



# OPEN Development of a bench system with capacitive sensor, sample compression, and TinyML for iron ore moisture measurement

Érica S. Pinto<sup>1,9</sup>, Saulo N. Matos<sup>2,3,4,9</sup>✉, Matheus Neiva<sup>1,5</sup>, Gabriel A. Santos<sup>1,5</sup>, Leandro S. Marcolino<sup>3,8</sup>, Jó Ueyama<sup>2</sup>, Thiago A. M. Euzébio<sup>5,6,7</sup>✉, Gustavo Pessin<sup>4,5</sup>, Philip V. Pritzelwitz<sup>4</sup> & Alan Kardek Rêgo Segundo<sup>1,5</sup>

In the mineral sector, many processes use water for ore beneficiation processes. A lack of sensing or control of water content can lead to operational problems in various mineral processing operations, especially in ore transport. Current instrumentation systems are either slow or inaccurate. Therefore, a novel bench system was developed to address this gap by achieving a fast response time and improved accuracy. The developed instrument measures the ore moisture by using the real-dual-frequency method (RDFM) to assess the ore's electrical conductivity and relative permittivity. Additionally, it takes into account the bulk density, the bench chamber level, and the compress torque. All these variables are used to create a tiny machine-learning (TinyML) model that evaluates the ore's moisture with a low time response. This process is done while the ore sample is compressed to reduce air bubbles inside the samples and improve measurement. Experiments were performed using the bench system in a mining company's physical analysis laboratory. The instrument was utilized to measure the moisture content in the ore, leading to the development of a dataset used to train and validate various tree-based tinyML models. The results indicate that ore compression enhances accuracy and that decision trees are effective for estimating moisture with a quicker response time.

An accurate measurement of process variables is essential across various industrial sectors. In ore mining operations, for example, the ore characteristics significantly impact stages such as exploration, processing, and transportation<sup>1</sup>. Among these characteristics, ore moisture is particularly important. Therefore, measuring the moisture content in iron ore is crucial for ensuring product quality and safety, determining pricing, and optimizing both processing and transportation operations<sup>2</sup>.

High-moisture ore materials can jeopardize the ore handling processes. Transporting, storing, and preparing the ore for processing becomes problematic when the ore is wet. For example, it can stick on the conveyor belt when transporting and even cause avalanches in feeders and piles<sup>3</sup>. Ore beneficiation processes are also affected. Moisture agglomerates fine ore particles, reducing the classification and comminution process efficiency<sup>4</sup>. Furthermore, it directly affects the transport of the final product by shipping. Millions of tons of iron ore are transported by sea annually, and the amount of water compromises the stability of ships to the point that there is a risk of capsizing because of cargo liquefaction<sup>2,5,6</sup>.

Between 2006 and 2016, a minimum of eight reported incidents involving bulk carriers may have been caused by shifting iron ore cargo<sup>6</sup>. Depending on the ore type, values around 8.5 to 9.5% of water already represent a borderline percentage, as highlighted by Clout and Manuel<sup>7</sup>. This threshold can render sea transportation hazardous, compounding the issue of transporting excess water with no commercial value.

<sup>1</sup>Programa de Pós-Graduação em Ciência da Computação, Universidade Federal de Ouro Preto, Ouro Preto, Minas Gerais, Brasil. <sup>2</sup>Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo (ICMC-USP), São Carlos, Brasil. <sup>3</sup>School of Computing and Communications, Lancaster, UK. <sup>4</sup>Instituto Tecnológico Vale, Ouro Preto, Minas Gerais, Brasil. <sup>5</sup>Programa de Pós-Graduação em Instrumentação, Controle e Automação de Processos de Mineração, Universidade Federal de Ouro Preto e Instituto Tecnológico Vale, Ouro Preto, Minas Gerais, Brasil. <sup>6</sup>Helmholtz-Zentrum Dresden-Rossendorf, Dresden, Germany. <sup>7</sup>Virtus-CC, Campina Grande, Brasil. <sup>8</sup>College of Engineering and Computer Science, VinUniversity, Gia Lâm, Hanoi, Vietnam. <sup>9</sup>Érica S. Pinto and Saulo N. Matos have contributed equally to this work. ✉email: saulo.matos@usp.br; t.melo-euzebio@hzdr.de

The inability to effectively manage the excess water can also negatively affect the commercial value of iron ore. Buyers and sellers of iron ore consider moisture content as a key factor in determining prices; generally, higher moisture content results in lower prices<sup>2</sup>.

The standard and most common method for determining moisture content in mining environments is the gravimetric method, also known as the drying oven method. This direct and absolute technique involves drying the material under test (MUT) in a temperature-controlled chamber at 105°C until it almost reaches a constant mass, which typically takes at least 4 hours<sup>8,9</sup>. Along with prolonged response times, this method is sample-destructive, highlighting the need for a more effective moisture measurement technique. In addition, this technique is labor-intensive, requiring personnel to weigh the samples before drying, place them in the oven, weigh them again after drying, and record the masses.

The current utilization of manual gravimetric drying methods for moisture determination in most mineral processing plants falls short in providing the necessary, timely, and accurate information essential for automatic control<sup>10</sup>. Researchers have suggested online methods to evaluate the material moisture. For instance, the use of microwave-based analyzers for moisture measurement<sup>11,12</sup>. However, using microwaves has limitations due to high costs, imprecision in industrial environments, and challenges in supporting, calibrating, and maintaining this type of equipment<sup>13</sup>.

There are also image-based approaches<sup>14</sup>. However, it lacks industrial validation. Cameras face challenges in industrial settings due to dust, rain, and poor lighting conditions, which negatively impact image processing accuracy<sup>1,15</sup>. Moreover, online methods can struggle to achieve high accuracy in practice due to variable presentation and sensor drift or degradation over time; consequently, plants often rely on more accurate offline systems.

Capacitive sensors are a viable option for measuring the moisture content of bulk materials by assessing the complex impedance of the MUT<sup>16,17</sup>. However, this technique is limited by its tendency to produce inaccurate measurements when the material contains air gaps. To overcome this issue, we propose an innovative bench system that efficiently measures the moisture content of iron ore samples with the bulk density standardized by the machine, allowing for moisture measurement in a compressed state. The system utilizes the real-dual-frequency method (RDFM) available in<sup>18</sup> and Tiny Machine Learning (TinyML) to measure the complex impedance and the moisture of an ore sample. This approach aims to minimize the impact of air gaps on capacitive sensor readings. Specifically, RDFM is a validated and patented technique and technology<sup>18,19</sup>. It is used in soil moisture measurement systems to accurately determine relative permittivity and conductivity. Key process variables—conductivity, relative permittivity, density, torque, and chamber level—are used as inputs to the machine learning model to assess the ore's moisture content of a sample accurately.

In this work, the developed instrumentation system was evaluated with iron ore samples containing moisture levels in the range of 7–9.5%, which corresponds to the critical interval for safe handling and transport. For the ore samples used in this research, moisture levels above 8.5% pose a liquefaction risk during shipping operations, highlighting the importance of accurate measurement in this range, especially since the critical threshold depends on the mineral type.

