

DIAGNOSIS EVALUATION OF THE COFFEE LEAF RUST DEVELOPMENT
STAGE IN THE COLOMBIAN CATURRA VARIETY INTEGRATING REMOTE
SENSING, WIRELESS SENSOR NETWORKS AND DEEP LEARNING

A Thesis

Submitted to the Coordination of the

Master in Engineering

Universidad EAFIT

by

Alejandro Sánchez Aristizábal

Sebastián Sarmiento Garavito

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

April 2019

Universidad EAFIT

Medellín, Antioquia

STATEMENT OF THESIS APPROVAL

Approved by:

M.Sc. David Velásquez Rendón

Tutor of the project, professor of the Computer Science Department at *Universidad EAFIT* and member of the GIDITIC research group.

This research project is dedicated to all Colombian coffee growers, who trust that the technological innovation is a key factor for this country to produce the best coffee worldwide again.

ACKNOWLEDGMENTS

We would first like to thank our thesis advisor M.Sc. David Velásquez Rendón of the Faculty of Systems Engineering at *Universidad EAFIT*. The door to his office was always open whenever we had questions about our research or writing. He consistently allowed this thesis to be our own work, but steered us in the right direction whenever he thought we needed it.

We would also like to thank the experts who were involved in the interviews for this research project: M.Sc. Leonardo Henao, Administrator Juan Carlos Guerra, M.Sc. Jhon Wilson Mejía, Administrator and Zootechnologist Juan Carlos Posada, Engineer Juan Carlos Álvarez, Engineer Mario Duque, Engineer Julio César Robledo, Head of Plant Pathology at *CENICAFÉ* Carlos Ariel Ángel, Instructor Andrés Ruiz and M.Sc. Juliana Maya. Their participation, input and recommendations were a fundamental part for the successful orientation and development of this research.

Furthermore, Engineer Carlos Mario Ospina is worthy of special thanks for its accompaniment during the visit to *CENICAFÉ*'s Experimental Station "El Rosario" as well as the coffee grower Nabor Giraldo, M.Sc. Diego Miguel Sierra and Dentist Samuel Roldán for helping by the procurement of the coffee plants and the biological matter containing Coffee Leaf Rust. Additionally, we would like to thank Ph.D. Alejandro Marulanda, Ph.D. Edwin Nelson Montoya, Engineer Hugo Murillo, *Universidad EAFIT* and *Colciencias* for the constant academic, infrastructural and financial support.

It is also important to highlight the enormous work and dedication of the project's Monitors Engineer Camilo Velásquez, Engineer Felipe Gutiérrez, Engineer Sebastián

Osorio, Engineer Juan Diego Zuluaga and Physics Engineering undergraduate student Sergio Jurado. In addition, we want to acknowledge the Biology team, conformed by the undergraduate students Alisson Martínez and Laura Cristina Moreno and led by Ph.D. Luisa Fernanda Posada, mention the Systems Engineering undergraduate students Alejandro Cano, Luis Javier Palacio and Sebastián Giraldo and thank the Engineers Ricardo Urrego and Luis Felipe Machado as well as the M.Sc. Sebastián Rodríguez. We are gratefully indebted to each of those persons for their patience, very valuable ideas, comments and passionate effort on the realization of this thesis.

Finally, we must express our very profound gratitude to our families, our girlfriends and friends for providing us with unfailing support and continuous encouragement throughout our years of study. This accomplishment would not have been possible without them. Thank you.

TABLE OF CONTENTS

	Page
LIST OF TABLES	xi
LIST OF FIGURES	xii
ABBREVIATIONS	xvii
ABSTRACT	xix
1 Introduction	1
1.1 Problem description	1
1.2 Objectives	4
1.2.1 General objective	5
1.2.2 Specific objectives	5
1.3 Contribution	6
1.4 Thesis structure	7
2 Methodology	9
3 Observation	12
3.1 Leonardo Henao	12
3.2 Juan Carlos Guerra	14
3.3 Jhon Wilson Mejía	17
3.4 Juan Carlos Posada	19
3.5 Juan Carlos Álvarez	20
3.6 Mario Duque	23
3.7 Julio César Robledo	24
3.8 Carlos Ariel Ángel	27
3.9 Andrés Ruiz	28
3.10 Juliana Maya	30
3.11 Summary	31

	Page
4 Research question	33
4.1 Research hypothesis	33
5 Theoretical and conceptual framework	34
5.1 Agro	34
5.1.1 Coffee	34
5.1.2 Life cycle and phenological phases of coffee	35
5.1.3 Coffee varieties	37
5.1.4 Caturra variety	38
5.1.5 Specialty Coffee	39
5.1.6 Coffee diseases	40
5.1.7 Coffee Leaf Rust	43
5.2 Technology	45
5.2.1 Internet of Things	45
5.2.2 Precision Agriculture	46
5.2.3 Serological analysis	47
5.2.4 Molecular analysis	48
5.2.5 Detection of visual symptoms	49
5.2.6 Volatile Organic Compounds	51
5.2.7 Remote Sensing	52
5.2.8 Spectroscopy	53
5.2.9 Wireless Sensor Networks	55
5.2.10 Biosensors	56
5.2.11 Computer Vision	57
5.2.12 Deep Learning	58
5.2.13 Web programming framework	65
6 State of the art	67
6.1 Remote Sensing	67
6.2 Visual Detection	70

	Page
6.3 Biological Intervention	71
6.4 Wireless Sensor Networks	73
6.5 Machine Learning	74
6.6 Commercial technologies	76
7 Design of experiment	79
7.1 Completely Randomized Design	79
7.2 Data collection	80
8 Conceptual approach	82
8.1 Product Design Specification	82
8.2 Black box	83
8.3 Functional structure	84
8.4 Morphological matrix	86
8.5 Evaluation of concepts	86
9 Detailed design approach	91
9.1 Mechanical Design	91
9.1.1 Coffee crop lots	91
9.1.2 pH measurement system	92
9.1.3 Rotary platform	94
9.1.4 Rotary arm	96
9.1.5 Rain system	98
9.2 Electronic Design	99
9.2.1 Arduino 1: Lot data collection	100
9.2.2 Arduino 2: General data collection	106
9.2.3 Raspberry Pi®	110
9.2.4 Electrical cabinet	112
9.3 Informatic Design	114
9.3.1 Arduinos	114
9.3.2 Communication commands Raspberry Pi®-Arduinos	117

	Page
9.3.3 Raspberry Pi [®]	119
9.3.4 Thingworx [®]	124
9.3.5 Academic Data Center	125
10 Prototyping and results	130
10.1 Mechanical Prototyping	130
10.1.1 Rotary arm	130
10.1.2 Coffee crop lots	137
10.1.3 Rain system	139
10.1.4 pH measurement system	141
10.1.5 Rotary platform	143
10.2 Electronic Prototyping	147
10.2.1 Electrical cabinet	148
10.2.2 Arduino 1	151
10.2.3 Arduino 2	155
10.2.4 Raspberry Pi [®]	158
10.3 Informatic Prototyping	161
10.3.1 Communication commands Raspberry Pi [®] - Arduinos	161
10.3.2 Raspberry Pi [®]	162
10.3.3 Thingworx [®]	169
10.3.4 ADC	171
10.4 Integration	178
10.5 Results	179
11 Analysis of results	181
12 Conclusions	184
13 Future work	186
REFERENCES	187
A PDS	197
B Functional structure	201

	Page
C Morphological matrix	203
D Advantages, disadvantages and solutions for both concepts	214
E Coffee crop lots	216
F pH measurement system	219
G Rotary platform	226
H Rotary arm	247
I Arm support structure	268
J Rain system	275
K Arduino 1: Lot data collection	285
L Arduino 2: General data collection	287
M Raspberry Pi®	289
N Electrical cabinet	291
O Collected data management, diagnostic model generation and performance evaluation	293
P Arduinos checklist template for laboratory	295
Q Arduinos checklist for laboratory	297
R Arduinos checklist template for greenhouse	300
S Arduino 2: General data collection - final version	302
T Raspberry Pi® checklist template for greenhouse	304
U Raspberry Pi® - final version	306
V Sub-models' architectures	308
W Comparative table for the statistical analysis	312

LIST OF TABLES

Table	Page
5.1 Main characteristics of the coffee varieties cultivated in Colombia [26]. . . .	37
5.2 Ranking for Specialty coffee [2].	40
6.1 Kinds of variables that describe the importance of Coffee plots in the presence of Coffee Leaf Rust [78]	71
8.1 Concepts scoring.	87
9.1 Functions of Arduino 1.	116
9.2 Functions of Arduino 2.	117
9.3 Communication commands Raspberry Pi®-Arduino 1.	118
9.4 Communication commands Raspberry Pi®-Arduino 2.	119
10.1 Checklist to test Raspberry Pi's components.	148
10.2 Functions of the <code>data_collection.py</code> script.	169
10.3 End-points for updating the data in the visualization dashboard.	170
10.4 Hyperparameters and F_1 -score for each generated sub-model.	175
10.5 Weights for the predictions of each sub-model.	176
10.6 F_1 -score reached by the individual sub-models and the composite model.	180
11.1 ANOVA table of the statistical analysis.	182

LIST OF FIGURES

Figure	Page
2.1 Major phases of DIR [14].	10
2.2 Diagram representation of the <i>Pahl & Beitz</i> methodology [15].	11
5.1 Map of Colombian coffee regions [24].	35
5.2 Phenological phases of coffee [25].	36
5.3 <i>Caturra</i> coffee tree.	39
5.4 Coffee Leaf Rust effect on the <i>Caturra</i> variety [7].	43
5.5 Coffee Leaf Rust on a leaf in its highest development stage [7].	44
5.6 Smart healthcare concept [41].	46
5.7 Evaluation of the opportunities of regulated deficit irrigation for saving water in southern Spain [42].	47
5.8 Simple ELISA kit developed at the Plant Virus Laboratory from Federal University of Ceará, Brazil [44].	48
5.9 Detection of Nanoviruses by PCR [46].	49
5.10 Coffee Leaf Rust detection through visual inspection [48].	50
5.11 Field sampling cotton bolls with a pest and damage detector developed to draw VOCs [50].	52
5.12 UAV equipped with RS devices for evaluation of weed infestation [45].	53
5.13 The three acquisition modes in spectroscopy: (a) Reflectance, (b) Trans- mittance and (c) Absorption [54].	54
5.14 WSN in agriculture for water management [57].	56
5.15 Analysis of carbon nanotube biosensors embedded in an <i>Arabidopsis thaliana</i> plant [61].	57
5.16 <i>Nvidia</i> [®] – <i>Drive PX</i> object recognition system [64].	58
5.17 Face detection using DL [68].	60
5.18 Fully connected, feed-forward neural network with two hidden layers [65].	61

Figure	Page
5.19 Example of a resulting cost function's gradient [69].	62
5.20 Forward propagation operations on a two-hidden-layer network [65].	63
5.21 The <i>Chain rule</i> [69].	64
5.22 Parameters update operations on a two-hidden-layer network [65].	64
5.23 MVC architectural pattern [71].	66
6.1 Reflectance (%) of the corn response during different flights under normal nitrogen rates and no weed control [75].	68
6.2 Characterisation of the NDVI with low cost solutions [76].	69
6.3 Water reflectance content difference in the plant [77].	70
6.4 Diagram of co-existing organisms [80].	72
6.5 Structure and sensors locations of the greenhouse [81].	73
6.6 Connection scheme of the WSN nodes [82].	74
6.7 Mean spectral reflectance of Coffee Leaf Rust infection levels with Sentinel-2 bands [83].	75
6.8 Correlation coefficients and significance between the spectroradiometer wave bands (B2 to B8) and coffee leas rust severity on individual leaves [84].	76
6.9 UAV with a parrot sequoia cameras and the results among crops [90]. . . .	77
6.10 Agrocloud app for detection of disease among African Oil Palm [92].	77
8.1 Black box of the prototype.	84
8.2 Final concept of the prototype.	90
9.1 Coffee crop lots.	92
9.2 Crop lot filtering.	93
9.3 pH measurement system.	94
9.4 Rotary platform movement.	95
9.5 Rotary platform's extension (rack-pinion) mechanism.	95
9.6 Rotary arm.	96
9.7 Rotary arm movement.	97
9.8 Rain system.	98
9.9 Final modelling of the mechanical part of the prototype.	99

Figure	Page
9.10 Analog pH sensor/meter pro kit for Arduino® [94].	101
9.11 Lux sensor [95].	101
9.12 Waterproof soil temperature sensor [96].	102
9.13 Analog capacitive soil moisture sensor [97].	103
9.14 Environmental temperature/humidity sensor [98].	103
9.15 Dynamixel AX-12A actuator [99].	104
9.16 Digital servo shield for Arduino® [100].	105
9.17 Electronic components of the rotary platform's extension [101, 102]. . . .	106
9.18 Electronic components of the rain system [103–105].	107
9.19 Weather station kit [106].	108
9.20 Electronic components of the rotary arm [107, 108].	109
9.21 Buck-boost switching regulator [109].	109
9.22 Raspberry Pi® [110].	110
9.23 USB 3.0 super speed hub.	111
9.24 C270 Logitech HD webcam [111].	111
9.25 Survey3 multi-spectral camera [112, 113].	112
9.26 Functional scheme of Arduino 1.	115
9.27 Functional scheme of Arduino 2.	115
9.28 Data collection pipeline.	123
9.29 Data visualization dashboard.	125
10.1 Rotary arm's support manufacture.	131
10.2 Central tube's support manufacture.	132
10.3 Finished rotary arm's support structure.	133
10.4 Rotary arm's main base.	134
10.5 Assembled rotary arm.	135
10.6 Installation of the rotary arm.	136
10.7 Installation of the multispectral cameras.	137
10.8 Construction of the crop lots.	138

Figure	Page
10.9 Crop lots on their respective positions inside the greenhouse.	139
10.10 Connection to the water source.	140
10.11 Rain system's L-structure.	141
10.12 Slope towards the filtering point.	142
10.13 Assembly of the pH measurement system.	143
10.14 Motors, micro-switches and wheels on the rotary platform.	144
10.15 Web-cam assembly.	145
10.16 Rotary platform placed on a lot.	146
10.17 Example of individual connector assembly.	147
10.18 Electrical cabinets' location.	149
10.19 First steps of the electrical cabinet.	150
10.20 ATX 5V power source.	151
10.21 Individual testing of sensors and actuators from Lot data collection Arduino.	152
10.22 Sensors in position on the greenhouse.	153
10.23 Power supply junction box.	153
10.24 Circuit boards from Arduino 1.	154
10.25 Arduino 1 installed on the electrical cabinet.	155
10.26 Testing Pololu (DC-motor) with GECKO driver.	156
10.27 Arduino 2 installed on the electrical cabinet.	157
10.28 Position of the implemented potentiometer with the rotatory arm.	158
10.29 Electrical cabinet new distribution with the USB hub's.	159
10.30 Encapsulated LDR.	160
10.31 Internal structure of an automatically created subdirectory.	162
10.32 RGB images captured in the lot 3.	163
10.33 Original and rotated multi-spectral images from lots 1 and 3.	164
10.34 Multi-spectral images on the Raspberry Pi®.	165
10.35 Documents from the same routine stored in the Raspberry Pi®.	167
10.36 Lot 3 directory's content after a data collection routine.	169

Figure	Page
10.37Content comparison of the same subdirectory locally and remotely. . . .	171
10.38Example of preprocessed JSON, RGB, RGN and RE files.	174
10.39Built cyber-physical prototype for data collection.	178
11.1 Data distribution of the observations for both treatments.	182

ABBREVIATIONS

ACE	Alliance for Coffee Excellence
ADC	Academic Data Center
AI	Artificial Intelligence
ANN	Artificial Neural Network
ANOVA	Analysis of Variance
ASECC	Colombian Association for Coffee Excellence
BS	Base Station
CBD	Coffee Berry Disease
CENICAFÉ	National Coffee Research Center
CNN	Convolutional Neural Network
CRD	Completely Randomized Design
CSV	Comma-separated values
CV	Computer Vision
DIR	Design Inclusive Research
DL	Deep Learning
d.p.c.	Dry parchment coffee
FCP	Colombia in Peace Fund
FNC	National Federation of Coffee Growers
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
HTTP	Hypertext Transfer Protocol
IoT	Internet of Things
JSON	JavaScript Object Notation
LAI	Leaf Area Index

LDR	Light-dependent Resistor
M2M	Machine to Machine
ML	Machine Learning
MLP	Multilayer Perceptron
MVC	Model-View-Controller
NDVI	Normalized Difference Vegetation Index
NIR	Near Infrared
OOM	Out-of-Memory
PA	Precision Agriculture
PWM	Pulse-width modulation
RE	RedEdge
ReLU	Rectified Linear Unit
RGB	Red-Green-Blue
RGN	Red+Green+NIR
RNN	Recurrent Neural Network
RPA	Remotely Piloted Aircraft
RS	Remote Sensing
SENA	National Learning Service
SFTP	Secure File Transfer Protocol
SSH	Secure Shell
Tanh	Hyperbolic Tangent
UAV	Unmanned Aerial Vehicle
UPS	Uninterruptible Power Supply
UTC	Coordinated Universal Time
VHR	Very High Resolution
VNC	Virtual Network Computing
VOCs	Volatile Organic Compounds
VPN	Virtual Private Network
WSN	Wireless Sensor Networks

ABSTRACT

Sánchez A., Alejandro; Sarmiento G., Sebastián M.Sc., Universidad EAFIT, April 2019. Diagnosis evaluation of the Coffee Leaf Rust development stage in the Colombian Caturra variety integrating Remote Sensing, Wireless Sensor Networks and Deep Learning. Major Professor: M.Sc. David Velásquez Rendón.

