10.1109/TIM.2022.3165740.
Zareiforoush, H., Komarizadeh, M.H.H., Alizadeh, M.R.R., 2009. Effect of moisture content on some physical properties of paddy grains. Res. J. Appl. Sci. Eng. Technol. 1 (3), 132–139.
Zhang, H., Wang, D., Jian, F., 2020. Movement and distribution of Sitophilus zeamais adults and relationship between their density and trapping frequency in wheat bulks under different grain temperatures and moisture contents. J. Stored Prod. Res. 87, 101590 https://doi.org/10.1016/j.jspr.2020.101590.