There is a lack of measurement equipment available to meet the industry demand for rapid, on-site determination of ore moisture with precision and accuracy comparable to the standard method, yet with short response times. This gap affects multiple stages of processing and transport; however, in this study, we focus on the shipping stage, where pre-shipment testing is critical due to cargo-liquefaction risk. In this context, adherence to the Transportable Moisture Limit (TML) is mandatory, underscoring the need for timely and reliable measurement.

To this end, we conducted experiments using the developed method and benchtop instrument at the Physical Analysis Laboratory located at Vale SA, Ponta da Madeira, São Luís, Maranhão, Brazil. As one of the leading mining companies in Brazil, Vale SA operates the Ponta da Madeira Maritime Terminal (TMPM), a port primarily designed for exporting iron ore, manganese, and pellets from its mines. These materials are then loaded into ships for transportation to various international destinations.

The equipment was incorporated into the iron ore moisture measurement routine at TMPM's physical-chemical analysis laboratory, which plays a crucial role in ore production and commercialization, and is a mandatory step according to ABNT (Brazilian Technical Standards Association), as presented in<sup>20,21</sup> to ensure transportation safety. The laboratory's operations are directly integrated into the industrial process, ensuring the experiments reflect real-world applications. The proposed equipment, designed as a bench-scale system, streamlines moisture analysis of sampled ore, enhancing efficiency by accelerating this critical step. Additionally, it provides valuable insights into the overall moisture trend of the transported ore, contributing to more effective process control and decision-making.

For the experiment, we used samples of IOCJ (Iron Ore Carajás) in the final product stage. At this phase, the ore has well-defined properties that vary in a well-defined range. These are established through regular physical-chemical analyses to meet client requirements regarding granulometry, iron content, and contaminant levels. These strict quality controls help standardize the moisture measurement process, minimizing errors.

According to data from the Brazilian National Waterway Transportation Agency (ANTAQ)<sup>22</sup>, IOCJ was responsible for 21.04% of all bulk cargo transported in 2023, totaling approximately 166.334 million tons. Given this high production volume, studying this specific material is relevant and justified, as it plays a significant role in the Brazilian economy.

We conducted the *in loco* experiment in two scenarios: compressing and not compressing the samples. We collected data on both scenarios, which were used to create two different ML models. We seek to achieve the objectives of our study, which include a) developing a methodology for measuring iron ore moisture with a lightweight ML algorithm, such as tree-based models, and b) investigating if ore compression affects moisture

measurement. The ML model must be lightweight enough to run in an 8-bit microcontroller since we intend to embed it in the development of a commercial moisture analyzer.

The proposed approach, based on compressed ore samples and TinyML, achieved a mean absolute error of 0.12 percentage points (pp) and an expanded uncertainty of  $\pm 0.35$  pp using a Decision Tree model. These results are sufficient to discriminate between safe and unsafe moisture levels, thereby providing practical value for industrial applications.

Therefore, we formulated two questions to guide our research. As a result, our methodology sought to answer the following research questions:

- **RQ1:** Which machine learning models achieve the best trade-off between accuracy, latency, and memory when measuring iron ore moisture in TinyML applications?
- **RQ2:** What is the impact of ore compression on the accuracy and reliability of machine learning-based moisture measurement systems?

The structure of this paper is organized as follows: Related Works provide an overview of existing studies and approaches that evaluate bulk material moisture. Fundamental Concepts present the capacitive principle of the proposed sensor. Methodology describes the instrumentation system's design, functioning, and calibration. The results and analysis are discussed in Results, highlighting key findings of the experiment in the industrial environment. The research questions that guided this study are addressed in Discussion. Finally, Conclusion summarizes the contributions, discusses limitations, and proposes potential directions for future work.

## Related works

This section distinguishes off-line and on-line approaches to bulk-material moisture measurement. Off-line techniques include the traditional gravimetric method, loss-on-drying (LOD) moisture analyzers, and capacitive sensors. In contrast, online methods—such as infrared/hyperspectral systems and computer vision approaches installed on conveyor belts—provide continuous, in-process monitoring for control. Our contribution falls within the offline category, targeting rapid and repeatable measurements on compressed iron-ore samples for pre-shipment decisions.

Miljak et al.<sup>23</sup> introduced an online moisture meter for iron ore that utilizes microwaves and operates on a conveyor belt. Similarly, Viana<sup>12</sup> developed a system for measuring the moisture content of aluminum ore on a conveyor belt. Zhu et al.<sup>24</sup> also applied microwaves to determine the moisture content of concentrated iron ore. Additionally, Indra et al.<sup>25</sup> employed the microwave sweep method, combined with stratified random sampling, to evaluate the moisture content of nickel ore being transported on conveyor belts. These works address on-belt, online measurement under dynamic conditions (varying bed height and loading, particle-size segregation, temperature, and density). In contrast, our pre-shipping analyzer performs an offline measurement on compressed samples under controlled industrial conditions.

Li et al.<sup>26</sup> integrated microwave spectrum analysis with machine-learning regression, evaluating several models. They also performed on-site validation on 27 groups of coal, reporting encouraging predictive performance. The authors conclude that the results were promising; however, further testing in real-world applications is needed to further validate their suitability for industrial use. Additionally, microwave instrumentation entails practical considerations, such as equipment cost and periodic recalibration, that can affect scalability in production settings<sup>13</sup>.

There are applications using computer vision. Tang and Esmaeili<sup>27</sup> utilized an unmanned aerial vehicle (UAV) equipped with thermal imaging sensors to capture surface radiation data and then correlate it with surface moisture through an empirical linear regression model. However, the methodology is limited to outdoor, open areas and relies heavily on consistent solar radiation, requiring measurements to be taken in the afternoon for uniform surface temperatures. Weather conditions such as rain or cloudy days render the method unusable. An image-based approach was suggested by Zhou and Liu<sup>14</sup>. It evaluated the moisture of iron ore green pellets using image feature extraction and polynomial regression. However, it lacks industrial validation. Buchczik et al.<sup>28</sup> developed a method for measuring the moisture content of fine copper ore, integrating computer vision and thermography techniques with a focus on evaluating the influence of particle size on water content determination. Computer vision techniques were employed to determine particle size, while the IR camera assessed moisture over a broad range of measurements and particle sizes. Three models – linear, quadratic, and spline regression – were evaluated on both uniform particle size samples and mixed granulometries. Although the results were promising, the IR method measures moisture only on the sample's surface, and the experiment has not been validated in an industrial environment. Additionally, cameras face challenges in industrial settings due to dust, rain, and poor lighting conditions, which negatively impact image processing accuracy<sup>1,15</sup>.

The patent developed by Andrade et al.<sup>29</sup> presents a method and system to determine the moisture content of iron ore transported on a conveyor belt. It uses hyperspectral cameras paired with computer vision and machine learning algorithms to identify the water content in the ore. However, hyperspectral cameras have a high acquisition cost and have challenges validating their use in industrial applications.

As offline approaches, there are the LOD analyzers for moisture determination<sup>30</sup>. However, they require heating cycles and a turnaround time of minutes, which limits sampling frequency immediately before shipping<sup>31</sup>. We propose a compressed capacitive method with embedded TinyML that provides a rapid offline pre-shipment check in the industrial laboratory, calibrated against LOD. Moreover, LOD analyzers determine moisture in less representative small subsamples (typically a few grams; instrument capacities are on the order of tens to at most 100–200 g), while the proposed instrument is capable of measuring samples of up to 700 g.