Coffee Leaf Rust is a fungal epidemic disease that has been negatively affecting coffee trees in Colombia since the 1980s. At the national level, it has caused massive defoliation and, in extreme cases, this disease has led to devastating losses reaching between 70 % and 80 % of the harvest. In this regard, the early diagnosis of key patterns that may indicate the presence of the disease would contribute strategically to minimize the impact on the crops and therefore strengthen the farmers' profitability. Different studies have proven the benefits of using technological methods for identifying those patterns. Hence, in this research, a technological integration of Remote Sensing, Wireless Sensor Networks and Deep Learning was developed with the purpose of evaluating its performance regarding the diagnosis of the Coffee Leaf Rust development stage in the Colombian *Caturra* variety. To this end, a cyber-physical data collection system, that automatically gathers data from a scale coffee crop and transfers it to a remote server over the Internet, as well as a diagnostic model, which achieved an F_1 -score of 0.775278, were implemented. The analysis of results revealed a p -value of 0.231, which indicates that the difference between the disease diagnosis made by means of a visual inspection and through the proposed technological integration is not statistically significant.

Keywords: Diagnosis, Coffee Leaf Rust, Development stage, Remote Sensing, Wireless Sensor Networks, Internet of Things, Deep Learning.

1. INTRODUCTION

This chapter describes the problem that gave rise to the current thesis, presents the general and specific objectives and exposes the contribution that was achieved with this research project. The structure of the thesis is also explained at the end.

1.1 Problem description

Coffee is one of the world's most popular drinks and merchantable commodities. Every year, over 500 billion cups are consumed and over 158 million bags of 60 Kg are produced. Coffee is grown in around 70 countries around the world in a region known as *The Bean Belt*, which is located between the Tropics of Cancer and Capricorn, and the world's main producers are Brazil (2.72 billion Kg/year), Vietnam (1.65 billion Kg/year) and Colombia (810 million Kg/year). Furthermore, the social impact of the coffee growing industry is very significant because the people who depend on this activity for all or most of their living exceeds 100'000,000 worldwide [1].

The market is divided into two groups, known as the standard and specialty markets, according to the quality of the final product, which depends on the cultivated coffee variety, the environmental conditions and the post-harvest process. This is measured with a score between 0 and 100, known as the cup quality. When the cup quality is less than 80 points, the coffee belongs to the standard market and its selling price depends primarily on the New York Commodity Exchange. On the other hand, when the coffee has a cup quality greater than or equal to 80 points, it belongs to the specialty market and its selling price is at least twice the standard coffee price [2]. Nevertheless, it is a fact that coffee, which is cultivated with a view to the specialty

market needs a more careful and judicious agronomic management.

Regardless of the product's target market, coffee growers around the world face three major challenges nowadays to preserve quality: i) unpredictable climate variations, ii) presence of nutritional deficiencies and iii) attacks of pests and diseases. With respect to the latter, for instance, the Coffee Leaf Rust, which is a disease considered the main phytosanitary problem for coffee crops, causes in Latin America losses of 30 % on the efficiency of each harvest [3].

In the Colombian context, coffee is the most exported agricultural product followed by cut flowers, bananas, cocoa and sugarcane [4]. In the country, there are more than 903,000 hectares dedicated to it and approximately 563,000 families depend directly on this economic activity. The Colombian coffee has been considered one of the best soft coffees in the world and this product has traditionally been of great importance for Colombian exports. Nowadays, 14'000,000 bags of 60 Kg are exported every year to the USA, Japan, Germany, among others countries [5].

In terms of employment generation and income distribution, coffee growing is a sector with superlative relevance for local economies and the maintenance of the social fabric in many regions of the country. For this reason, it is justified to contribute through solutions that strengthen the profitability of families engaged in this activity and improve their life quality, either by increasing the selling price of the product, reducing production costs or increasing the number of units produced per unit of cultivated area.

One of the most relevant initiatives that promote the materialization of those solutions is the *Fondo Colombia en Paz* (FCP by its acronym in Spanish). It serves as the focal point for all institutional and financial efforts aimed at Colombia's transition towards a sustainable and durable peace. This instrument supports the realization

of economic, social and environmental dividends associated with the interventions on sustainable rural development, conservation of biodiversity and fight against climate change. For achieving this transition, the FCP is leveraged in the work of national and international institutions, which contribute to the same objective through 17 strategic axes [6].

One of those strategic axes is called *Innovation for Rural Development*, which consists in the realization of research and innovation projects for the development of agricultural and non-agricultural activities at the farm level, where efforts of the rural inhabitant, local entities and academy are articulated [6]. With the purpose of contributing to this particular axis, the present project was born.

Among the main threats for strengthening the coffee growing families' profitability, nutritional deficiencies and phytosanitary problems stand out. Phytosanitary problems are caused by pests such as the Coffee Borer Beetle and diseases such as the Coffee Leaf Rust, whose proliferation increases due to the drastic climate changes (from long drought periods to extended rainy seasons) that occur in Colombia. In the case of Coffee Leaf Rust, on occasions when the climatic conditions have been unfavorable and the agronomic management deficient, at least 20 % of the total expected harvest was not able to be collected and the quality of coffee has deteriorated dramatically, reducing the marketing price and increasing the costs associated with its control [7]. In extreme cases of Coffee Leaf Rust in Colombia, the disease has caused devastating losses that have represented between 70 % and 80 % of the total harvest.

Although it is a disease of vertiginous spread and highly negative repercussions for the coffee farmers' economy, its detection and diagnosis are carried out by means of visual inspections walking through the crops. The person in charge of administrating the plantations, which is usually the farm's butler, makes a zig-zag route through

each cultivated hectare and randomly selects 60 trees to count the number of diseased leaves. Subsequently, he divides that value by the total number of leaves present in the 60 trees and determines thereby the average infection in that particular area. Then, depending on how advanced the Coffee Leaf Rust is, he decides which and how much fungicide must be applied [7]. Unfortunately, because the process consists in a visual inspection which is not done with enough regularity, most of the time, the detection of the development stage of the disease is late, its control turns more difficult and large economic losses are inevitable.

In view of the fact that the determination of the infection percentage of the crop is a tedious task, which is also laborious, time-consuming and subject to human error and inconsistency [8], the objective of this research project is to evaluate to what extent it is possible to diagnose the Coffee Leaf Rust development stage in the Colombian *Caturra* variety (the most susceptible to the disease) through a technological integration system that involves Remote Sensing (RS), Wireless Sensor Networks (WSN) and Deep Learning (DL). An adequate management of Coffee Leaf Rust could preserve the quality and selling price of the final product, reduce production costs by rationing control costs and protect the productivity. In that sense, the purpose is to contribute from the academy to facilitate the management of the most dangerous disease in the Colombian coffee production in order to strengthen the profitability of the rural inhabitant.

1.2 Objectives

This section presents the general and specifics objectives of the thesis. Their definition is important to clarify the pursued purpose, guide the research process and evaluate the obtained results.

1.2.1 General objective

To evaluate to what extent it is possible to diagnose the Coffee Leaf Rust development stage in the Colombian *Caturra* variety through the integration of RS, WSN and DL in order to facilitate the disease management to the coffee grower.

1.2.2 Specific objectives

- To analyze the results obtained in interviews with experts, through the consolidation of the most relevant information, in order to detail the knowledge about the study context and the Coffee Leaf Rust, and identify its repercussions in the Colombian coffee growing industry.
- To define the most useful and relevant technological integration of RS and WSN based on patent and papers, through creative design sessions, in order to build a prototype that allows data collection in the field and its transfer to the Internet.
- To build the diagnostic model of the disease, through DL techniques using the data collected and stored on the Internet, in order to subsequently evaluate the Coffee Leaf Rust development stage in unknown field data.
- To create a simple user interface, through a web programming framework, in order to visualize the results of the Coffee Leaf Rust development stage diagnosis using the generated model.
- To compare the results of the diagnostic model with those of a traditional visual inspection, through the use of unknown data from the same scale coffee crop and from the same real crop, in order to measure the performance in terms of the calculated development stage of the disease.

1.3 Contribution

The motivation of this research project is to contribute to the rural development through technological innovation in order to strengthen the profitability of Colombian coffee growers. Colombian coffee participates strongly in the standard market but it has the potential in terms of environmental conditions and diverse ecosystems to generate a giant portfolio of exotic products which would be better valued in the specialty market [9], where sales are increasing by 20 % per year having Kenya and Ethiopia as the protagonists [1]. Nevertheless, most coffee growers prefer to cultivate varieties that are not the most suitable to produce specialty coffee because their resistance to Coffee Leaf Rust makes the agronomic management easier and the associated risk lower. In that sense, this Master thesis evaluates a technological approach that pretends to facilitate the diagnosis of the disease and mitigate its risk in order to motivate the cultivation of varieties like *Caturra*, that help to achieve higher cup qualities.

To support that contribution, this project provides empirical evidence of a prototype that integrates RS, WSN and DL, which materializes the proposed technological integration architecture for diagnosing the Coffee Leaf Rust development stage. Through this prototype, the performance evaluation is made and the results are presented in order to conclude to what extent it is possible to diagnose the disease in question. This information can be useful for coffee growers to determine whether an integration of RS, WSN and DL could positively impact their profitability or not.

The problematic that the project, object of this Master thesis, is addressing has an increasing interest, because the achieved results could have relevance at different levels. At the institutional level, *Universidad EAFIT* is working on ambitious projects related with energy efficiency and Precision Agriculture (PA) in the frame of a government call for improving the quality of the Colombian Higher Education Institutions named *Colombia Científica* (Scientific Colombia) [10]. Additionally, at

the regional level, the Department of Antioquia, where the project was developed, is the main agricultural producer in Colombia and thus, one of the main stakeholders in technological innovation processes for the rural development [11]. Finally, at the national level, there are several calls for research oriented to the use of science and technology in order to innovate and contribute to the construction of a sustainable and durable peace, such as *Jóvenes Investigadores e Innovadores por la Paz 2017* (Young Researchers and Innovators for Peace 2017) [12].

Indirectly, this project contributes to the 9th Sustainable Development Goal of the 17 proposed by the *United Nations*. It aims to support the goal “Build resilient infrastructure, promote sustainable industrialization and foster innovation” considering that the technological innovation promotes investments in infrastructure, which are crucial to empower the rural community and improve their living standards and activities in terms of progress, productivity and income generation. Finally, it is noteworthy that this Master thesis is aligned with the *United Nations* thought, which states that “without technology and innovation, industrialization will not happen, and without industrialization, development will not happen” [13].

1.4 Thesis structure

The thesis is structured according to the *Design Inclusive Research* (DIR) [14] and *Pahl & Beitz* [15] methodologies. Every chapter is intended to represent each of the particular stages that constitute them.

In that sense, the next chapter describes the methodology used to structure and develop this Master thesis. Chapter 3 of this document presents the observation that was helpful to understand the study context and the Coffee Leaf Rust repercussions. The research question is stated in Chapter 4 whereas the theoretical and conceptual framework and the state of the art are elaborated in Chapters 5 and 6, respectively.

Chapter 7 describes the design of experiment used to give statistical support to the obtained results. The conceptual approach as well as the detailed design approach are described in Chapters 8 and 9. Chapter 10 details the prototyping and presents the outcomes of the proposed approach facing the problematic. Chapter 11 contains the analysis of results and Chapter 12 states the conclusions. Finally, Chapter 13 deals with the future scope of this research with some insights into further extensions.

2. METHODOLOGY

This chapter presents the methodology that structures the research project. It is a fundamental part for the realization of this Master thesis because it points and details the different phases that systematize the achievement of results of the scientific study in question.

This research is driven by the *Design Inclusive Research* (DIR) methodology, whose objective is to provide a sound theoretical foundation and a robust methodological approach for designerly inquiry to meet scientific rigor. DIR systematically integrates the domains of research and design, offering the possibility to embed design as a research means (such as an artifact, a process, an entity, etc.) and creating new opportunities for exploring and constructing contextualized knowledge. In that sense, DIR was selected as the framing methodology for this project because it allows to structure the research in such a way that hardware, software, service products, among others, can be used to create testable instantiations that support the validity and feasibility proof of the initial ideas. The benefit of using DIR resides in the combination of rigorous research and creative design because it integrates knowledge of multiple domains and lends itself to multi-disciplinary approaches, which harmonize the internal goals of research with the societal context of designing [14]. The general process of DIR is presented in Figure 2.1.

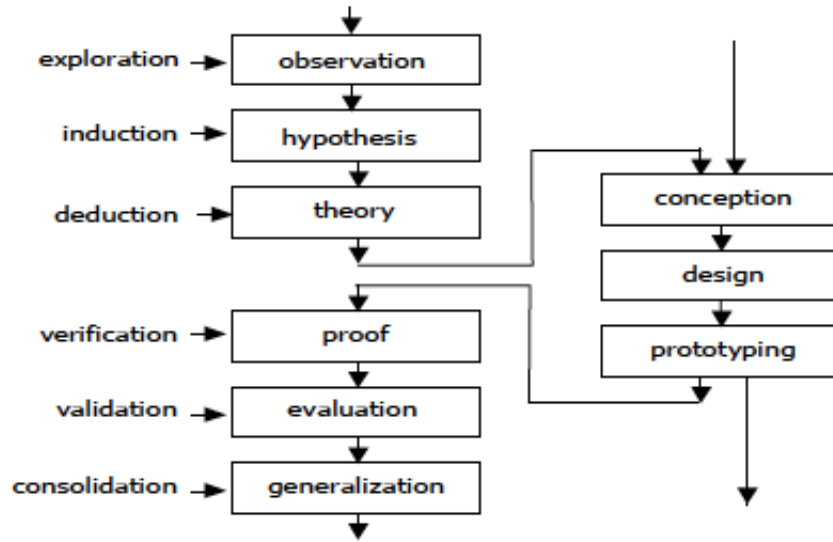


Fig. 2.1.: Major phases of DIR [14].

This methodology divides the research process into three phases: i) pre-study, ii) embedded design process and iii) post-study. The first phase is composed by the exploratory research actions (exploration, induction and deduction), where knowledge is aggregated, the current understanding of the phenomenon and existing approaches are formulated, the research question is asked, the hypotheses are developed, the objectives of the design activities are set and theories are built in order to solve the research problem. The second phase is recurrent and comprises the creative design actions (conception, design and prototyping), where concepts, models and technologies are invented, prototypes are created to proof the validity and feasibility of the ideas and experiments are carried out to enhance the understanding of the problematic. Finally, the confirmatory research actions (proof, evaluation and generalization) are executed on the third phase, where the hypotheses, theories and outcomes of the design process are verified, the external validity of the research findings is discussed and the results are consolidated, analyzing how they can be integrated to the existing body of knowledge and generalized towards other applications [14].

It is important to mention that for DIR's second phase (embedded design process), the *Pahl & Beitz* methodology was used. This methodology is correctly coupled to the mentioned phase because it is split into similar stages and is strategic to increase the probability of technical and economic success for the design process. *Pahl & Beitz* has an initial stage for planning and clarifying the task, which seeks to clearly establish the objectives of the solution to be conceived. There is another stage for the conceptual design, where several concepts for solving the task are generated and evaluated in order to select the winning one. A further stage is used for the embodiment design, where the selected concept is clarified and refined using external feedback. Finally, the last stage of the methodology is the detailed design, which specifies and structures the arrangement, forms, dimensions, properties, materials and technologies for the realization of the individual components that will constitute the refined concept. The end result of this methodology, as seen in Figure 2.2, has the necessary information for the materialization of the proposed solution [15].

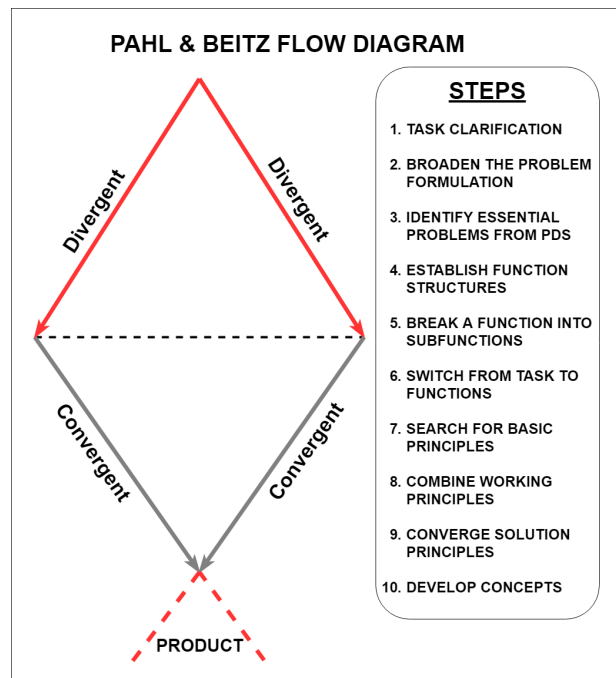


Fig. 2.2.: Diagram representation of the *Pahl & Beitz* methodology [15].

3. OBSERVATION

To start with the materialization of the research project in question, ten interviews were conducted with experts of the Colombian coffee industry in order to analyze and understand the study context. Directors of exports, along with agro-industry instructors, masters in agricultural science and agrobusiness, agricultural and mechanical engineers, entrepreneurs, phytopathology leaders and quality assurance coordinators, helped to better guide the research with their comments and experience in the Colombian coffee field.

Thanks to their explanations, it was possible to determine which would be the target coffee variety and disease combination for this study and which considerations would be important. The information collected with every interviewee as well as the consolidated conclusions are presented in this chapter.

3.1 Leonardo Henao

He is the director of exports of *Café de Santa Bárbara*[®] since 2011. This company is in charge of buying dry parchment coffee (d.p.c.) from small producers, threshing and roasting it and analyzing the coffee cup quality in order to commercialize/export only specialty coffees (more than 17,400 coffee bags per year). He has twelve years of experience regarding the coffee industry and studied a Master in Coffee Fermentation. 98 % of the company's coffee, on which they work with, comes from the *Caturra* variety due to its high quality and productivity. The company also buys *Bourbon*, *Geisha* and a little bit of *Colombia* but it has no interest in the *Castillo* or *Catimor* varieties, because they produce standard coffee. *Café de Santa Bárbara*[®] only buys coffee with a cup quality greater than 84 points since specialty coffee can access a

higher selling price (greater than 1'000,000 COP per 125 Kg bag) and it is not as volatile as standard coffee, which has been sold between 360,000 and 1'000,000 COP.

According to Leonardo, coffee growers seek to obtain certifications as *Rainforest* [16] or *UTZ* [17] to give added value to their product and protect it from the price fluctuation. Additionally, regarding the productivity and quality, he says that the biggest challenges are the Coffee Borer Beetle and the Coffee Leaf Rust, which increment their reproduction exponentially with the climatic variations. In Colombia, the Coffee Leaf Rust is the worst phytosanitary problem since 1983 and it can spoil a whole harvest. When the coffee suffers a severe attack of this disease, it may not meet the *Excelso UGQ 1260* norm (export quality requirements) [18] causing its selling price to drop drastically. Thus, coffee growers prefer to renew their plantations with the *Castillo* variety, which is resistant to the Coffee Leaf Rust, although it can harm their cup quality.

Leonardo explains that the Coffee Leaf Rust infection is calculated counting the number of affected leaves. In order to prevent the disease, sulfocalcium broth can be applied after the flowering and before the harvest as an organic fungus management. Traditionally, a commercial fungicide called *TIMOREX* or *ALTO 100 SL* is used. There is not a regular monitoring of the Coffee Leaf Rust but the crop inspection is carried out when there is a strong evidence or according to climatic variations. In addition, Leonardo knows that there are initiatives to detect the disease in a non-invasive way by analyzing the plant gas emissions or measuring the chlorophyll and nitrogen levels with electromagnetic radiation.

The Director of Exports of *Café de Santa Bárbara*® affirms that, in general, the Coffee Leaf Rust is detected in an advanced development stage, when the fungus has already sporulated. Under these conditions, the vascular system of the plant is invaded and the effectiveness of applying a fungicide is very limited. He explains that the risk

factors associated with the disease are a nitrogen-cooper ratio greater than 6:1 and high humidity conditions. Furthermore, the answer of the plant to the presence of the fungus is a decrease in the chlorophyll levels and photosynthetic activity, which results in defoliation.

For Leonardo, the best coffee variety to cultivate is *Caturra*, if the crop is located at minimum 1,600 m.a.s.l. and the climatic conditions are favorable. Otherwise, the recommended variety is *Castillo*, which is more productive, has lower production costs and will not be affected by diseases with the proper agronomic management. Either way, he thinks that the accompaniment to the coffee growers is one of the most important tasks, but he admits that it is impossible to visit each of them because of their amount.

Leonardo is concerned that the differential which is added to the commercialization of standard coffee according to the country is 12–15 USD cents for Colombia whereas it is 1.20 USD for Guatemala and 2.80 USD for Kenya. Besides, Colombia was the leader in cup quality until 1960, when the production came from *Bourbon*, *Typica* and *Caturra* varieties. After the renewal with *Castillo*, the coffee economic relevance decreased (from 65 % of the GDP in 1980 to 0.52 % nowadays) and the country lost the reputation of producing the best soft coffee worldwide.

3.2 Juan Carlos Guerra

He is a business administrator who has been working for the last ten years with coffee and its international commercialization. Juan Carlos is the owner of *Green Hills Coffee*[®], a family business founded five years ago. The production farms are located in Ciudad Bolívar, Antioquia. The company is in charge of producing, toasting, commercializing and exporting coffee and his owner trusts in adding value to the product

to achieve a differentiated result.

The variety Juan Carlos mainly works with is *Caturra*, because it is the one that has given him the best results regarding the relation between cup quality, productivity and sustainability. He also cultivates *Castillo*, *Red* and *Yellow Catuai*, *Red* and *Yellow Colombia* and *Bourbon*. Juan Carlos affirms that the *Caturra* variety is very susceptible to the Coffee Leaf Rust and therefore requires a better agronomic management. Furthermore, he mentions that the nutritional deficiencies due to poor fertilizations and the uncontrollable climatic variations (*El Niño* and *La Niña* phenomena) negatively affect the productivity and quality of the Colombian coffee. Moreover, the international price of Colombian coffee makes that the producer does not generate any profit many times and even loses money, which directly harms his sustainability.

Juan Carlos knows that the Coffee Leaf Rust is the main problem for the *Caturra* variety. In addition, he mentions that the Coffee Borer Beetle is transversal to all varieties and will always be present in the crops, which is why it is important to keep it under control. As the negative pest/disease consequences, Juan Carlos highlights that the excessive application of insecticides and fungicides can cause future rejections in the chemical analyses of the exported coffee. Additionally, the late detection of those phytosanitary problems requires shock plans to try to save as much as possible, which are more expensive, laborious and deteriorate the cup quality, affecting the coffee growers' profits noticeably. There are also situations, where the plant is so infected that a renewal by stumping is cheaper than trying to cure the tree, which at the end results in time and productivity loss.

To manage the Coffee Borer Beetle and impede its reproduction in the farms, the workers use a technique called "ReRe", consisting in collecting and repelling ripe coffee beans, or traps with methyl alcohol. Additionally, the workers carry out a

visual inspection every week to detect pests/diseases or nutritional deficiencies and they also perform various other tasks according to a calendar of activities (weed management, fertilization, fumigation, harvest, etc.). To determine if an application of pesticide/fungicide is necessary, the workers walk through the crop, take samples and measure the pest infestation or disease incidence levels.

Juan Carlos would like to incorporate more technology to his coffee company to characterize the status of the crops and decide what action to take in order to prevent, not to cure. The most technological tasks that he knows are the soil and foliar analyses to detect nutritional deficiencies, but he is concerned that it takes more than fifteen days to get the results. He affirms that technology is required for the Coffee Leaf Rust continuous monitoring and early detection, because of the drastic climatic changes. It would be very significant in order to avoid situations like the one in which the workers of the farms realized that 50 % of the crop was already infected by the disease.

The owner of *Green Hills Coffee*[®] mentions that excessive rainfalls are favorable to Coffee Leaf Rust onset, and hot and dry weather promotes the Coffee Borer Beetle reproduction. Juan Carlos highlights the importance of controlling the Coffee Leaf Rust, because it negatively affects the plant's photosynthetic activity. The disease reduces the radiation area, takes advantage of the leaves' nutrients and causes severe defoliation. Furthermore, he says that a poor production planning also increases the plant susceptibility to those phytosanitary problems.

Juan Carlos affirms that the best coffee variety is *Caturra* because it allows him to reach two times the price of the standard coffee, but he admits that the Coffee Leaf Rust management is the main problem associated to its production. He also clarifies that the cup quality depends not only on the coffee variety but also on the pre- and

post-harvest processes. Juan Carlos adds that the delimitation and identification of areas in the crop facilitates the agronomic management.

3.3 Jhon Wilson Mejía

He is an Agricultural Engineer and studied a Master in Agricultural Science. He worked for twelve years at the National Coffee Research Center in Colombia (*CENICAFE*) and nowadays works as a consultant in *Oro Molido S.A. / African Plantation Kilimanjaro*[®], a coffee company in Antioquia (Colombia), Santa Marta (Colombia) and Tanzania (Africa). The company is in charge of the whole process from the sowing until the time before the threshing. *Oro Molido S.A. / African Plantation Kilimanjaro*[®] cultivates mainly *Castillo* variety (90 % – 95 %), its regional derivatives and *Caturra*. It prefers the *Castillo* variety because it is resistant to the Coffee Leaf Rust and more productive.

According to Jhon Wilson, the main problems for the coffee production are associated with biotic (pests and diseases) and abiotic factors (nutritional deficiencies, plant intoxication by pesticides/fungicides, flooded or dry soils, etc.). In addition, the volatility of the standard coffee price and the USD exchange rate are risks to the final selling price. The production costs are negatively affected by the control of Coffee Leaf Rust and Coffee Borer Beetle, which are the main phytosanitary problems. With the favorable climatic conditions and poor quality agronomic interventions, these problems can also harm the productivity up to 70 %. Jhon Wilson explains that the Coffee Leaf Rust reduces the coffee quality and causes the loss of 20 % of the total harvest in five years. He also states that the presence of pests and diseases as well as stricter environmental regulations require developing strategies to contribute to the social and agro-ecological resilience.

Jhon Wilson mentions that diseases like the Coffee Leaf Rust and Brown Eye Spot and pests like the Coffee Borer Beetle, Coffee Leaf Miner and Red Spider damage the coffee grower's income function, because they act against the agronomy's objectives: the control cost increases, the selling price decreases due to bad ripening of the coffee berries, and the number of units produced per unit of cultivated area gets reduced. For instance, the Coffee Leaf Rust decreases the photosynthetic tissue, parasitizes the photoassimilates of the tree, which are used in the process of filling the grains, and reduces the absorption capacity of the coffee berry.

In order to protect the crops from diseases and take care of the sustainability, Jhon Wilson recommends planting resistant varieties, keeping the right nutritional status and using fungicides effectively. The application of fungicides can be preventive or curative and it is done according to a fixed calendar of activities or the measured incidence levels. To determine the disease incidence level, the farm workers walk every month through the crop, take ten trees per hectare and ten branches per tree with more than ten leaves. Having calculated that level, they relate it with the days after the flowering and decide which type of fungicide to use. If the disease monitoring is carried out monthly, it is possible to detect it in an early stage, which do not exceed the economic damage threshold (5 % – 10 %). An early detection allows the usage of protective fungicides to prevent the sporulation, which are 80 % cheaper than the curative ones to eradicate the fungus.

The consultant of *Oro Molido S.A. / African Plantation Kilimanjaro*[®] indicates that the nutritional deficiencies and the variable climatic conditions (temperature, relative moisture, sun brightness, rainfalls and wind speed) are closely related to the plant susceptibility to pests and diseases. The crop age and the sowing density are other important risk factors. Jhon Wilson mentions that it is important to know that the natural albedo (measure of the diffuse reflection of solar radiation) of the coffee crops

changes when the Coffee Leaf Rust attacks them.

Under normal circumstances, Jhon Wilson prefers the *Castillo* variety. Nevertheless, he admits that he would choose *Caturra* if the crop was located at minimum 2,000 m.a.s.l. However, he knows that the cup quality is a multifactorial variable since it is the result of the interaction between the coffee variety, the environmental conditions, the agronomic management, the harvest and the post-harvest processes. As additional information, he would like to have a study of images of his crop in order to detect the focus of infection, decide where to apply fungicide, avoid applying it uniformly over the whole plantation and save costs (every liter costs 45 USD).

3.4 Juan Carlos Posada

He is an Agricultural Administrator, Zootechnologist and he has a postgraduate's degree in Marketing Management. The National Federation of Coffee Growers in Colombia (*FNC*) invited him to participate in a coffee plantation renewal program and he began with 5,000 trees. Nowadays, he has 20,000 trees and exports his own brand coffee (*Nova Ola*[®]) by his own, not through Cooperatives. Juan Carlos received training in the production of specialty coffees, he experiments with diverse post-harvest processes and now he sells his product for 14 USD/lb. The owner of *Nova Ola*[®] works from the sowing until the international commercialization.

Juan Carlos cultivates a regional derivative of the *Castillo* variety named *Castillo - El Rosario*, one of the latest varieties developed by the *FNC*. It is very resistant to pests/diseases, premature and offers a cup quality of 85 points. He does not work with *Caturra* because it is more susceptible to pest/disease attacks, but he admits that it is one of the varieties that allow the coffee grower to reach the best cup qualities. He would be interested in cultivating it because he believes that the future of

the Colombian coffee industry is the production of specialty coffee.

For Juan Carlos, one the main problems in his job are the economic losses caused by the volatility of the standard coffee selling price. Additionally, he is concerned of the extreme climatic conditions in the country, due to which he had to replace 5,000 trees in 2016. Furthermore, he mentions that the coffee berry–d.p.c. ratio is unfavorable because it is 5:1 and then, 20 % gets lost during threshing and another 20 % during roasting.