Lage et al.<sup>16</sup> developed a bench system specifically for measuring the moisture of iron ore. The system operates on the principle of a current-to-voltage converter (I-V) with a trans-impedance amplifier to assess the

relative permittivity of the ore, which is then used to determine its moisture content. For future developments, they propose using a manual press to compact the ore within the instrument chamber, thereby minimizing the effects of air gaps on capacitive sensor readings. However, as noted by Rego<sup>18</sup>, the I-V method may not be the most suitable for industrial applications where the material under test (MUT) has high conductivity due to its lack of precision.

An improvement in evaluating the permittivity of high-conductivity materials, such as ore, is the real-dual-frequency method (RDFM). This method effectively reduces the impact of high conductivity on relative permittivity measurement by using affordable low-frequency hardware<sup>18</sup>. Pinto et al.<sup>17</sup> applied this approach with a capacitive sensor to measure moisture on a conveyor belt, while Santos et al.<sup>32</sup> used it similarly to assess ore moisture in a train wagon prototype. However, neither application compared machine learning models nor experimented with measuring the ore's moisture under compression.

Table 1 summarizes different works focused on ore moisture measurement. Our method uniquely utilizes a capacitive sensor combined with ore sample compression, leveraging a TinyML approach to estimate moisture content. Unlike previous studies, we conducted complex impedance measurements while compressing ore samples. Additionally, earlier research on capacitive sensors primarily focused on linear regression for moisture estimation. In contrast, we use multiple input features to evaluate lightweight tree-based machine-learning algorithms to enhance the instrument's accuracy and ensure the regressor is suitable for execution on a microcontroller.

Fundamental concepts

This Section presents fundamental concepts of the capacitive principle of the proposed sensor. In electromagnetism, the complex permittivity describes the ability of a material to store and dissipate energy in response to an applied electric field. For a lossy dielectric medium, this property can be represented by:

ε˙ = ε0εr - jσω, (1)

where ε0 is the permittivity of free space (=8.854 × 10<sup>-12</sup> F m<sup>-1</sup>), εr is the relative permittivity (dimensionless), σ is the conductivity (S m<sup>-1</sup>) and ω is the angular velocity (rad s<sup>-1</sup>).

According to electrical circuit theory, considering the coupling of the capacitive sensor to the medium under test, the complex permittivity becomes related to the complex impedance as<sup>17</sup>:

Z˙ = kg<sup>-1</sup>[σ<sup>-1</sup> - j(ωε0εr)<sup>-1</sup>], (2)

where kg is the geometric constant of the capacitor (m). Hence, the impedance can be rewritten as:

Z˙ = G<sup>-1</sup> - j(ωC)<sup>-1</sup>, (3)

where G is the conductance (S), which is the inverse of resistance (G = R<sup>-1</sup>) and C is the capacitance (F).

The complex impedance can be measured using the two plateaus' current-to-voltage circuit using a transimpedance amplifier, which excites the sensor with sinusoidal signals at two distinct frequencies and calculates the impedance by measuring each voltage gain<sup>17,33</sup>. The system's frequency response exhibits a characteristic curve with two well-defined plateaus and a 20 dB/dec interval between them. The first plateau, corresponding to low frequencies, allows the measurement of the material's conductance by suppressing the capacitive effect. The second plateau, at high frequencies, enables the determination of the capacitance while minimizing the influence of conductance.

In 3, as angular frequency ω increases, the influence of conductance on the total impedance diminishes. In this regime, the capacitive effect becomes dominant, making it more straightforward to measure. However, the

Work	Microwave	Computer vision	Capacitive sensor	Capacitive sensor + compression + TinyML
Miljak et al. <sup>23</sup>	●	○	○	○
Viana <sup>12</sup>	●	○	○	○
Zhu et al. <sup>24</sup>	●	○	○	○
Indra et al. <sup>25</sup>	●	○	○	○
Li et al. <sup>26</sup>	○	●	○	○
Tang and Esmaili <sup>27</sup>	○	●	○	○
Zhou and Liu <sup>14</sup>	○	●	○	○
Buchczik et al. <sup>28</sup>	○	●	○	○
Andrade et al. <sup>29</sup>	○	●	○	○
Lage et al. <sup>16</sup>	○	○	●	○
Pinto et al. <sup>17</sup>	○	○	●	○
Santos et al. <sup>32</sup>	○	○	●	○
Our work	○	○	○	●

Table 1. Related Works Focusing on ore moisture measurement.

plateau-based approach exhibits limitations when applied to lossy media, as the requisite frequency range expands proportionally with the material's conductivity. Consequently, the associated circuitry becomes increasingly costly and complex at higher frequencies. To overcome this challenge, the measurement methodology RDFM incorporates a mathematical model of the capacitive sensor's frequency response. This approach enables accurate permittivity measurements at comparatively lower frequencies than traditional methods, even for material under test with high conductivity<sup>18,19</sup>.

## Methodology

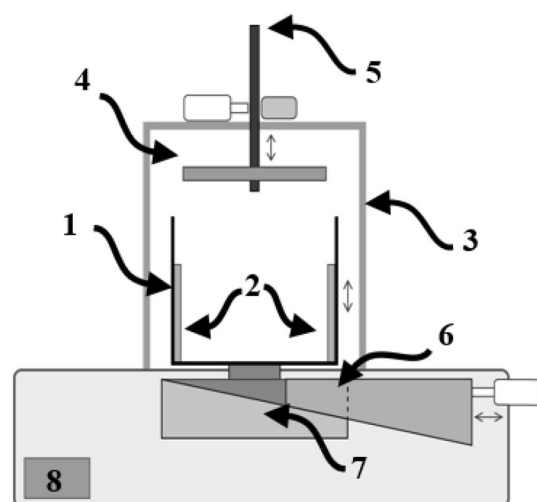
### Equipment operational principles

The measurement system incorporates a capacitive moisture sensor and a method for standardizing ore compression, consisting of these key components:

- **Measuring Chamber (1):** Made from acrylic, this box-shaped chamber with a hollow center and no top contains the stainless steel electrodes (2) and holds the ore sample for measurement. Additionally, a ruler is included for measuring the ore height, allowing volume determination.
- **Electrodes (2):** Two stainless steel plates are attached parallel to the interior sides, forming the capacitive sensor with floating terminals, meaning neither electrode is grounded. The electrodes are 48 mm apart and have a geometric constant of 0.083 m.
- **Metallic Support Structure (3):** Provides the framework for the compression system, ensuring stability during operation.
- **Press (4):** Serves as a surface for compressing the ore sample, aiding in consistent measurements.
- **Threaded Spindle and Digital Torque Wrench (5):** Enables vertical displacement of the compression system and includes a torque wrench to standardize the applied force.
- **Load Cell (6):** Measures the sample mass, allowing for bulk density determination.
- **Hood-Shaped Niche (7):** Designed to house a support block, this feature prevents overloading of the load cell during the ore compression process.
- **Measurement Circuit Box (8):** Houses the electronics and circuitry necessary for the measurement process. The complex impedance measurement employs an 8-bit microcontroller to implement the RDF method, as described by<sup>17,18</sup>.

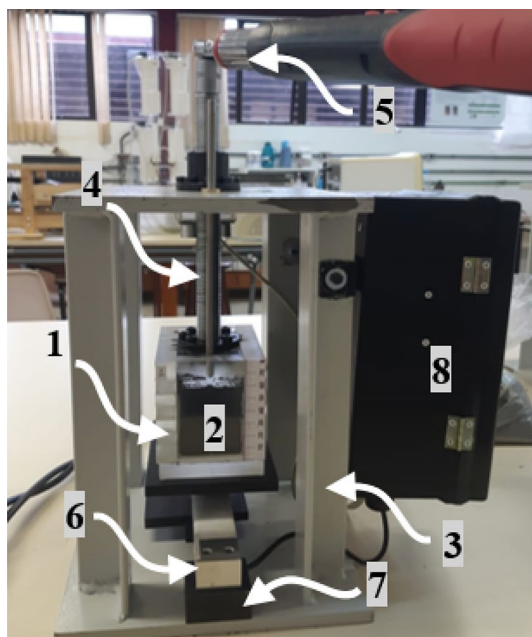
Figure 1 shows the schematic drawing of the setup, while Figure 2 presents the prototype used during the measurement campaign. The equipment operating methodology involves placing an approximately 700 g IOCJ sample in the measuring chamber, weighing it, and measuring its initial complex impedance before starting the compression process. After each compression, the complex impedance is re-measured in a continuous cycle until the operator reaches five compressions, and then the process can be restarted. The equipment's grounded metallic structure shields the sensor from external electromagnetic fields, while the climate-controlled environment minimizes concerns about temperature and humidity variations.

The accuracy of the sensor in measuring relative permittivity and conductivity was verified using aqueous substances as references (Table 2). Conductivity measurements at 20 °C were obtained using a Del Lab DL-150P conductivity meter. These substances, including mixtures of water and alcohol and varying concentrations of water and NaCl, were used to cover a range of relative permittivity and conductivity values. It is important to note that our measurement system achieved the highest accuracy for conductivity values  $> 90 \mu\text{S}/\text{cm}$ , which encompasses the range observed in iron ore samples. The higher relative errors observed for very low conductivity samples ( $< 10 \mu\text{S}/\text{cm}$ ) are outside the primary application range of this instrument. The complex



**Fig. 1.** Schematic drawing of the bench equipment for ore moisture measurement.





**Fig. 2.** Real prototype of the bench equipment for ore moisture measurement.

Samples	Reference $\varepsilon_r$	Measured $\varepsilon_r$	Reference $\sigma (\mu S/cm)$	Measured $\sigma (\mu S/cm)$
Air	1.0	1.8	0.07	7.63
Ethyl alcohol	25.2	22.3	4.33	5.40
25% water + 75% ethyl alcohol	39.3	37.0	4.28	1.61
50% water + 50% ethyl alcohol	54.0	54.8	4.03	3.19
75% water + 25% ethyl alcohol	69.6	69.6	4.83	8.55
Distilled water	80.2	77.6	3.34	10.76
Drinking water	80.2	78.2	26.11	13.45
Water + NaCl (1)	80.5	77.4	90.90	91.41
Water + NaCl (2)	80.3	82.5	266.50	260.70
Water + NaCl (3)	80.0	79.7	432.00	431.74
Water + NaCl (4)	80.0	80.1	851.30	853.15

**Table 2.** Reference Samples.

impedance measurement circuit used in the equipment is the same as presented by<sup>17,18</sup>, deploying the RDFM on an 8-bit microcontroller with 500 kHz and 8 MHz signals applied to the sensor.

### Iron ore moisture measurement

The second part of the sensor calibration involves collecting data from different but moisture-known iron ore samples. We carried out *in loco* tests for 6 days at the Ponta da Madeira port of Vale SA located in São Luiz, Maranhão. We used real ore samples called Iron Ore Carajás (IOCJ), which represent about 90% of all material loaded at the port. It was used 55 ore samples with different moisture percentages. All these samples were analyzed before ship transportation.

The partial IOCJ samples used in the experiment come from five different shipments. These samples exhibit a *Fe* variation between 64.87 and 65.89%, *SiO<sub>2</sub>* between 1.38 and 1.66%, *Al<sub>2</sub>O<sub>3</sub>* between 0.87 and 1.44%, *Mn* between 0.07 and 0.16%, *P* between 0.05 and 0.07%, *CaO* between 0.01 and 0.03%, *MgO* between 0.03 and 0.05%, *S* between 0.008 and 0.012%, *PPC* between 2.97 and 3.99%, *TiO<sub>2</sub>* between 0.078 and 0.114%.

It is important to note that while all samples used in this study are classified as the same product type (IOCJ sinter feed), the observed variations in measurement results cannot be primarily attributed to data noise. Our sensor validation with homogeneous substances demonstrates high measurement repeatability with minimal standard deviations<sup>17,18</sup>. Rather, the variability in our iron ore moisture measurements is derived from three main factors: (1) natural compositional differences between samples collected from different loading lines, with variations in iron content, silica, alumina, and other constituents, as detailed above; (2) the initial arrangement of particles within the measuring chamber before compression, which affects the initial complex impedance

readings; and (3) the variable effectiveness of compression in eliminating air gaps due to differences in particle size distribution.

Vale's physical analysis procedure includes the determination of particle size down to an aperture size of 150  $\mu\text{m}$ , with results showing that, on average, 77.65% of the particles are coarser than this size, while 22.35% pass through the 150  $\mu\text{m}$  sieve. The compression process, although standardized by torque measurements, interacts differently with samples exhibiting distinct granulometric profiles. This behavior supports our observation that sample compression enhances measurement accuracy by reducing, though not entirely eliminating, material-dependent variations. Figure 3 presents the average particle size distribution (PSD) curve of the samples, plotted as cumulative percent undersize versus particle size ( $\mu\text{m}$ ). The results indicate a progressive increase in cumulative undersize with decreasing particle size, showing that most of the material is concentrated in the finer fractions. This distribution is consistent with the expected granulometric behavior of iron ore, where a significant portion of particles is below 2 mm. Although the  $D_{50}$  values ranged from 0.5 mm to 1.0 mm, small pebbles up to 16 mm in diameter were occasionally observed, contributing to notable granulometric heterogeneity within individual samples.