Although Juan Carlos has never experienced pest or disease attacks in his crops at 1,900 m.a.s.l., he knows that the Coffee Borer Beetle and Coffee Leaf Rust are the main phytosanitary problems. They proliferate to a greater extent when the crops are at 1,600 m.a.s.l. or less and intensify the application of pesticides and fungicides. Moreover, the Coffee Borer Beetle harms the grain directly and the Cooperatives punish the purchase price, which negatively affects the coffee grower’s economy.

In order to protect the crops against pests and diseases, Juan Carlos prepares his own pesticide/fungicide using tobacco, hot pepper and garlic and applies them locally when there are signs or alarms. Additionally, he uses the “ReRe” technique every twenty days and carries out visual inspections. For Juan Carlos, the best way to protect the crops is the careful and regular monitoring. He does not use technology because he believes that it is unnecessary for the size of his plantation. The only technological method that Juan Carlos applies to each of his crop lots is a yearly soil analysis to schedule the necessary nutrition with an agronomist.

3.5 Juan Carlos Álvarez

He is a Mechanical Engineer and was always passionate about working in the countryside. His grandfather and father worked with coffee and he decided to continue

with the tradition. Nowadays, Juan Carlos is a coffee grower and he has farms in Andes and Betania, Antioquia. In those farms, coffee is cultivated, harvested and prepared for the threshing. Finally, it is sold to private buyers or to the Los Andes Cooperative, which is able to export the product. In the future, Juan Carlos would like to venture into the roasting business.

In 95 % of Juan Carlos' coffee crops, the *Catimor* variety (a cross between *Caturra* and *Timor Hybrid*) is sown. The remaining 5 % contains *Castillo* and *Caturra*. He prefers the *Catimor* variety because its root is strong and the coffee berries are bigger and more homogeneous. Although it is more productive, vigorous and easier to manage than the *Caturra* variety, Juan Carlos admits that he has sacrificed the cup quality. In the past, he decided to reduce the *Caturra* percentage in his crops because *Catimor* and *Castillo* varieties would need less agronomic management in terms of controlling the Coffee Leaf Rust, which would be better for the business. However, he noticed that the cup quality reached with the *Castillo* variety was very low.

According to Juan Carlos, one of the main problems by coffee growing is the management of Coffee Leaf Rust, because it proliferates with the spontaneous climatic variations and adapts even to the resistant varieties, causing severe defoliations. Due to this disease, he had to make once two applications of fungicide in a matter of three months. Additionally, Juan Carlos mentions that the standard coffee selling price and the USD exchange rate are very volatile and explains that, if any country presents overproduction, the coffee price would decrease immediately, due to the fact that the level of production is equal to the level of consumption nowadays.

Juan Carlos explains that there are two phytosanitary problems that are significant to the coffee grower's economy: the Coffee Borer Beetle and the Coffee Leaf Rust. For instance, he mentions that, in the past, 90 % of the coffee berries in his *Caturra* crops did not reach its ripeness due to the Coffee Leaf Rust. As a result, the quality

of the final product was very low and he had to sell it at low prices too. However, he highlights that his main economic losses have been caused by the Coffee Borer Beetle.

In order to early detect pests, diseases and nutritional deficiencies, the workers of the farms inspect the crops weekly and observe their general conditions. In most cases, they detect pests/diseases in an advanced stage, when the damage is visible and curative control is needed. In the past, they protected the plantations using copper-oxychloride but they suspended it due to high costs. In addition, Juan Carlos applies some soil analyses, but he admits that the results strongly depend on the extraction place and cannot be extended to the whole crop. He mentions that he would prefer an analysis of the farm's micro-maps that would indicate him what is missing or left over in each lot. Regarding technological methods for detecting and controlling pests/diseases, Juan Carlos has heard about drones that apply fungicide and satellite images that use the reflectance in different wavelengths to characterize the crop status without invading it.

Juan Carlos explains that the Coffee Leaf Rust as well as the Coffee Borer Beetle proliferate more under high temperatures and altitudes below 1,600 m.a.s.l. Furthermore, he mentions that the early detection of the Coffee Leaf Rust allows to make timely applications in order to avoid massive defoliation and reduce the damage of the disease. If Juan Carlos had the opportunity, he would like to cultivate *Caturra* again due to its suitability to produce specialty coffee with higher cup qualities. He sees the Coffee Leaf Rust management as the main risk, but he admits that he has some friends that continued working with that variety, when he decided to use *Castillo*, and achieve better economic results today.

3.6 Mario Duque

He is an Agricultural Engineer, works with coffee since 2006 and has a production farm in Ituango, Antioquia. In his farm, the activities range from the sowing until the commercialization and he sells the d.p.c. to the regional Cooperative.

Mario works with *Castillo* and *Colombia* varieties. He used to cultivate *Caturra* until 2012 but he was forced to eliminate it due to attacks of several diseases. Mario affirms that the varieties he cultivates now are resistant to the Coffee Leaf Rust, but he admits that the associated cup quality and selling price are not as high as the *Caturra* ones. Furthermore, he mentions that *Caturra* can be 10 % more productive than *Castillo* if the Coffee Leaf Rust is managed correctly. In addition, Mario explains that the *Colombia* and *Castillo* varieties present root problems and excessive elongation, respectively.

According to Mario, the main risks in coffee growing are the changing climatic conditions, which impact the coffee life cycle and the selling price fluctuation, negatively affecting the coffee grower's profitability. Additionally, he says that the productivity decreases under 400 grams per tree when the coffee plantations are old, causing further economic losses.

Mario mentions that the only really limiting coffee disease is the Coffee Leaf Rust, which attacks plants in any thermal floor, produces defoliation, impedes the total ripening and dries the berries. In terms of pests, the Coffee Borer Beetle is also economically limiting. Mario also states that the dynamics in coffee production were changed by the onset of the Coffee Leaf Rust, because the chemical management, which was very expensive and polluting, led to the replacement of traditional varieties by resistant ones. Due to this disease, complete tree defoliations occurred in Mario's crops in less than two months and 70 % of the production was constituted

by dried berries. The commercialization of the final product went down to a selling price equivalent to 20 % of the standard coffee price i.e. approximately 1,200 COP/lb.

In order to manage the Coffee Leaf Rust easily, Mario cultivates resistant varieties. Regarding the Coffee Borer Beetle, he applies the “ReRe” technique and uses some fungi when the infestation levels are high. The only localized chemical control is made during the seedling phase to protect the plant from the Damping-off disease. He inspects his crops to identify color, form and size changes in the leaves and prefers not to use chemicals, because they damage the biological balance. The only technological method that Mario incorporates in his crops are soil analyses every three years to diagnose the nutritional status and program the fertilizations.

According to Mario, frequent rainfalls are favorable to the Coffee Leaf Rust, which begins by attacking the bottom of the trees. The plant’s response to its presence is to abort leaves. In order to avoid critical phytosanitary problems, facilitate the agronomic management and preserve the cost-benefit relation, Mario plans to continue cultivating the *Castillo* variety.

3.7 Julio César Robledo

He is an Agricultural Engineer and works as the Quality Assurance Coordinator at the *FNC*. The *FNC* manages the resources of the *National Coffee Fund*, which draws from the tax that is charged to all the exported coffee in Colombia (6 USD cents/lb). This institution makes sure that coffee growers have a nearby Cooperative where they can sell their product at a fair price. Furthermore, they work together with *CENI-CAFÉ* to transfer more clearly the research results to the coffee growers and educate the rural families.

The *FNC* promotes the *Castillo* variety and its regional derivatives as well as *CENI-CAFÉ 1*, which is the most recent variety developed, because they can be cultivated in any coffee region of the country and they are more productive and resistant to the Coffee Leaf Rust and the Coffee Berry Disease. In addition, these varieties allow a higher sowing density and shorter cycles.

Julio César mentions that intense summers, rainfalls and natural disasters negatively affect the coffee production. Furthermore, he states that, although fumigations are carried out, the Coffee Leaf Rust is still the main phytosanitary problem when susceptible varieties are used. There are also bad post-harvest practices that cause waste and economic losses, along with environmental damage.

Among the pests and diseases that attack the coffee plant, Julio César highlights the Coffee Borer Beetle and the Coffee Leaf Rust. There are other organisms like the Coffee Leaf Miner, the Nematodes and the Pink Disease, but they are not economically limiting in Colombia. He affirms that the Coffee Leaf Rust is a critical problem for the specialty coffee production because it can cause economic losses between 30 % - 40% due to photosynthetic inefficiencies and severe defoliations. Moreover, the application of high doses of fungicide to control the disease impacts the quality of the final product and the food, water, air and health of the people living near the plantations.

According to Julio César, pests and diseases can be managed using resistant and adaptive varieties, good agronomic practices and, in extreme cases, pesticides/fungicides. He admits that this is the most difficult activity by coffee growing because those organisms develop resistance and mutate during their life cycle. Additionally, he recommends monitoring the climatic conditions frequently to identify whether they are favorable for the pest or disease proliferation and decide if there is necessary to intervene. The current method for detecting pests/diseases consists in a visual inspection,

where the person in charge walks through the crop and counts the number of infected plants to determine an affectation percentage. Another method is to use the flowering records to define critical moments to be alert to possible attacks. However, Julio César admits that not every coffee grower has the culture of registering his crop's development and taking actions based on real facts and information. Instead, the average coffee grower decides to use chemicals whenever any anomaly appears.

Regarding technological methods for detecting pests and diseases, Julio César knows that there are vaccination of substrates with antagonistic fungi and an application developed by *CENICAFÉ*, which performs a diagnosis based on photos of coffee leaves. He also mentions that there are satellite methods to measure crops reflectance intensities in the electromagnetic spectrum.

According to Julio César, the early detection of phytosanitary problems depends on the regularity of the care that every coffee grower gives to his plantation. He says that it is important to have a calendar of activities but, due to the drastic climatic variations that may alter the development of the plants, he affirms that the frequent monitoring is a better and more accurate alternative to avoid severe attacks. The actions have to be taken according to the real state of the crop and they should be preventive, not curative. Julio César adds that he would like to have an early alert system that notifies the infection levels of Coffee Borer Beetle and Coffee Leaf Rust, as well as the respective climatic conditions.

The intensification of climatic factors as temperature, rainfall and relative moisture propitiates the pest and disease proliferation, which increases the economic damage. In the case of the Coffee Leaf Rust, Julio César mentions that sufficient water and high temperatures are enough for its rapid spread. Thus, he prefers the *CENICAFÉ 1* variety, because it brings together all the research of *CENICAFÉ* and was specially designed to meet the requirements of the Colombian coffee industry.

3.8 Carlos Ariel Ángel

He is Head of Plant Pathology at *CENICAFÉ*, which is the leading institution worldwide in coffee research and exists since 1938. This research center promotes the crops renewal in order to only use resistant varieties. According to Carlos Ariel, 76 % of the Colombian coffee plantations consists of non-susceptible materials.

Carlos Ariel mentions that trying to manage susceptible varieties is a very hard task, which inevitably leads to economic losses. Instead, he recommends working with resistant varieties because it avoids spending money controlling diseases. However, he admits that even non-susceptible varieties as *Catimor* present attacks of Coffee Leaf Rust.

Among the main problems that impact coffee growing, Carlos Ariel highlights the malnutrition and climatic stress due to droughts and excessive rainfalls. As a result, pests like the Coffee Borer Beetle and diseases like the Coffee Leaf Rust and the Brown Eye Spot proliferate, increasing the production costs, decreasing the productivity and negatively affecting the physical and sensory quality.

In order to detect a phytosanitary problem like the Coffee Leaf Rust, Carlos Ariel mentions that the traditional methods are based on the plant's physiology, the flowering periods and the weather. Additionally, the *FNC*, which works closely with *CENICAFÉ*, inspects 6,000 lots quarterly at the national level to determine the infection of Coffee Leaf Rust and Coffee Borer Beetle, as well as nutritional deficiencies. Furthermore, Carlos Ariel says that some approaches using drones, spectral and thermal cameras and other sensors have been studied. As additional information, he mentions that the detection of the Coffee Leaf Rust is late when the disease is visible and countable.

For Carlos Ariel, it is essential to know the real coffee production system, understand the plant physiology and learn about the dynamics of the Coffee Leaf Rust before developing any solution against phytosanitary problems. For any development to be well supported, there must be an integration with the actual study context and its difficulties. Finally, he mentions that he would be interested in the remote measurement of physical variables using technology to aid the explanation of biological processes.

3.9 Andrés Ruiz

He is an Instructor of Agroindustry and Head of Specialty Coffee at the National Learning Service in Colombia (*SENA*). He is involved in the processes of adding value to the coffee i.e. not in the agricultural phase but from the post-harvest onwards until the final commercialization.

Andrés works with *Caturra* and *Caturra Chiroso* varieties because the cup quality is the main factor for him and he knows that these varieties have a big influence on it. The coffee coming from *Caturra* and *Caturra Chiroso* is excellent, exotic and differentiated, which allows to reach a higher selling price. However, Andrés says that quality depends not only on the variety but also on the agronomic management and post-harvest processes.

According to the Head of Specialty Coffee at *SENA*, one of the risks related to the sale of coffee is the dependence on the international price, which is very volatile. For that reason, he aims to the market of specialty coffee, where susceptible varieties are used. Nevertheless, he admits that the risk of pests and diseases increases and the management needs to be intensified. Andrés also mentions spontaneous climatic variations and malnutrition as problems associated with coffee productivity and quality.

Regarding pests and diseases, Andrés highlights the Coffee Borer Beetle and the Coffee Leaf Rust as the greatest threats. He says that the Coffee Leaf Rust is very aggressive, causes severe defoliations and inhibits the complete physiological ripeness of coffee berries. Due to this disease, curative controls using chemical applications have to be carried out and thus, the sensory quality of the final product gets negatively affected. When the attack is epidemic, coffee growers must even renew their crops by stumping or sowing.

On the one hand, Andrés mentions that the traditional method for detecting pests and diseases consists in visual inspections of the plantations. On the other hand, he says that there are technological approaches, that use drones in other crop types, for detecting the focus of an infection and preventively notifying the administrators. In addition, he knows that there is a research project that aims to manage the Coffee Leaf Rust by controlling the plant's nutrition. Andrés says that he would like to have a system for the localized control of pests and diseases and fertilization, based on real information, in order to increase the productivity and quality. He believes that a coffee grower can use susceptible varieties as long as he takes conscious care of his crops regularly.

The Head of Specialty Coffee at *SENA* prefers the *Caturra* variety because his expectations are high cup quality, variety of flavors and exoticism. He admits that the risks are greater but he says that he is interested not only in coffee but also in the generation of experiences around it.

As additional information, Andrés mentions that the market of specialty coffee grows 30 % every year whereas the market of standard coffee grows only 2 % in the same period. Furthermore, he explains that Colombia dramatically decreased the selling price of its coffee, because the country decided to focus on the standard market, and, nowadays, it is not able to compete against specialty coffee from Kenya, Ethiopia or

even Guatemala. He is concerned about this issue because he states that Colombia has the environmental conditions and ecosystems with which a very broad product portfolio could be generated. For Andrés, the Colombian coffee industry should start pointing to the market of specialty coffee, where the demand is lower but the exclusivity is greater.

3.10 Juliana Maya

She is an Electronic Engineer, Specialist in Agroindustrial Management and Development and she has a Master in Innovation in Agribusiness. Juliana works as the Director of Production in a family business founded eleven years ago, the farms are located in Salgar, Antioquia and they contain 2'000,000 coffee trees. The company is in charge of the whole process from the sowing until the commercialization and export to Russia, Italy and USA. Juliana has also her own brand coffee and store named *Café Terrón*[®].

The variety Juliana works with is *Castillo* but she also has a small percentage of *Tabi* since 2015. She mentions that those varieties are semi-resistant to the Coffee Leaf Rust, their adaptability is high and their management is easier considering the extension of her crops (300 hectares). The family business produces 875,000 Kg/year and its core is the standard coffee.

For Juliana, the main ally and enemy by coffee growing is the climate. In addition, she mentions that the majority of coffee growers in Colombia are unorganized, they fertilize at undue moments and fumigate when pests and diseases are already in a very advanced stage. Juliana affirms that the coffee growing industry needs more technological development because the innovation in the last seventy years has been minimal.

Juliana explains that there are two main phytosanitary problems that must be always controlled: the Coffee Leaf Rust and the Coffee Borer Beetle. Regarding the Coffee Leaf Rust, Juliana knows that the disease caused once a reduction of 60 % of the harvest in Andes, Antioquia, which decreased the selling price drastically. Additionally, in 2015 the Coffee Borer Beetle negatively affected the quality of Juliana's coffee and the company suffered economic losses of 350'000,000 COP. Other difficulties that must be taken into account are the Red Spider and the nutritional deficiencies.

According to Juliana, the applications of fungicide for controlling Coffee Leaf Rust are usually made when its presence is observed while inspecting the crops, as well as considering the records of rainfall and flowering. The traditional method for detecting phytosanitary problems consists in regular visual inspections using standardized protocols. In addition, there is a calendar of agronomic activities for every farm. Besides, Juliana mentions that, although she does not use technology in her plantations for managing diseases, she knows that there is a technological park for coffee research called *TECNICAFÉ*, which monitors the climatic conditions in every lot and works in innovation projects with *CENICAFÉ*.

Juliana affirms that prolonged rainfalls and droughts, together with nutritional deficiencies, are the main risk factors for the proliferation of pests and diseases. Thus, she prefers to cultivate *Castillo* because she knows that her production is massive and she wants to keep it under control. However, Juliana admits that the early detection of pests and diseases is important to reduce economic losses.

3.11 Summary

Summarizing the content of the ten interviews, it is possible to notice that there are two positions regarding coffee growing. One of them is preferring the *Caturra* variety and aiming the specialty coffee market. The other one is preferring the *Castillo*

variety and pointing the standard coffee market. *Caturra* is severely affected by the Coffee Leaf Rust and needs more care whereas *Castillo* is resistant to that disease and the agronomic management is easier. However, the cup quality and selling price that can be reached using *Castillo* are generally not as high as the ones associated with *Caturra*.

In addition, the phytosanitary problems by cultivating coffee, that have the highest impact on the coffee grower's profitability, are the Coffee Leaf Rust, the Coffee Borer Beetle and the nutritional deficiencies. The most common method for detecting these problems consists in walking through the crops carrying out visual inspections and using standardized protocols. Nevertheless, the visual inspections are normally not very frequent and there is a high probability of detecting phytosanitary problems in advanced stages, which limits the effectiveness of their control. Furthermore, it is noteworthy that drastic climatic variations in Colombia as well as poor agronomic managements are critical risk factors for the increase of pests, diseases and nutritional deficiencies.

Finally, the comments and experiences of the ten interviewees helped by selecting *Caturra* and Coffee Leaf Rust as the target variety and disease combination for this study. Being able to diagnose the disease in an early development stage would be highly relevant for strengthening the coffee grower's profitability.

4. RESEARCH QUESTION

After understanding the severe negative impacts of the Coffee Leaf Rust on the world-wide coffee production and knowing how this disease harms the coffee grower's profitability in the national context, an initial literature review was carried out. Its purpose was to examine existing approaches for diagnosing the Coffee Leaf Rust and conceive, in general terms, how this problematic could be effectively faced by the project at hand.

As a result, a research question was stated in order to clearly establish in which direction it was necessary to point the study and where the efforts needed to be concentrated. In that sense, the question that motivated the realization of this Master thesis as well as the initial idea regarding the possible outcome are presented in the following lines.

To what extent is it possible to diagnose the Coffee Leaf Rust development stage in the Colombian Caturra variety through the integration of Remote Sensing, Wireless Sensor Networks and Deep Learning?

4.1 Research hypothesis

Through the integration of RS, WSN and DL, it is possible to diagnose the Coffee Leaf Rust development stage in the Colombian *Caturra* variety with significant performance.

5. THEORETICAL AND CONCEPTUAL FRAMEWORK

There are several relevant definitions and explanations that are necessary for understanding the study context and serve as the conceptual background of the theory that was reviewed and used for the realization of this research project. The information presented in this chapter is intended to explain the main agricultural and technological concepts that were found in the literature review, as well as the ones that were actually incorporated for the development of this Master thesis.

5.1 Agro

5.1.1 Coffee

For more than a 1,000 years and even today, the most important consumed drink around the world has been coffee, with more than 500 billion cups per year [19]. More than a 100 species belonging to the genus *coffea* exist, but only two are used for the preparation of the drink, namely *Coffea Arabica* and *Coffea Canephora*. The former is considered as the most delicate of all coffee species and provides 75 % of the worlds coffee production, while the latter, considered to be more acid but more resistant to pests and diseases, provides the remaining 25 % [20]. On the one hand, *Coffea Arabica*, indigenous from Ethiopia, develops healthy and optimal at high altitudes (600–2,000 m.a.s.l.). On the other hand, *Coffea Canephora* plantations adapt well to altitudes lower than 600 m.a.s.l [21]. This is why the only cultivated species in Colombia is *Coffea Arabica*, since the country’s rugged topography is ideal for it and offers variety of climates bonded with high altitude zones [22]. The Colombian coffee plantations are mainly distributed among three regions (Northern, Southern

and Central), as seen in Figure 5.1 [23].

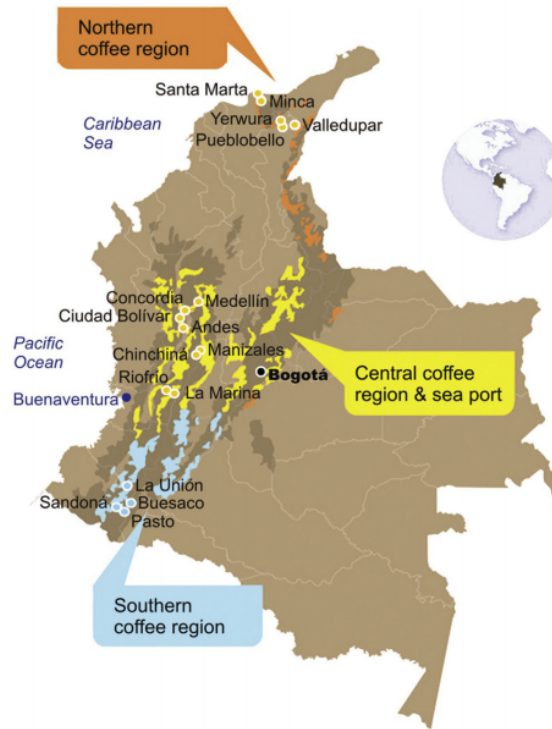


Fig. 5.1.: Map of Colombian coffee regions [24].

5.1.2 Life cycle and phenological phases of coffee

The success of the produced coffee relies strongly on the quality of its growth. A coffee tree has around 20–25 productive years, depending on the conditions of the crop system. Its productive life starts on the first year after the sowing and last for 24 more periods with mandatory renewals by stumping every 6–8 years. The Figure 5.2 illustrates the first three years of coffee's life cycle. During the first year (aka vegetative development), the formation of non-reproductive structures like roots, branches, knots and leaves occur. After this period, the reproductive development begins, where the process of flowering, fruit growth and ripening takes place. Besides, the phenological phases of coffee strongly depend on the rainfall dynamics. The formation of

knots, leaves and fruits is related to humid periods whereas flowering occurs at the end of dry seasons. Lastly, after approximately seven years of activity, the tree starts decreasing its productivity towards levels of low profitability (aka senescence phase) and thus, should be stumped to begin a new cycle [25].

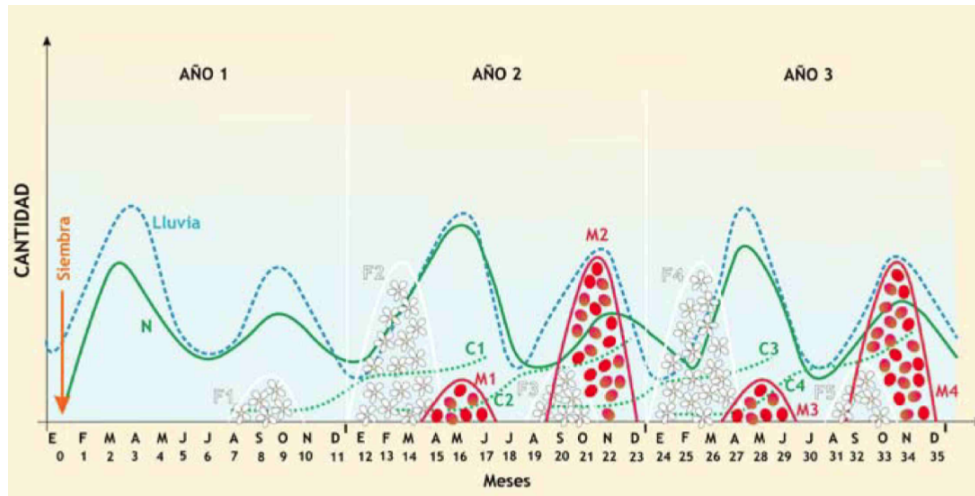


Fig. 5.2.: Phenological phases of coffee [25].

The availability of water and sunlight in the coffee regions and the changes in temperature, along with the interaction with genetic, nutritional and hormonal factors of the coffee tree, determine the pace of growth of the plant's tissues and functional structures. This growth begins with the seedling stage, which takes two months, continues with the “Almácigo” stage, which starts with the transplant of the seedlings to sacs and lasts up to four months, and finishes with the sowing on the field. After 5–6 months, the first flowering occurs and the tree enters its reproductive development. It should also be noted that the speed of the senescence phase depends on the intensity of production, the availability of nutrients, the presence of pests/diseases and the environmental stress [25].

5.1.3 Coffee varieties

The *Coffea Arabica* species is sub-divided among different varieties with specific characteristics. Those varieties are classified according to the type of bearing, which determines the number of units produced per unit of cultivated area. The Table 5.1 describes the main characteristics of the *Coffea Arabica* varieties present in Colombia, namely *Typica*, *Bourbon*, *Maragogipe*, *Tabi*, *Caturra* and *Colombia* [26].

Table 5.1.: Main characteristics of the coffee varieties cultivated in Colombia [26].

CHARACTERISTICS	TALL BEARING				LOW BEARING	
	<i>TYPICA</i>	<i>BOURBON</i>	<i>MARAGOGIPE</i>	<i>TABI</i>	<i>CATURRA</i>	<i>COLOMBIA</i>
Leaves color ("Cogollos")	Tan or reddish	Green	Green	Green or tan	Light green	Tan
Susceptible to <i>Hemileia vastatrix</i>	Yes	Yes	Yes	No	Yes	No
Grain size (% of <i>Supreme</i>)	70 %	46 %	90 % - 100 %	80 %	66 %	80 %
Grains in pod	4	4	4	3,7	5	5
Sowing density (trees per hectare)	2.500	2.500	2.500	Until 3.000	Until 10.000	Until 10.000
Production per tree (in the sun) [Kg]	0,9	1,2	0,7	0,7	0,5	0,5
Production per hectare [Kg]	180	240	140	140	169 - 272	400
Height (24 months) [cm]	226	237	> <i>Bourbon</i>	243	137	165
Canopy diameter (24-30 months) [cm]	180	177	> <i>Typica</i>	200	112	133
Sowing density (in the sun)	2.500	2.500	2.500	2.500	Until 10.000	Until 10.000
Sowing density (with 25% - 50% shadow)	1.500	1.500	1.500	1.500	Until 4.000	Until 4.000

The *Colombia* variety comes from a cross made in Colombia between *Caturra* and *Timor-Hybrid*. It is similar in shape, tree size and productivity to the *Caturra* variety, its leaves are brown and it is resistant to the Coffee Leaf Rust. In addition, its sowing density is about 10,000 trees per hectare, it has an acceptable grain density and a good cup quality. The *Castillo* variety, along with its regional derivatives, is the product of decades of genetic improvement of the *Colombia* variety. The name of these derivatives are *Castillo-Naranjal*, *Castillo-El Tambo*, *Castillo-Santa Bárbara*, *Castillo-El Rosario*, *Castillo-La Trinidad*, *Castillo-Paraguaicito* and *Castillo-Pueblo Bello*, which have been adapted to the particular conditions of each coffee region in

the country. *Castillo* is also resistant to the Coffee Leaf Rust, its productivity, grain size and sowing density are high and its cup quality is good [26,27].

For its part, the *Tabi* variety is derived from the cross between *Typica*, *Bourbon* and *Timor-Hybrid*. Its trees are tall, their leaves are green and it has a resistant component against the Coffee Leaf Rust. *Tabi*'s sowing density is about 3,000 trees per hectare, its productivity is lower than the national average but its cup quality is very good. The *Typica* variety (aka *Pajarito* or *Nacional*) is characterized by tall bearing with elongated red-brownish leaves. Although its susceptibility to the Coffee Leaf Rust makes it less productive, this variety provides large grains, its sowing density is between 2,500 and 3,000 trees by hectare and its cup quality is exceptional. Moreover, the *Bourbon* variety is also known for being of tall bearing, its leaves are rounded and light green and it is vulnerable to the Coffee Leaf Rust too. *Bourbon*'s grain size is average and its productivity is better than the *Typica* one. However, in terms of sowing density and cup quality, this variety is very similar to it [26,27]. The *Caturra* variety, which is the target variety of this research, is detailed in a separate section on this chapter.

5.1.4 Caturra variety

This variety was originated in Brazil as a natural mutation of *Bourbon*. It is characterized by low bearing (1.80 m), small internal knots, high stem thickness and short but abundant branches, which make it look vigorous and compact. Its leaves are large, wide, rounded and darker than the *Bourbon* ones. Additionally, *Caturra* has a very well developed root system and produces large and dense berries. Its nutritional and agronomic management requirements are high, and its sowing is recommended for altitudes over 1600 m.a.s.l. This variety is still having relevant presence in Colombia because of its outstanding productive potential (272,169 Kg/ha). Furthermore, its cup quality is excellent, as the first five winners of the contest "Cup of Excellence"

2017 in Colombia suggest, who used *Caturra*, a blend of *Caturra* and *Castillo*, and *Geisha* [7, 27]. The Figure 5.3 shows a coffee tree of the *Caturra* variety.



Fig. 5.3.: *Caturra* coffee tree.

Approximately 24 % of the coffee areas in Colombia are sown with susceptible varieties, mostly *Caturra*. Coffee growers, who work with this variety, know that combining their genetic properties with the appropriate environmental offer and a judicious and timely agronomic management, it is possible to obtain extraordinary cup qualities that allow them to sell a differential product [26].

5.1.5 Specialty Coffee

The excellence of the coffee beverage relies on many factors, including genetic origin, altitude, latitude, climate, phytosanitary care, agronomic management and quality of the harvest practices. Even the post-harvest processes, among threshing, storage, roasting, milling and type of preparation, can enhance or deteriorate the characteristics of the final product. The quality of coffee is defined giving a score (from 0 to 10) to each element belonging to its sensory features, such as fragrance, aroma, flavor, acidity, body, balance, sweetness, uniformity and cleanliness. Summing those scores

up gives a resulting score in the $[0, 100]$ range, known worldwide as the *Coffee Cup Quality*. In essence, the cup quality is directly proportional to the coffee's selling price and it defines the target market for the obtained product. Thereby, if the calculated score is under 80 points, the coffee is classified as standard. Otherwise, the product can be sold in the specialty coffee market [28]. The Table 5.2 summarizes the mentioned ranking.

Table 5.2.: Ranking for Specialty coffee [2].