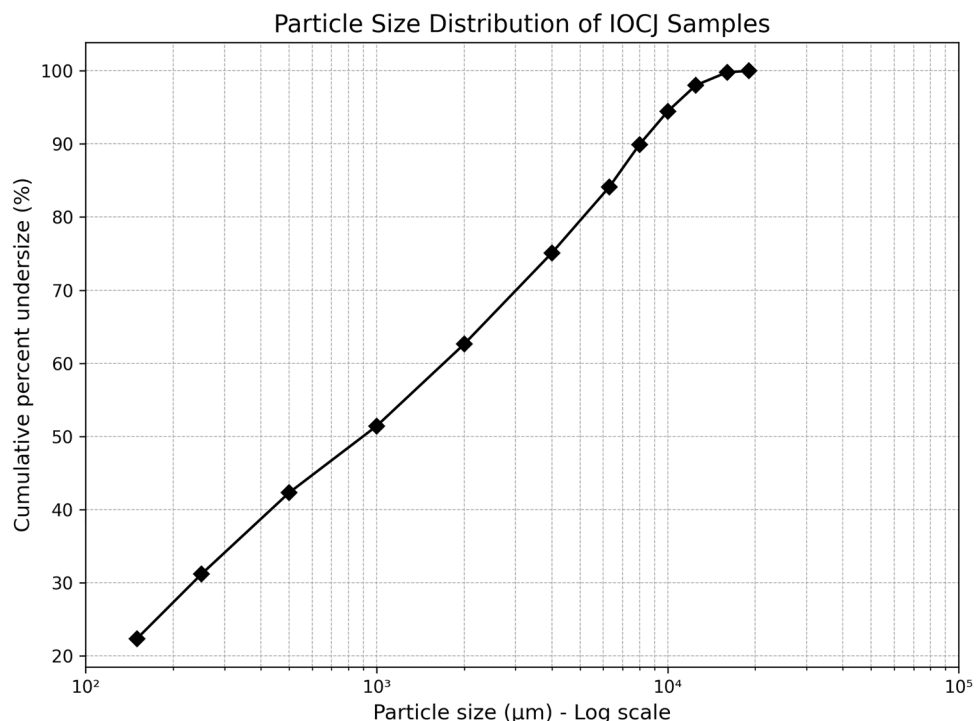
The Vale SA analysis laboratory uses gravimetric techniques<sup>8</sup> standardized to evaluate the ore moisture. We used these values as the ground truth of our models. Furthermore, we collected data without compressing the sample and then gradually compressing it. In this process, we measured variables such as electrical conductivity, relative permittivity, density, torque, and chamber level.

The measurement process with iron ore begins when we place the sample of the mineral in the measuring chamber, a load cell determines its mass, and we put the chock into the hood-shaped niche. Initially, we measure the variables without compression. Then, we apply torques of approximately 1.5 Nm, 1.7 Nm, 2 Nm, 5 Nm, and 10 Nm to the ore sample. The choice of torque was made empirically to ensure measurement accuracy while preserving the integrity of the acrylic chamber, which has a maximum limit of 15 Nm. Since greater compression reduces empty spaces in the sample, improving the precision of complex impedance measurement, we set a safe limit of 10 Nm to prevent structural damage and variations in electrode spacing. Figure 4 illustrates the variation of sensor capacitance with applied torque, highlighting how increasing compression influences the dielectric response.

The moisture associated with the evaluated ore samples was between 7.0% and 9.5%, considering the wet base. These values represent common moisture for the rainy season and represent a large part of the company's measurement of interest range. Table 3 shows the summary of the statistics of the dataset.

We developed two datasets—one with the measurements without compression and another with compression, consequently with variations on the applied torque. For model development, it is crucial to have a lightweight algorithm, as it must operate on microcontrollers, which characterizes a TinyML application.

Given the deployment target on microcontrollers (TinyML constraints), we prioritized methods with modest computational and memory footprints. We evaluated five supervised regressors:



**Fig. 3.** Particle size distribution of IOCJ samples showing the cumulative percent undersize as a function of particle size ( $\mu\text{m}$ ).

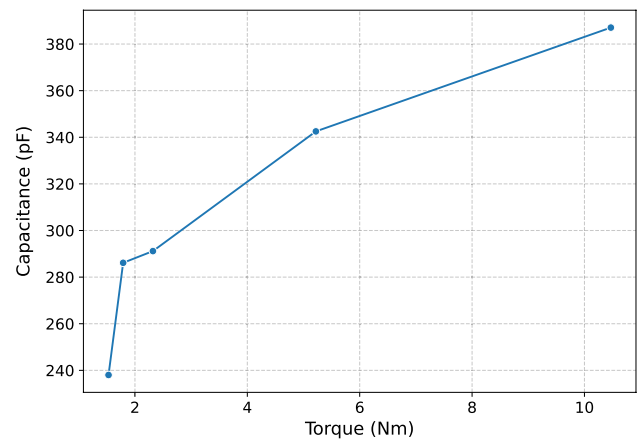


Fig. 4. Capacitance as a function of torque..

Variable	Unit	Mean	Std. Dev	Min	Max
Torque	Nm	3.99	3.20	1.52	10.80
Level	mm	64.83	3.92	55.40	76.10
Density	g/cm <sup>3</sup>	2715.17	138.52	2129.95	3077.37
Conductivity	μS/cm	55.70	15.83	25.87	109.66
Relative Permittivity	-	49.94	16.32	19.34	93.97
Moisture	%	8.36	0.46	7.03	9.49

Table 3. Descriptive statistics of the dataset.

- (i) Multiple Linear Regression (MLR) <sup>34</sup>,
- (ii) a Decision Tree (DT) <sup>35</sup>, and
- (iii) Random Forests (RF) <sup>36</sup> with 5, 10, and 15 trees, denoted RF5, RF10, and RF15, respectively.

These models differ in how they use variables: MLR assumes linear additivity of predictors, whereas DT/RF capture interactions and nonlinearities through hierarchical splits and ensembling, respectively.

To assess generalization and guard against overfitting, we employed Leave-One-Out Cross-Validation (LOOCV). In LOOCV, each sample serves once as the test case while the remaining  $n - 1$  samples are used for training; the held-out prediction  $\hat{y}_i^{(-i)}$  is compared with the reference  $y_i$ . This protocol provides an essentially unbiased, sample-level estimate of out-of-sample performance; overfitted models would exhibit low training error but elevated LOOCV error. Although computationally intensive, LOOCV is advantageous for small datasets because it maximizes the training data per split<sup>37,38</sup>. All reported metrics are computed only from the LOOCV predictions: Mean Absolute Error (MAE), Mean Squared Error (MSE), coefficient of determination  $R^2$ , and an uncertainty estimated as twice the standard deviation of the residuals.

We consider the input variables  $X = \{\text{Torque (Nm)}, \text{Level (mm)}, \text{Density (g cm}^{-3}\text{)}, \text{Conductivity (}\mu\text{S/cm)}, \text{RelativePermittivity}\}$  and the output variable  $y = \text{Moisture (\%)}$ . To ensure fairness and interpretability, all models were trained on the same input set  $X$  and evaluated under the identical LOOCV protocol. This design avoids confounding due to model-specific pre-selection of variables and provides a transparent comparison of predictive performance. The code and data are available in the GitHub repository<sup>39</sup>.

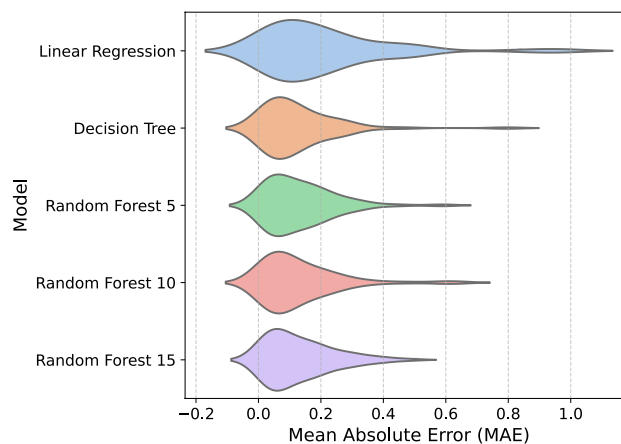
We selected the algorithm with lower error and best performance on a microcontroller, addressing the **RQ1**. We used the 8-bit microcontroller ATmega2560, running at 16 MIPS, to embed the models. This choice was made due to its common availability, cost-effectiveness, and suitability for transforming models into C code for microcontroller deployment. This was done by means of Python *micromlgen* library, which converts machine-learning models from Python into C implementations. Each model was deployed to the microcontroller, and key metrics such as processing time, flash memory usage, and RAM consumption were recorded. We chose the algorithm that best performs on the prediction.