Score	Quality	Classification
90.00 - 100.00	Outstanding	Specialty
85.00 - 89.99	Excellent	
80.00 - 84.99	Very Good	
< 80.00	Below Specialty Quality	Not specialty

Proof of the international significance of the cup quality is the “Cup of Excellence” contest. It is considered the most prestigious contest for selecting the best coffee worldwide and seeks for the most unique specialty coffee based on its taste, smell and the experiences that its consumption generates on the judges. With the support of the Colombian Association for Coffee Excellence (ASECC) and the Alliance for Coffee Excellence (ACE), the contest guarantees a vast variety of samples from all over the world to be able to compete for the highest recognition [29].

5.1.6 Coffee diseases

For producing a high quality product, coffee requires a judicious and precise agroeconomic management during its development. Nevertheless, there are several biotic and abiotic factors that can negatively affect the coffee crops and hinder the achievement of the expected results. For example, a healthy leaf can last an average of 10–15 months on a coffee plantation under shade conditions, but 9–14 months under full

sun exposure. This happens because the constant exposure to the sun can harm the follicular development of the plants, which is essential for the photosynthesis. In addition, drastic climatic variations can cause premature senescence and loss of foliage due to the interruption of the necessary humid/dry periods. Malnutrition issues, such as nitrogen and magnesium deficits, may also induce lower chlorophyll levels, resulting in severe defoliation [30].

In the case of the presence of pests and diseases, the mentioned conditions aggravate their consequences, which may lead to large scale production and quality losses, complete defoliation and even death of the coffee trees. The diseases, for instance, reduce the radiation area on the leaves, impeding the efficient absorption of essential components (nutrients and water) by the roots for photosynthetic processes. As a result of the blocked substances flow inside the plant, the filling of the grains slows down, the diseases consume the vital components and the ripening cannot be completed [30]. The main coffee diseases are explained below. The Coffee Leaf Rust, which is the disease to diagnose through this project, is presented in a separate section on this chapter.

- **Tip-over:** it is a fungal disease that rots the seed or the young stem before the seedling emerges. The seedling displays dark-brown patches, which weaken its tissues and cause the initial leaves to fall. These injuries are evident 15 days after the sowing [31].
- **Brown Eye Spot:** it is a disease that commonly manifests on the leaves and it occurs both on seedling and adult plants. It is characterized by brown spots surrounded by a yellow halo. Furthermore, the Brown Eye Spot may attack unripe fruits, causing them to dry out, lose weight or fall prematurely, which negatively affects the quality and quantity of the harvest. The disease can even

dry the mucilage of the berry, increasing the production of non-exportable coffee (aka “Pasilla”) [32].

- **“Mal Rosado”**: this disease causes severe damage in production, especially within areas of high sowing density. Generally, it occurs when the oxygenation is low and high precipitations as well as cold-humid environmental conditions predominate. This pathogen develops on dried stems and on the leaf litter that accumulates in the lower part of the bushes [33].
- **“Llaga Macana”**: it is a fungal disease that affects the stem and causes death among the plants in all their life cycle stages. As a result, the population of coffee trees decreases, as well as the harvest volume. The whole cultivated varieties in Colombia are susceptible to it. In addition, it also negatively affects rubber, citrus and cocoa plants [34].
- **Descending Death**: it is a disease that appears in all coffee varieties during the renewal of the plants’ leaves, hinders the production and causes harvest reductions. The Descending Death begins with the appearance of black and opaque small spots on the newly formed leaves and continues covering them until reaching their entirety, which generates premature defoliation. Both seedling and adult plants are negatively affected by this disease [35].
- **“Ojo de Gallo”**: it attacks the coffee plantations with excess of shade and high percentage of environmental humidity. This fungal disease, which can appear during the whole life cycle of the coffee plant, begins with the presence of dark-brown spots, followed by perforations and defoliation of the bush. It also weakens the nutritional status of the trees, which directly impairs productivity [36].
- **“Crespera”**: this disease negatively affects the normal development of the plants, especially their leaves and flower buds. It is characterized by the appearance of curved and thin leaves that grow massively, replacing the healthy

leaves and damaging the floral meristems. It arises mainly in the oldest tree branches and, in very severe cases, it can cause dwarfism on the plants by reducing the distance between knots and disabling the flowering [37].

- **Coffee Berry Disease:** although this disease does not exist in Colombia yet, it is considered one of the most limiting phytosanitary problems for *Coffea Arabica* plantations, along with the Coffee Leaf Rust. This is based on the fact that this disease can lead to losses up to 80 % of the production. The Coffee Berry Disease, which directly attacks the grains and occasionally the near leaves, has most severe consequences in plantations between 1,200 and 2,000 m.a.s.l. [38].

5.1.7 Coffee Leaf Rust

The fungus *Hemileia Vastatrix* is the cause of the Coffee Leaf Rust disease, which is the major phytosanitary problem for coffee crops. Once high levels of severity are reached, the corrective actions can be very limited. An inappropriate management of the disease can harshly compromise the coffee plants, as seen in Figure 5.4, generating that only few leaves remain on the trees, which has a direct negative impact on the quantity and quality of the harvest [7].



Fig. 5.4.: Coffee Leaf Rust effect on the *Caturra* variety [7].

Four essential factors are necessary for the appearance of the Coffee Leaf Rust: i) the host, ii) the pathogen, iii) the climatic conditions and iv) the agronomic management. The host implies that the coffee variety must be genetically susceptible to the disease and its physiological state is propitious for the fungus. The pathogen agent, which triggers the Coffee Leaf Rust, is the fungus *Hemileia Vastatrix* – Race 2. This fungus parasitizes living plant's cells reaching the reproductive stage in approximately 50 days and it is unable to survive among the soil or inert organic matter. Regarding the climatic conditions, constant rainfall is necessary for the sprouting of the fungus, which subsequently begins spreading on leaves, as seen in Figure 5.5, and surrounding plants. Additionally, temperatures between 16 °C and 28 °C, along with low solar brightness, are favorable for it. With respect to the agronomic management, high sowing densities and the incorrect application of fungicides are suitable factors for the development of the Coffee Leaf Rust. *CENICAFÉ* provides technical information to coffee growers in this regard and recommends monitoring the disease regularly independently of the crop altitude [7].



Fig. 5.5.: Coffee Leaf Rust on a leaf in its highest development stage [7].

Coffee Leaf Rust progresses gradually in time and reaches three noticeably phases. The first one, called “slow phase” (severity $\leq 5\%$), is where the first structures

responsible for the production of spores emerge and low levels of infection are evident. The second one, which is named “fast or explosive phase” ($5\% < \text{severity} \leq 30\%$), starts with the fungus sporulation and is represented by more plants getting sick in a short period. The final phase is called “maximum or terminal phase” ($\text{severity} > 30\%$) and occurs when most of the leaves are severely attacked and a small amount of healthy leaves remains. At that moment, the epidemic stops on the host, due to the lack of biological matter for continuing the infection. When the Coffee Leaf Rust is not controlled and the climatic conditions are favorable, the disease can develop at a daily rate of $0.19\% - 0.38\%$, reduce the impact of the chemical controls and cause significant economic damage [7].

5.2 Technology

5.2.1 Internet of Things

The Internet of Things (IoT) refers to the use of intelligently connected devices and systems to leverage data gathered by embedded sensors and actuators in machines and other physical objects [39]. It enhances the ability to distribute and analyze the collected data for transforming it into relevant information, knowledge and appropriate actions in the real world.

IoT enables, for example, GPS-based tracking, controlling and monitoring of devices, machine-to-machine (M2M) communication, connected cars, etc. [40] and allows, by means of the interoperability through the Internet, improvements in decision-making in several fields like energy efficiency, security, health, education, manufacturing, retail and agriculture [39]. IoT makes everyday things “smart” because it gives them the ability to sense, compute and interact with remote servers, applications, processes and persons.

For instance, as shown in Figure 5.6, smart health sensors can be used to collect and send comprehensive physiological information to the Internet, in order to analyze it and communicate the results wirelessly to caregivers for further review [41].

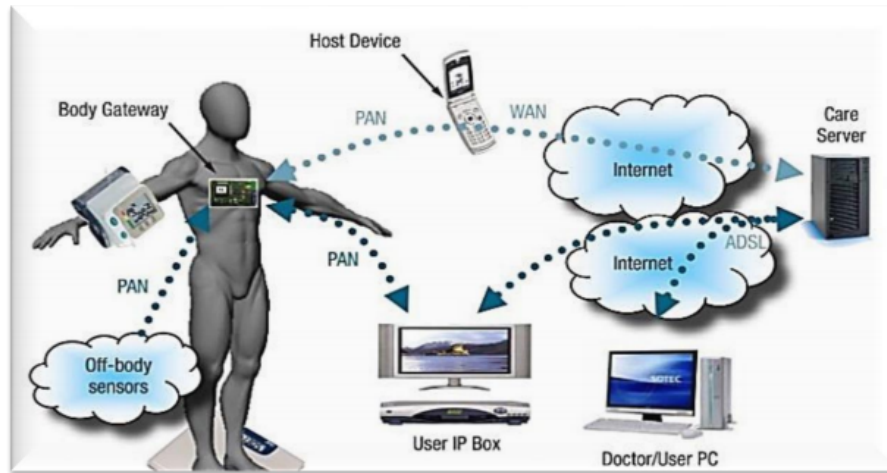


Fig. 5.6.: Smart healthcare concept [41].

5.2.2 Precision Agriculture

Precision Agriculture (PA) is a farming management concept that uses information technology to observe, measure and respond to crop site-specific variability. It consists in applying the right treatment in the right place and at the right time in order to adjust the agronomic management to the actual crop needs. In general, PA contributes to the wider goal concerning sustainability of agricultural production [42].

Through the analysis of satellite positioning (GNSS) and RS data, as well as data gathered using sensor technologies, the farmer can optimize the yield on inputs (machinery, labor, fertilizer, chemicals, seeds, water, energy, etc.) while reducing environmental impacts [42]. Facing the increasing demand for agricultural products and considering the drastic environmental challenges nowadays, PA can positively affect important variables like performance, productivity, quality, profitability, sustainabil-

ity and working conditions.

As an example, the Figure 5.7 shows how RS methods, based on manned and unmanned vehicles (RPAs and UAVs), enabled the acquisition of very high resolution (VHR) thermal imagery to provide maps of the spatial variability of water stress at the tree level [42].

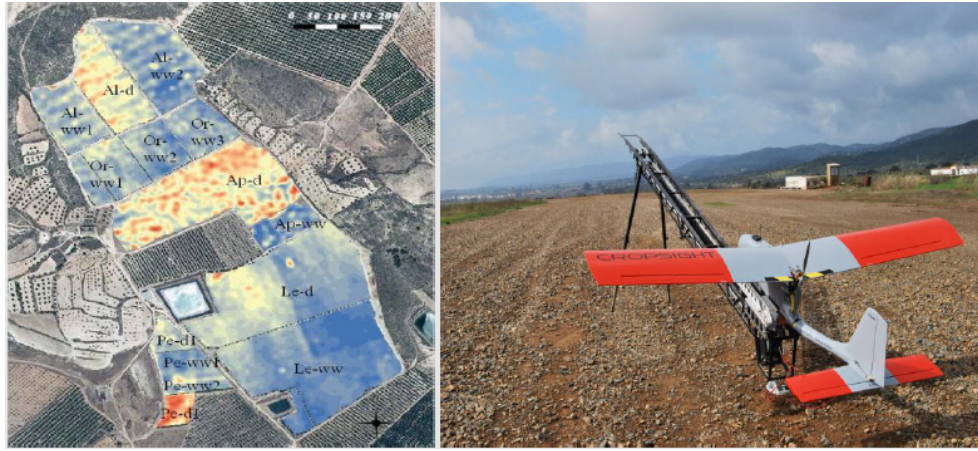


Fig. 5.7.: Evaluation of the opportunities of regulated deficit irrigation for saving water in southern Spain [42].

5.2.3 Serological analysis

The serological analysis is a method that is carried out in the laboratory for the diagnosis of phytosanitary problems. It consists in the direct detection of a pathogen from the analysis of the specific reactions that occur between the antigens, the proteins coming from the infectious agent causing the disease, and the antibodies present in the organism under study [43]. The antibodies are the immune response of an organism when it is attacked by an external agent.

It is noteworthy that specific antigens causes the production of specific antibodies, which will only generate a reaction by recognizing the presence of those particular antigens. Some of the techniques of serological analysis are Enzyme-Linked ImmunoSorbent Assay (ELISA), Tissue Blot ImmunoAssay (TBIA) and Quartz Crystal Microbalance Immunosensors (QCMD) [43]. The Figure 5.8 shows an example of a simple ELISA kit that was developed for detecting six virus species in infected plant tissues [44].



Fig. 5.8.: Simple ELISA kit developed at the Plant Virus Laboratory from Federal University of Ceará, Brazil [44].

5.2.4 Molecular analysis

Molecular methods are also used for the diagnosis of plant diseases in the laboratory and they can be applied when genetic information is available. The molecular analysis allows to amplify and manipulate an original DNA fragment in order to increase the probability of detecting the pathogen causing the disease i.e. bacterias, fungus or viruses [43].

The techniques associated with this type of analysis, such as Polymerase Chain Reaction (PCR), Isothermal amplification and Microarray (Oligonucleotide array), enable a rapid and accurate detection and quantification [43]. However, the sample preparation is critical and requires complex protocols. All molecular methods for detecting plant pathogens are based on the accurate design of oligonucleotides and probes [45]. As an example of a molecular analysis, the Figure 5.9 shows the detection of different Nanoviruses by a broad primers set using PCR [46].

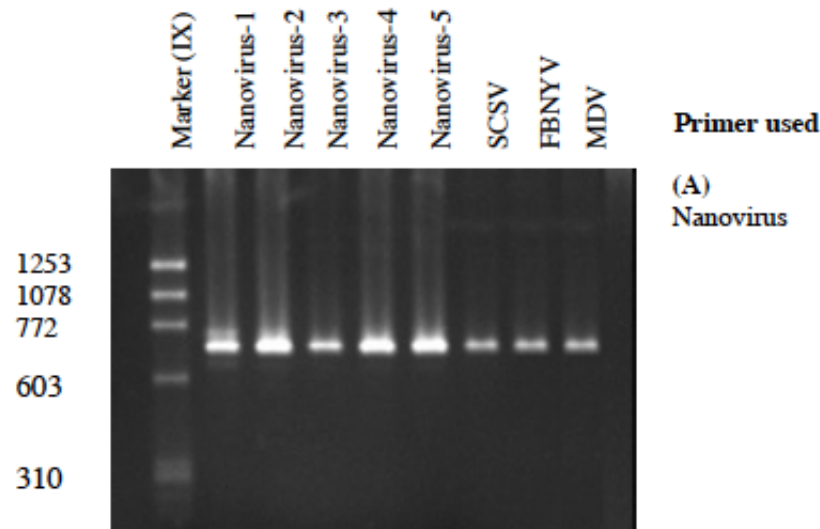


Fig. 5.9.: Detection of Nanoviruses by PCR [46].

5.2.5 Detection of visual symptoms

This method refers to the recognition of plant diseases using visual inspections, development scales and standard severity diagrams for their measurement [47]. People in charge of the crops, walk through them watching and touching the plants in order to identify symptoms associated with the particular disease that produces them and calculate infection levels [45].

The detection of visual symptoms uses the changes in the plant's appearance (colors, forms, lesions, spots, etc.) as an indicator of it being attacked by a disease or pest [45]. For instance, for the evaluation of the Coffee Leaf Rust's infection percentage in a specific lot, the number of diseased leaves in 60 random trees has to be divided into the total number of leaves in those trees and multiplied by 100 (see Equation 5.1). A leaf is considered diseased with Coffee Leaf Rust when chlorotic spots or orange dust are observed on it. As shown in Figure 5.10, a visual inspection can be carried out for detecting the presence of chlorotic spots on the leaves, which are then used for measuring incidence and severity of the disease [7].

$$\text{Average infection \% in the lot} = \frac{\text{Number of diseased leaves in the 60 trees}}{\text{Total number of leaves in the 60 trees}} * 100 \quad (5.1)$$



Fig. 5.10.: Coffee Leaf Rust detection through visual inspection [48].

5.2.6 Volatile Organic Compounds

Plants emit many Volatile Organic Compounds (VOCs) into their immediate surroundings that serve essential functions in growth, communication, defense and survival. Those compounds, emitted from leaf surfaces, are the result of the plant's metabolism and exist generally at ultra-low concentrations below the human olfactory threshold [45]. VOCs can serve as an indirect indicator of a plant's physiological health status because they are correlated not only with physicochemical factors such as humidity, temperature and fertilization, but also with biological ones such as growth stage and presence of pests and diseases [49].

VOCs of an infected plant are expected to vary from the ones released under normal health conditions. Thus, changes in the volatile profile caused by abiotic and biotic stresses can be monitored in order to detect a disease in real-time. The Gas Chromatography (GC) and Electronic Nose systems are two common techniques used for assessing the mentioned volatile profile [49].

For instance, the Figure 5.11 shows a portable Electronic Nose developed in the USA for detecting a pest, that damages cotton fields, called *Nezara viridula*. The device was able to capture and analyze VOCs from the head space of a cotton boll over carbon black-polymer composite sensors [50].



Fig. 5.11.: Field sampling cotton bolls with a pest and damage detector developed to draw VOCs [50].

5.2.7 Remote Sensing

Remote Sensing (RS) is based on the interaction of electromagnetic radiation with any material. In the case of agriculture, it involves the non-contact measurement of the reflected radiation from soil or plants in order to assess different attributes such as Leaf Area Index (LAI), chlorophyll content, water stress, weed density, crop nutrients, among others. Those measurements can be made using satellites, aircrafts, drones, tractors and hand-held sensors [51].

In addition to measuring reflected radiation, there are other two RS techniques that analyze fluorescent and thermal energy emitted by the leaves. However, the most common technique is reflectance, because the amount of reflected radiation from the plants is inversely related to the radiation absorbed by their pigments and this can serve as an indicator of their health status [51].

In specific wavelengths of the electromagnetic spectrum, the reflectance of plants varies. For instance, under normal conditions, plant pigments such as chlorophyll

strongly absorb radiation in the visible region (400 nm – 700 nm), whereas the plant reflectance is high in the near infrared (NIR) region (700 nm – 1,300 nm) [51]. These variations are due to the specific radiation that plants require for their photosynthetic activity [52]. RS helps the indirect detection of problems in agricultural fields, since this method captures unusual behaviors in crops reflectance, which can be caused by factors as nutritional deficiencies, pests and diseases and water stress.

As an example, the Figure 5.12 shows an UAV–RS system for the evaluation of weed infestation using a point-and-shoot camera and a six-band multispectral camera [45]. A multispectral camera is able to capture reflected radiation in multiple bands of the electromagnetic spectrum.



Fig. 5.12.: UAV equipped with RS devices for evaluation of weed infestation [45].

5.2.8 Spectroscopy

Spectroscopy is the study of matter using electromagnetic radiation. It deals with the production, measurement and interpretation of the spectra arising from the radiation-matter interaction. The spectroscopic methods differ with respect to the species to be

analyzed, the type of interaction to be monitored (such as reflectance, absorption or transmittance) and the region of the electromagnetic spectrum to be studied. In traditional food analysis laboratories, for example, the most common studies are related to reflectance or absorption of radiation in the ultraviolet (UV), visible (Red-Green-Blue, RGB), infrared (IR) and radio (Nuclear Magnetic Resonance, NMR) frequency ranges [53].

Regardless of the selected spectral band for the analysis, spectroscopy is based on the fact that the energy content of matter is quantized and that photons of radiation may be absorbed or reflected by matter [53]. Therefore, it is attempted to measure the amount of that radiation by an object of study using an optical unit, such as a spectrograph, in order to characterize its spectral signature and make an indirect assessment of its physicochemical features simultaneously. Spectroscopy has gained widespread recognition for product quality control and process monitoring because it combines simplicity, precision and a limited level of sample preparation. The Figure 5.13 shows the acquisition modes for spectroscopic analyses [54].

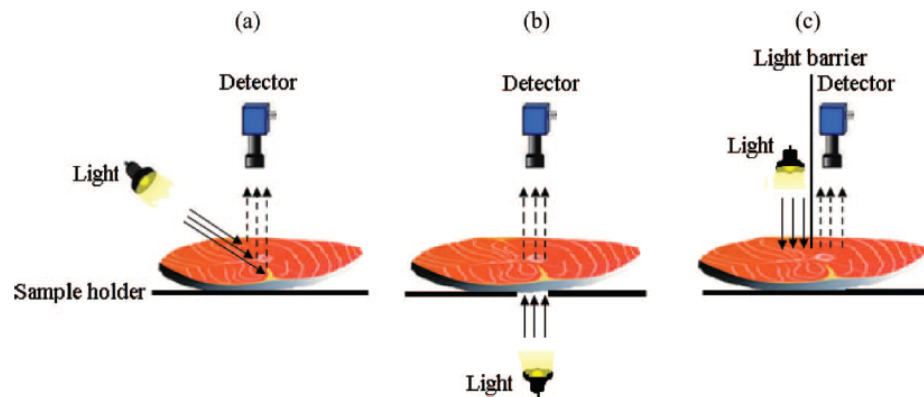


Fig. 5.13.: The three acquisition modes in spectroscopy: (a) Reflectance, (b) Transmittance and (c) Absorption [54].

5.2.9 Wireless Sensor Networks

Wireless Sensor Networks (WSN) is a technology that is being used in many countries worldwide in order to monitor different agricultural characteristics in real-time and remotely. It consists of multiple non-assisted embedded devices, called sensor nodes, that collect data in the field and communicate it wirelessly to a centralized processing station, which is known as the Base Station (BS). The BS has storage, processing and fusion capabilities and it is in charge of transmitting the received data to the Internet for presenting it to an end user [55].

Once the collected data is stored on a central server on the Internet, further analysis, processing and visualization techniques are applied to extract valuable information and hidden correlations that can help to detect changes in crop characteristics. These changes could be used as indicators of phytosanitary problems such as nutritional deficiencies, pests, diseases and water stress. WSN is a powerful technology since information of remote and inaccessible physical environments can be easily accessed through the Internet, with the help of the cooperative and constant monitoring of multiple sensors [56].

The sensor nodes in a WSN setup can vary in terms of their functions. Some of them can serve as simple data collectors that monitor a single physical phenomenon, while more powerful nodes may also perform more complex processing and aggregation operations. Even some sensors can have GPS modules that help them determine their particular location with high accuracy [56]. The most common sensors used in WSN for agriculture are the ones which collect climate data, images and frequencies. For instance, the Figure 5.14 shows a WSN configuration that was developed for a decision support system for water management in order to increase crop health and yield [57].



Fig. 5.14.: WSN in agriculture for water management [57].

5.2.10 Biosensors

A biosensor is a particular type of chemical sensor that uses the highly sensitive recognition properties of biological components, such as an enzyme, antibody, nucleic acid, microorganism or cell [58]. Biosensors are portable and they have been implemented for various applications including the early identification of crop diseases. This method is considered one of the most recent developments in the field [59].

Biosensors are self-contained devices, constituted by a biological and a physiochemical component, that are capable of providing information about the concentration of a target species, microorganism or virus within a test sample in real-time. The biological component has recognition receptors for determining when an interaction with the target element occurs (sensing). For its part, the physicochemical component is in charge of the transduction i.e. converting the chemical or physical variations in the biological component, due to the interaction with the target element, into a measurable physical quantity [60]. The Figure 5.15 is an example of a near-infrared

microscope reading the output of embedded carbon nanotube biosensors [61].

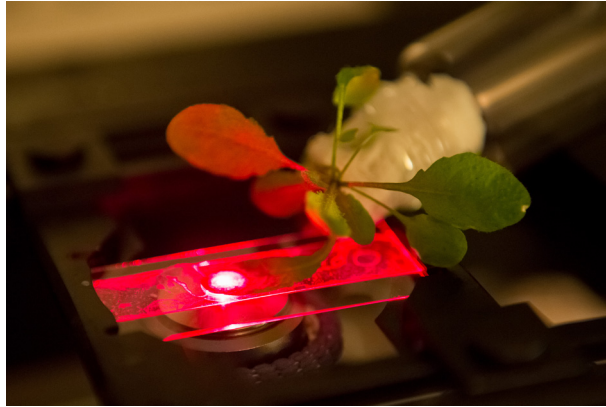


Fig. 5.15.: Analysis of carbon nanotube biosensors embedded in an *Arabidopsis thaliana* plant [61].

5.2.11 Computer Vision

Computer Vision (CV) encompasses the mathematical techniques for recovering the three-dimensional shape and appearance of objects in imagery. These techniques aim to describe the content of one or more images and to reconstruct its perceptual components as shape, illumination, color distribution, among others. Nowadays, CV is being used in multiple real-world applications such as reading handwritten postal codes, automatically recognizing number plates, rapidly inspecting mechanical parts for quality assurance, reconstructing 3D models from aerial photographs, detecting unexpected obstacles on the streets in the automotive safety field, etc. [62].

CV allows to transform data from a still or video camera into whether a decision or a new representation. An example of a decision might be “there are 14 tumor cells on this picture”, whereas a new representation might mean converting a color image into a gray-scale one. However, the kind of tasks that CV systems solve are difficult because they pretend to analyze and perceive situations as the human brain does,

using a noisy grid of numbers coming from the real world (a screen-shot) as the only input [63]. The Figure 5.16 shows an example of the object recognition system of an *Nvidia*[®] autopilot prototype called *Drive PX*. Through CV, the prototype recognizes and locates cautionary, stationary, moving and trivial objects on images that are captured while the car is in motion [64].

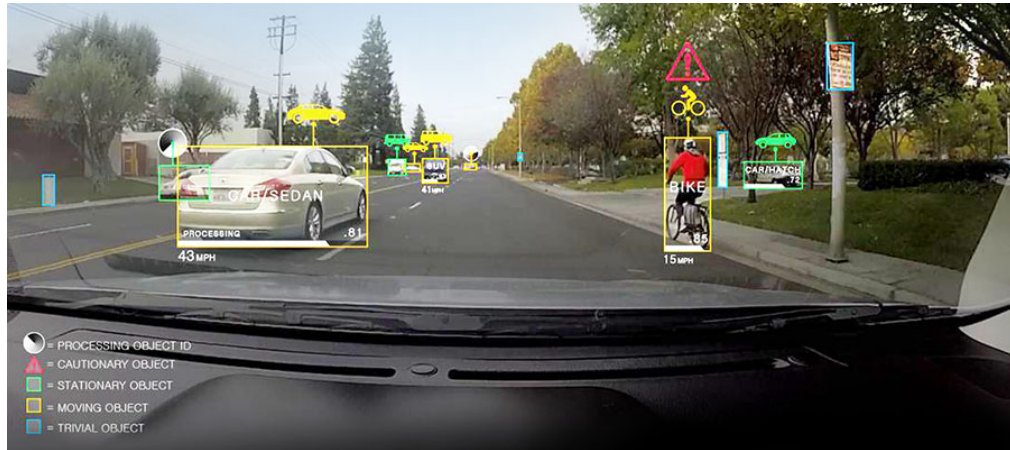


Fig. 5.16.: *Nvidia*[®] – *Drive PX* object recognition system [64].

5.2.12 Deep Learning

The area in charge of building intelligent machines that can perform certain tasks just like a human, is called Artificial Intelligence (AI) [65]. One of the main subareas of AI is Machine Learning (ML), which aims to extract complex patterns from large amounts of raw data automatically in order to predict future behaviors. When the extracting process of those patterns is taken to a more detailed level, where computers learn complicated real-world concepts by building them out of simpler ones in a hierarchical way, ML enters one of its most relevant subsets: Deep Learning (DL) [66].

The functionality of DL is an attempt to mimic the activity in layers of neurons in the human brain. The central structure that DL uses is called Artificial Neural Network

(ANN), which is composed by multiple layers of neurons and weighted connections between them. The neurons are excitable units that transform information whereas the connections are in charge of rescaling the output of one layer of neurons and transmitting it to the next one to serve as its input [67]. By passing data such as images, videos, sound and text through the ANN, DL builds hierarchical structures and levels of representation and abstraction that enable the identification of underlying patterns [65].

DL makes possible to learn a feature representation of the input data at multiple levels of abstraction with few or none human intervention and to discover the optimal parameters for building an accurate statistical predictive model. Using an image as an example of input data, the first layer would learn the first order features such as color and edges, the second layer would learn higher-order features such as corners, the third layer would learn about textures and so on. At the end of the ANN, the final layer generates a prediction either for a classification or regression task [65].

Thanks to the most recent improvements in mathematical formulas, increasingly powerful computers and the availability of large-scale datasets, DL is successfully applied in a wide range of fields nowadays, including Natural Language Processing (NLP), object visual recognition, medical diagnosis, trading, network security, among others [65]. For instance, the Figure 5.17 shows the process of detecting a face on an image through different levels of representation [68].

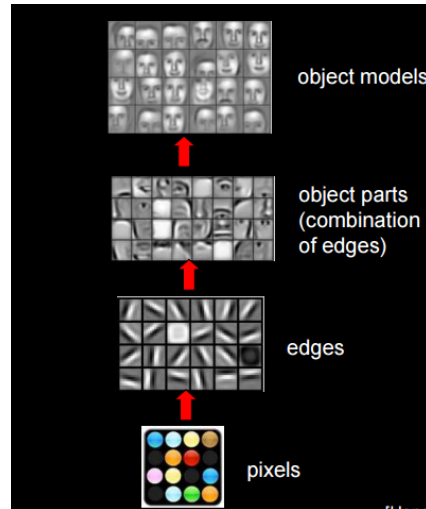


Fig. 5.17.: Face detection using DL [68].

For a deeper understanding about how DL works under the hood, the most relevant concepts as well as the popular standard models called Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are explained below. The Figure 5.18 shows the simplest ANN, which is called Multilayer Perceptron (MLP). It is essentially defined as having one input layer, one output layer and a few hidden layers (intermediate layers). The adjacent layers are fully connected and each connection has a specific weight, so that each artificial neuron can calculate a weighted sum of all its incoming inputs by multiplying the signals with the mentioned weights and adding a bias (offset). Then, the resulting value will go through a differentiable function called *activation function*, which makes the neuron decide if the output signal for the next layer should be fired, depending on whether a threshold have been reached.

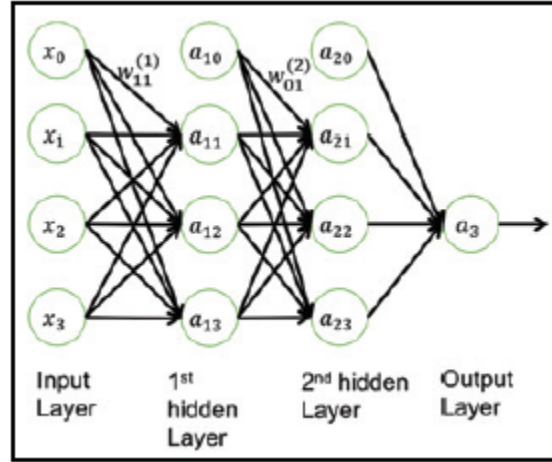


Fig. 5.18.: Fully connected, feed-forward neural network with two hidden layers [65].