To address **RQ2**, we trained two models: one with sample compression and the other without. We compared their performances to investigate the impact of ore compression on moisture measurement. Performance metrics were calculated for each model, and statistical tests were performed for comparison.



Model	MSE	MAE	R <sup>2</sup>	Uncertainty
MLR	0.0826	0.2058	0.5568	±0.58
DT	0.0302	0.1241	0.8211	±0.35
RF5	0.0288	0.1308	0.8266	±0.34
RF10	0.0318	0.1283	0.8221	±0.36
RF15	0.0283	0.1301	0.8464	±0.34

**Table 4.** Performance metrics of machine learning models. MAE and uncertainty are reported on the original moisture scale (percentage points) using LOOCV predictions..



**Fig. 5.** Violin plot of the Mean Absolute Error (MAE) for each model.

Comparison	Conclusion
RF15 vs MLR	Reject null hypothesis
RF15 vs DT	Fail to reject null hypothesis
RF15 vs RF 5	Fail to reject null hypothesis
RF15 vs RF 10	Fail to reject null hypothesis

**Table 5.** Paired t-test conclusions for model comparisons.

## Results

### Model development

We employed leave-one-out cross-validation for training and testing on the dataset with ore's compression and calculated metrics such as MSE, MAE, R<sup>2</sup>, and uncertainty. Table 4 shows the obtained metrics for each model.

The violin plot of the Mean Absolute Error (MAE) for each model is presented in Fig. 5. Except for multiple linear regression, the models obtained very similar behavior. We also observed that the RF15 exhibited less variability compared to the others. Hence, we performed a Welch's t-test to determine whether the RF15 is statistically different from the others. Table 5 shows that Random Forest 15 significantly outperforms Linear Regression; however, according to the test, there are non-significance differences between other algorithms. Thus, the algorithms with non-significance differences were evaluated by their computational costs.

### Evaluation of microcontroller performance

We have collected metrics related to the microcontroller process during its functioning. The processing time refers to the duration taken by the microcontroller to execute a single prediction. Flash memory usage indicates the amount of non-volatile storage required to house the model code and associated data. RAM consumption measures the dynamic memory used during the model's execution.

The decision tree algorithm exhibits a slight increase in flash and RAM usage when compared to the random forest algorithms. However, it demonstrates a significantly lower execution time. Fig. 6 presents the execution time comparison for each model evaluated. Since the t-test revealed no significant differences in error rates between the decision tree and the other algorithms, this suggests that the decision tree is more efficient for this task.

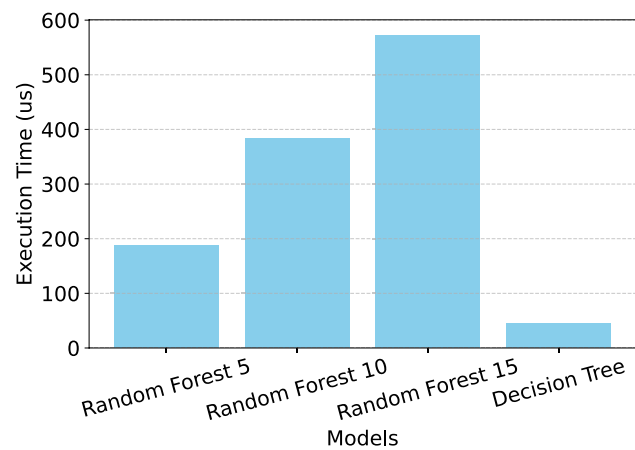


Fig. 6. Execution time comparison.

Metric	Compressed	Non-compressed
MAE (pp)	0.12	0.47
Uncertainty (pp)	±0.35	±1.3
R <sup>2</sup>	0.82	0.54

Table 6. Comparison of models based on compressed versus non-compressed ore samples. MAE and uncertainty are reported on the original moisture scale (percentage points) using LOOCV predictions..

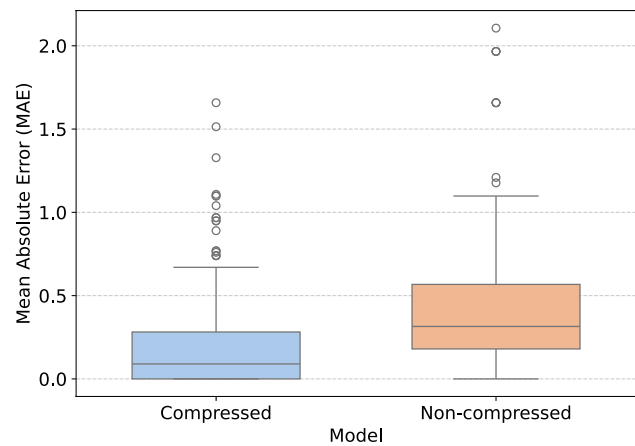


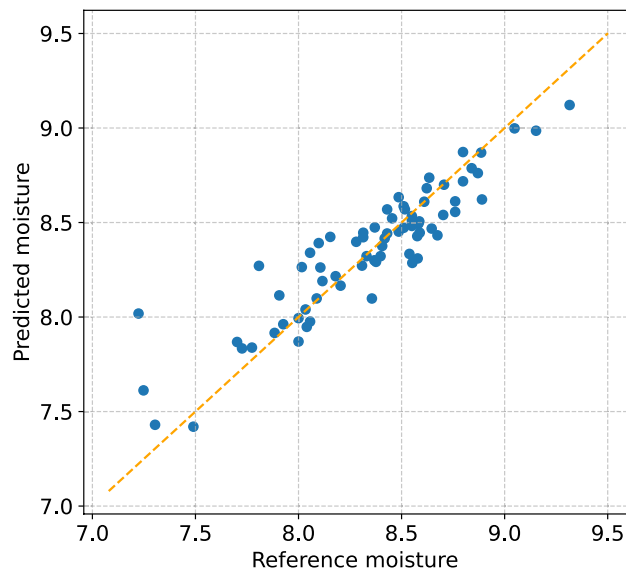
Fig. 7. Box plots representing the evaluation results of the compressed and non-compressed models, showing the distribution of prediction errors (MAE) for each case.

Impact evaluation of ore compression

We selected the decision tree algorithm to examine the effects of sample compression on moisture measurement. Using data gathered from non-compressed ore samples, we developed another decision tree model. To evaluate the non-compressed model, we also used LOOCV. During each evaluation phase, we calculated metrics to assess the performance of the processes. The outcomes of these evaluations are detailed in Table 6.

The compressed model demonstrates superior performance, achieving a substantially lower MAE of 0.12 pp compared to 0.47 pp for the non-compressed model. This indicates that compressing ore samples leads to more accurate predictions on average. Moreover, the uncertainty of the compressed model is markedly lower, suggesting that its predictions are more consistent and reliable. The higher  $R^2$  value further supports its effectiveness, showing that the compressed model explains a larger proportion of the variance in the data. Overall, these metrics highlight the advantage of applying compression to ore samples. Figure 7 presents the box plot of MAE for the two models, emphasizing the performance differences between them.

The Mann-Whitney U Test was conducted to compare these groups and assess the significance of their differences. The results indicate a statistically significant difference between the two groups, as evidenced by a



**Fig. 8.** Plot of the model for compressed samples, showing predicted and reference moisture values across all data points..

U-statistic of 1137.5 and a p-value of  $6.061 \times 10^{-13}$ . This suggests a meaningful difference in the median values of the two groups under study. The result allows us to reject the null hypothesis that there is no difference between the groups, supporting the alternative hypothesis of a significant difference in their central tendencies. These highlight the significant improvement in the measurement of compressed ore samples.

Figure 8 presents the model for compressed samples: leave-one-out predictions versus reference moisture values. Each point corresponds to a sample, plotted with the reference moisture content on the x-axis and the model prediction on the y-axis. The orange dashed line represents the line of equality  $y = x$ , where perfect agreement would occur. The majority of points are distributed closely around the equality line, indicating that the model predictions are largely unbiased across the studied range. This confirms that compression improves model stability and that the predictive performance is sufficient to discriminate safe versus critical moisture levels.

## Discussion

A response can be given to the research questions based on the results obtained.

**RQ1:** Which machine learning models achieve the best trade-off between accuracy and latency when measuring iron ore moisture in TinyML applications?

Previous work has explored the use of multiple linear regression; however, it did not yield satisfactory results in our experiments. As a result, we trained and validated other machine-learning algorithms. We tested decision trees and random forests using configurations of 5, 10, and 15 trees. These tree-based algorithms demonstrated significant improvement compared to multiple linear regression. When comparing the performances of the different algorithms, we found that their results were quite similar. However, the random forests with 15 trees achieved slightly better outcomes. To verify if there are statistical differences between the ML error populations, we performed a t-test, which failed to reject the null hypothesis. This indicates that, based on prediction errors, the algorithms are statistically similar. Hence, we embedded each model on the ATmega2560 and collected the microcontroller's performance metrics, such as execution time, free RAM, and flash memory consumption. RAM and flash showed no significant changes, but the execution time of the decision tree was much lower than the others. Consequently, we chose the decision tree algorithm due to its efficiency based on the prediction's errors and response latency.

**RQ2:** What is the impact of ore compression on the accuracy and reliability of machine learning-based moisture measurement systems?

We compare two decision tree models: one trained on compressed ore samples and another on non-compressed samples. The compressed dataset demonstrates significant improvements over the non-compressed dataset across all evaluated metrics. Using compressed samples leads to a 73.76% reduction in MAE, indicating substantially higher accuracy. The uncertainty is reduced by 73.33%, reflecting enhanced reliability in the predictions. Furthermore, the  $R^2$  shows a 52.00% improvement, underscoring the superior ability of the compressed dataset to explain variance in the data. These results collectively highlight the advancements in precision, consistency, and predictive capability of compressing samples when measuring complex impedance for moisture prediction.

## Conclusion

The work presents a bench system for measuring iron ore moisture using complex impedance with standardized compressions. It aims to improve decision-making in the iron ore production chain, particularly during ship loading, by providing a real-time, high-accuracy measurement method. Currently, the mineral industry lacks such a method, which affects decision-making. The study proposes a solution by compressing ore samples and utilizing the RDFM method along with a TinyML model to assess moisture content.

The proposed system, tested with IOCJ samples at Vale SA's Ponta da Madeira port, demonstrated that compression significantly improves prediction accuracy, achieving an average MAE of 0.12 pp and reducing error by 73.76% compared to non-compressed samples. In addition, decision-tree models enabled near-instant predictions (44  $\mu$ s), cutting the turn-around time from hours, as in conventional methods, to microseconds.

These results indicate that the method not only ensures high precision but also allows real-time decision-making during ship loading, reducing delays and associated costs. Faster and more accurate measurements can increase loading capacity, improve scheduling, and enhance global competitiveness at high-volume ports.

This work marks significant progress in transforming the technology into industrial equipment. Given TPM's high cargo volume, faster measurements can enhance loading capacity, reduce costs, and improve global competitiveness. Although the equipment is calibrated for a single ore product, establishing reliable technology at one of Latin America's largest ports presents a key opportunity to maximize efficiency. As moisture measurement speed and accuracy are crucial in the loading process, an improved system can greatly accelerate this step.

For future work, we aim to further enhance the hardware of the bench system to make it fully automated, improving efficiency and reducing the need for manual intervention. Concerning the compression process, we intend to use industrial servomotors designed for precise control of torque, speed, and position, ensuring a robust and reliable solution to maintain process consistency and enhance the equipment's trustworthiness.

We also plan to assess model update techniques, as different types of ore necessitate specific model calibrations, and to address model degradation. We will evaluate incremental learning methods, such as just-in-time learning, since traditional moisture measurements provide ground truth data, even though they respond slowly.

## Data availability

All data generated or analyzed during this study are included in this published article.

Received: 3 June 2025; Accepted: 30 October 2025

Published online: 28 November 2025

## References

- Matos, S. N. et al. An evaluation of iron ore characteristics through machine learning and 2-d lidar technology. *IEEE Trans. Instrum. Meas.* **73**, 1–11. <https://doi.org/10.1109/TIM.2023.3342220> (2024).
- Munro, M. C. & Mohajerani, A. Moisture content limits of iron ore fines to prevent liquefaction during transport: Review and experimental study. *Int. J. Miner. Process.* **148**, 137–146 (2016).
- Klippel, E. et al. Deep learning approach at the edge to detect iron ore type. *Sensors* **22**, 169 (2021).
- Abazarpour, A., Halali, M., Hejazi, R. & Saghaeian, M. Hprg effect on the particle size and shape of iron ore pellet feed using response surface methodology. *Mineral Processing and Extractive Metallurgy* **127**, 40–48 (2018).
- Ferreira, R. F., Policarpo, D. L. V., Padula, V. P. & Ferreira, M. T. S. Limite de umidade transportável de minérios de ferro: Aspectos regulatórios e técnicos. *Tecnologia em metalurgia, materiais e mineração* **14**, 16–23 (2017).
- Munro, M. C. & Mohajerani, A. Determination of the transportable moisture limit of iron ore fines for the prevention of liquefaction in bulk carriers. *Mar. Struct.* **40**, 193–224 (2015).
- Clout, J. & Manuel, J. Mineralogical, chemical, and physical characteristics of iron ore. In *Iron Ore* 45–84 (Elsevier, 2015).
- Associação Brasileira de Normas Técnicas (ABNT). *Norma ABNT NBR ISO 3082 - Minério de ferro - Processo de amostragem e preparação de amostras* (ABNT, Brasil, 2019).
- Associação Brasileira de Normas Técnicas (ABNT). *Norma ABNT NBR ISO 3087 de 08/2021 - Minérios de ferro - Determinação do teor de umidade de um lote* (ABNT, Brasil, 2021).
- Cancilla, P., Barrette, P. & Rosenblum, F. On-line moisture determination of ore concentrates a review of traditional methods and introduction of a novel solution. *Miner. Eng.* **16**, 151–163 (2003).
- Miljak, D. G., Bennett, D., Kazzaz, T. & Cutmore, N. G. On-line microwave moisture measurement of iron ore and mineral concentrates in conveyor applications. In *2006 IEEE Instrumentation and Measurement Technology Conference Proceedings*, 183–186, (2006). <https://doi.org/10.1109/IMTC.2006.328366>
- Viana, S. A. Microwave-based moisture measurement of bauxite ore on conveyor belts. *Eng. Min. J.* **214**, 44 (2013).
- Segundo, A. K. R. et al. Sistema e método para medição de umidade de uma amostra de solo ou minério (2019). Patent submission number: BR 10 2019 016715 7.
- Zhou, S.-y. & Liu, X.-y. A simple image-based method for online moisture content estimation of iron ore green pellets. *Journal of Iron and Steel Research International* 1–13 (2023).
- Nienhaus, K., Warcholik, M., Büschgens, C. & Müller, D. Belt positioning and skewing prevention in lignite mining using long-wavelength infrared cameras. In *Proceedings of the 12th International Symposium Continuous Surface Mining-Aachen 2014* 11–20 (Springer, 2015).
- Lage, V. N. et al. Bench system for iron ore moisture measurement. In *2018 12th International Conference on Sensing Technology (ICST)* 139–142 (IEEE, 2018).
- Pinto, É. S., Santos, G. A., Magalhães, P. H. V. & Segundo, A. K. R. A novel two degrees of freedom system for measuring iron ore water content on a conveyor belt. *IEEE Sens. J.* **21**, 16291–16300 (2021).
- Rêgo Segundo, A. K., Silva Pinto, É., Almeida Santos, G. & de Barros Monteiro, P. M. Capacitive impedance measurement: Dual-frequency approach. *Sensors* **19**, 2539 (2019).
- Segundo, A. K. R., Santos, G. A., Érica Silva Pinto & Monteiro, P. M. D. B. Método real-dual-frequency para medição de impedância capacitiva utilizando circuito conversor tensão-corrente (2021). Assignee: Universidade Federal De Ouro Preto - Ufop and Vale S.A.
- ABNTa. Abnt nbr iso 3082 - minério de ferro - procedimentos de amostragem e preparação de amostras. Tech. Rep., Associação Brasileira de Normas Técnicas (2019).

21. ABNTb. Norma brasileira abnt nbr iso 3087 - minérios de ferro - determinação do teor de umidade de um lote. Tech. Rep., Associação Brasileira de Normas Técnicas (2021).
22. Agência Nacional de Transportes Aquaviários (ANTAQ). Estatístico aquaviário - movimentação portuária de cargas a granel em 2023 (2025). Access: 03/19/2025.
23. Miljak, D. G., Bennett, D., Kazzaz, T. & Cutmore, N. G. On-line microwave moisture measurement of iron ore and mineral concentrates in conveyor applications. In *2006 IEEE Instrumentation and Measurement Technology Conference Proceedings*, 183–186, <https://doi.org/10.1109/IMTC.2006.328366> (2006).
24. Zhu, J. D., Chen, Z. P., He, C. J. & Xu, E. D. A moisture detection system of iron ore concrete. *IET Conference Publications* **2015**, <https://doi.org/10.1049/CP.2015.1563> (2015).
25. Indra, W. A., Hassim, N. B. & Aziz, S. A. C. Performance of on-line microwave moisture measurement for nickel ore. *ARPN Journal of Engineering and Applied Sciences* **13**, 2335–2339 (2018).
26. Li, M. *et al.* Intelligent measurement of coal moisture based on microwave spectrum via distance-weighted knn. *Applied Sciences* **2022**, Vol. 12, Page 6199 **12**, 6199, <https://doi.org/10.3390/APP12126199> (2022).
27. Tang, M. & Esmaili, K. Mapping surface moisture of a gold heap leach pad at the el gallo mine using a uav and thermal imaging. *Mining, Metallurgy and Exploration* **38**, 299–313. <https://doi.org/10.1007/S42461-020-00332-4/METRICS> (2021).
28. Buchczik, D., Budzan, S., Krauze, O. & Wyzgolik, R. Moisture determination for fine-sized copper ore by computer vision and thermovision methods. *Sensors* **23**, 1220. <https://doi.org/10.3390/S23031220/S1> (2023).
29. de Andrade, I. A. P. F., Canaverde, H. S., Machado, L. H. & Oliveira, V. Método, sistema e programa de computador para identificação do teor de umidade online presente em minério de ferro passante sobre uma correia transportadora. Patent WO2022147602A1. BR Patent. Número Internacional: PCT/BR2021/050142 (2021).
30. METTLER TOLEDO. Moisture analysis solutions (2025). Accessed: 20 Sep 2025.
31. METTLER TOLEDO. Moisture content determination (2025). Accessed 20 Sep 2025.
32. Santos, G. *et al.* Real-time capacitive sensor applied to a train wagon prototype for measuring iron ore moisture. *IEEE Lat. Am. Trans.* **21**, 595–601 (2023).
33. dos Santos, E. N. *et al.* Dual-modality wire-mesh sensor for the visualization of three-phase flows. *Meas. Sci. Technol.* **26**, 105302 (2015).
34. Montgomery, D. C., Peck, E. A. & Vining, G. G. *Introduction to linear regression analysis* (John Wiley & Sons, 2021).
35. Moore, D. H. Classification and regression trees, by leo breiman, jerome h. friedman, richard a. olshen, and charles j. stone. brooks/cole publishing, monterey, 1984, 358 pages, 27.95 (1987).
36. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
37. Matos, S. N. *et al.* Machine learning techniques for improving multiclass anomaly detection on conveyor belts. In *2024 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)* 1–6 (IEEE, 2024).
38. Souza, F. A., Araújo, R. & Mendes, J. Review of soft sensor methods for regression applications. *Chemom. Intell. Lab. Syst.* **152**, 69–79 (2016).
39. Matos, S. N. Moisture sensor data and analysis scripts. <https://github.com/sauloneves/moisture-sensor> (2025). Accessed: 2025-09-24.

## Acknowledgements

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001; Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq); Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG); Vale S.A.; Instituto Tecnológico Vale ITV; Universidade Federal de Ouro Preto. The publication costs of this research were supported by Helmholtz-Zentrum Dresden-Rossendorf. For open access purposes, the authors have assigned the Creative Commons CC BY license to any accepted version of the article.

## Author contributions

É.S.P. and S.N.M. contributed equally to the design and conceptualization of the study. É.S.P. conducted the main experiments, and S.N.M. performed the data analysis. G.A.S., and M.N. supported the data collection and assisted with preliminary analysis. L.S.M., J.U., and P.V.P. contributed to theoretical discussions and methodology refinement. G.P. and A.K.R.S. supervised the project, provided guidance on experimental design, and contributed to manuscript writing. T.A.M.E. helped with the interpretation of results and contributed to manuscript editing. All authors reviewed and approved the final version of the manuscript.

## Funding

Open Access funding enabled and organized by Projekt DEAL.

## Declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to S.N.M. or T.A.M.E.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025