The non-linearity nature of the models built with ANNs is mainly contributed by the shape of the activation function, which belongs usually to a *sigmoid*, hyperbolic tangent (*Tanh*) or Rectifier Linear Unit (*ReLU*). It is also relevant to mention that for classification problems in particular, the output layer has n neurons for n possible categories and applies a *softmax* function (see Equation 5.2) on the resulting values vector to output the probability of the input being each of them. In that sense, the higher probability indicates the model's prediction [65].

$$\sigma(\mathbf{z})_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad \text{for } j = 1, \dots, K \quad (5.2)$$

In addition, learning from the data in the context of DL means minimizing a cost function through a technique called *Gradient Descent*. The cost function takes the weights and biases as parameters and measures the total distance between the predictions of the model and the reality. For its part, *Gradient Descent* is the repetitive process of calculating the current gradient of the cost function, which is a vector in the direction of its steepest ascent, and taking a step in the opposite direction proportional to its magnitude until converging to a local minima. In that way and considering that the calculated cost covers the whole data, the outputs of the ANN

start nudging and moving from random to more accurate predictions [69].

The algorithm used in order to calculate the cost function's gradient efficiently is called *Backpropagation*. Its result indicates in which direction and how much every single parameter of the network needs to be adjusted in average, so that the total error decreases as much as possible [69]. An example of the output of a *Backpropagation* execution is shown in Figure 5.19. Each entry of the gradient corresponds to the required adjustment for each parameter of the network. In this example, the first parameter should be increased by 0.19 whereas the third parameter needs to be decreased by 0.85.

$$-\nabla C(\dots) = \begin{bmatrix} 0.19 \\ 0.86 \\ -0.85 \\ \vdots \\ -0.05 \\ 1.65 \\ 1.54 \end{bmatrix}$$

All weights
and biases

Fig. 5.19.: Example of a resulting cost function's gradient [69].

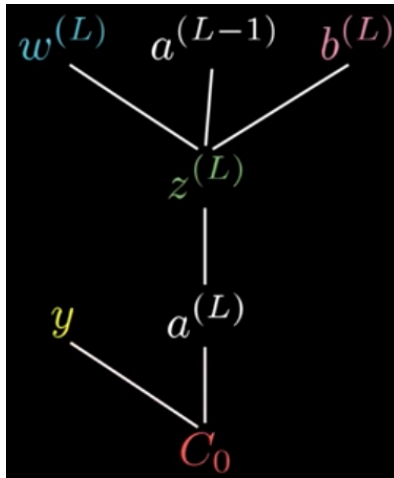
In essence, all the training does is to adjust the parameters (weights and biases) to best capture the patterns from data in order to successfully output the approximation of the target value. For this purpose, the training starts with the initialization of the parameters at random. Then, as shown in Figure 5.20, *Forward propagation* is executed, which consists in passing the input data through the network multiplying the signals (a_i) by the network's weights (w_i), adding the biases (b_i) and then going through the activation function (σ), from layer to layer until the output one is

reached [65].

$$\begin{aligned}
 z_1 &= a_0 * w_1 + b_1 \\
 a_1 &= \sigma(z_1) \\
 z_2 &= a_1 * w_2 + b_2 \\
 a_2 &= \sigma(z_2)
 \end{aligned}$$

Fig. 5.20.: Forward propagation operations on a two-hidden-layer network [65].

Once a prediction is made, *Backpropagation* takes action. First, every neuron on the output layer determines its contribution to the total error and the necessary adjustment according to the variation of the cost with respect to the variation of its output (through partial derivatives). Afterwards and using the determined adjustments, it is possible to calculate the corresponding corrections on the non-activated values (z_i), those that were used for the activation of the output layer's neurons (a_i), based on the variation ratio between a_i and z_i . Once the corrections are calculated, the necessary modifications on the parameters, which contribute to the non-activated values, are established with the help of those corrections and the variation ratio between the non-activated values and each parameter. Thereby, this procedure, known as the *Chain rule* and represented in Figure 5.21, generates the gradient's components that correspond to all parameters on the connections between the last hidden layer and the output one. Executing it recursively backwards through the network results in the necessary modifications for each remaining parameter of the model [69].



(a) Graphical representation.

$$\frac{\partial C_0}{\partial w^{(L)}} = \frac{\partial z^{(L)}}{\partial w^{(L)}} \frac{\partial a^{(L)}}{\partial z^{(L)}} \frac{\partial C_0}{\partial a^{(L)}}$$

(b) Formal representation.

Fig. 5.21.: The *Chain rule* [69].

The last step of the training process is where the network actually learns from the data. As shown in Figure 5.22, it consists in updating the network's weights and biases according to the calculated gradient and using a hyperparameter called the *learning rate* (η), which controls in which percentage the parameters are adjusted. The lower the value, the slower and more controlled the steps towards the local minima [65].

$$w_1 \leftarrow w_1 - \eta \cdot \nabla w_1$$

$$b_1 \leftarrow b_1 - \eta \cdot \nabla b_1$$

$$w_2 \leftarrow w_2 - \eta \cdot \nabla w_2$$

$$b_2 \leftarrow b_2 - \eta \cdot \nabla b_2$$

Fig. 5.22.: Parameters update operations on a two-hidden-layer network [65].

Additionally, it is relevant to mention two popular standard models, namely CNNs and RNNs, which are variants of MLP for specific tasks. On the one hand, CNNs

have been proven to be very effective in areas such as image recognition and classification. They have been successfully applied for identifying things on images thanks to their capacity to exploit spatially-local correlation by enforcing a local connectivity pattern (aka weights sharing) between neurons of adjacent layers that have spatially contiguous receptive fields. The main difference with standard MLP is the presence of layers that execute convolutions. A convolution is the process of taking an image and sliding over it block-wise while applying filters (detectors of different features), in order to aggregate the original input and transform it into a more compact but deeper representation [65].

On the other hand, RNNs are designed specifically for handling inputs which come in sequence, such as sentences, genomes, handwriting, speech, data from sensors and stock markets, among others, and predicting the next value in line or the subject, which often greatly depends on the past context and its order. In that sense, a RNN can be thought of as a chain of repeating cells in the dimension of time, in which each time step can be treated as a different layer, meaning that any cell of the network (aka state) is influenced not only by the current input but also by its recent past [65].

5.2.13 Web programming framework

A web programming framework is a well-defined structure of reusable, interchangeable and customizable software components for the development of a web system. It can be understood as an incomplete generic application, which serves as a solid base and is configurable and extensible for building a concrete and unique application. The components of a web programming framework are designed to facilitate and speed up the organization and implementation of systems than can be accessed through the Internet [70].

The use of a programming framework for web applications is intended to accelerate the software development process, reuse existing code, promote the inclusion of architectural patterns and facilitate maintenance. In general, every web programming framework contains components for URL abstraction, user's session management, data access, authentication/authorization, internationalization and separation between design and content [70].

The web programming frameworks are organized according to the programming language they support. The most common ones are Django, Flask and Tornado for *Python*, Angular.js, React.js and Ember.js for *JavaScript*, and Rails and Sinatra for *Ruby*. The Figure 5.23 shows an architectural pattern called Model-View-Controller (MVC), on which Ruby on Rails and Django are based [71].

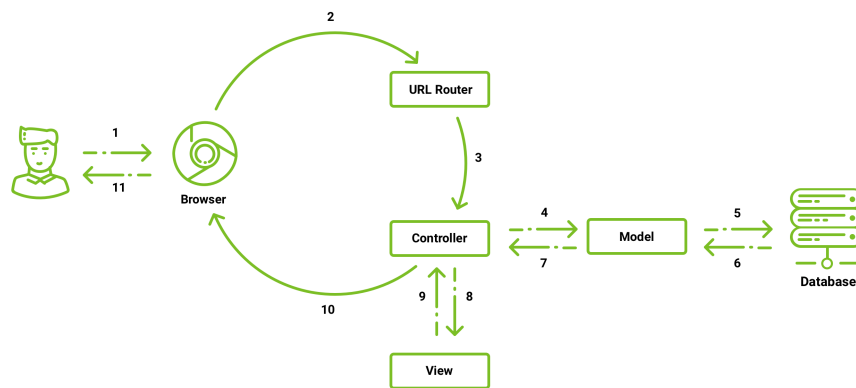


Fig. 5.23.: MVC architectural pattern [71].

6. STATE OF THE ART

Plenty of research has been done around applying technological methods and strategies to diagnose diseases [72], detect pests [73], obtain nutritional information [74], among other objectives, for different types of crops. The phytosanitary status of the plantations is closely related to different crucial factors in their ecosystem such as weather, altitude, type of soil, etc. Therefore, several biological and engineering studies aim to implement practical solutions based on these factors to improve farming techniques to preserve healthy crops.

The most commonly used methods for monitoring the phytosanitary status efficiently, including those which make use of technology, are *Visual Detection*, *Biological intervention*, *RS*, *WSN* and *ML*. Thus, this chapter is intended to present recent relevant studies based on the mentioned methods for detecting anomalies on the plantations. Research towards the diagnosis of the Coffee Leaf Rust, which is the target disease of this Master thesis, are also exposed.

6.1 Remote Sensing

In 2003, Goel et al [75] studied the detection of changes in the spectral response in corn (*Zea mays*) due to nitrogen application rates and weed control. For that purpose, the researchers employed a hyper-spectral sensor called Compact Airborne Spectrographic Imager (CASI) and analyzed the reflectance values of 72 bands with a wavelength between 409 and 947 nm, which comprise part of the visible and near-infrared (NIR) regions of the electromagnetic spectrum. The obtained results demonstrated the potential of detecting weed infestations and nitrogen stress using the hyper-spectral sensor CASI. Specifically, the researchers found that the best fitting bands for the

detection were the wavelength regions near 498 nm and 671 nm respectively, as seen in Figure 6.1.

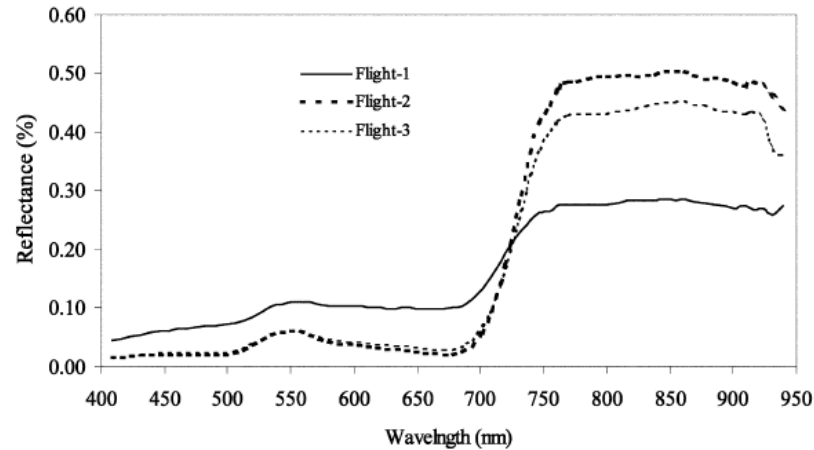


Fig. 6.1.: Reflectance (%) of the corn response during different flights under normal nitrogen rates and no weed control [75].

Regarding the phytosanitary status of the plants, the water and the type of soil are two components that play an important role in their health. In 2017, Bolaños et al. [76] proposed a characterization method using visible and infrared spectrum to identify these components, through low-cost cameras with two different filters Roscolux #19 and Roscolux #2007, and a multi-rotor air vehicle. Through this method, and using portable and highly qualified devices, those hard-to-reach places were monitored and analyzed in order to detect anomalies that may cause diseases in the crop. This monitored phase provided a characterization of the Normalized Difference Vegetation Index (NDVI) as seen in Figure 6.2(a). As a result of this phase, a color based chart determined three essential characteristics such as crop health, diseased plants or soil, and water or others, as seen in Figure 6.2(b).

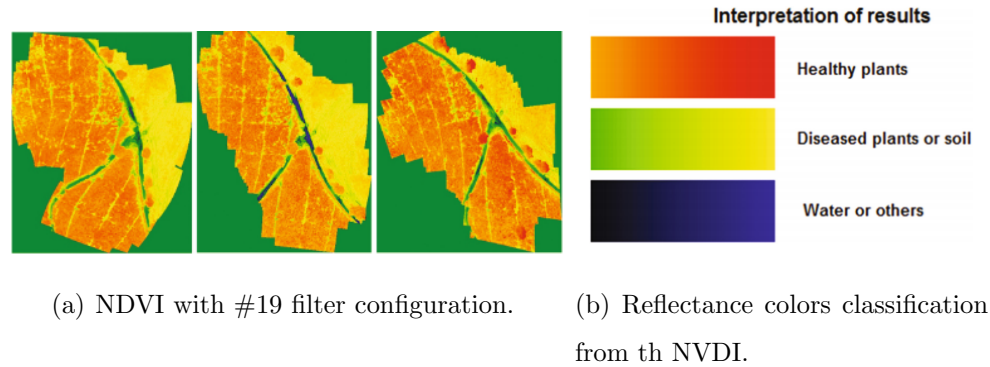


Fig. 6.2.: Characterisation of the NDVI with low cost solutions [76].

In 2017, Chemura et al. proposed a method to early predict the presence of diseases and pests among coffee trees based on unnoticeable water stress. For that purpose, multi-spectral scanners with wavebands from the visual spectrum and near the infrared region filters were placed on an Unmanned Aerial Vehicle (UAV) [77]. The wavebands scanner results showed inflections points between the regions 430nm, and 705–735 nm near infrared due to the water content in coffee trees, see Figure 6.3. These results underlined the importance of a suitable irrigation plan according to the water requirements of the trees, causing an improvement in productivity sole. Although the latest region indicated relevant values, the waveband of 430 nm was the most relevant band of remote sensing for predicting the water plant content directly related to the stress of it. However, in [77] the authors remark that although the results were promising there were missing some valid components that could allow the model to be totally suitable and testable in real conditions. For that purpose, they recommended using hyperspectral cameras which provide a more precise waveband measured results.

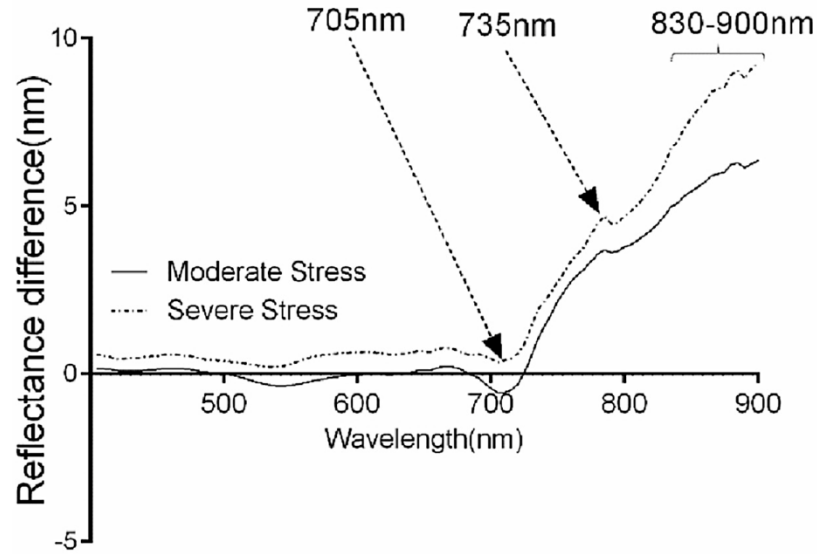


Fig. 6.3.: Water reflectance content difference in the plant [77].

6.2 Visual Detection

To understand the conditions conducive to the development of Coffee Leaf Rust and subsequently, refine the disease control, Avelino et al. [78] monitored such development on 73 coffee crops in Honduras for 1-3 years. Thereby, through the analysis of production situation variables such as climate, soil components, coffee tree characteristics, and crop management patterns, the researchers aimed to seek to establish a relationship with the rust presence. The result of this research indicated that Coffee Leaf Rust epidemics depend on the diverse production situations based on Table 6.1, linked as well to the local conditions of the plantation. For this reason, these results reflect the need of consideration of a certified growing system that aims for sustainability, taking into consideration production situations and, thus, preventing developing of pests and diseases.

Table 6.1.: Kinds of variables that describe the importance of Coffee plots in the presence of Coffee Leaf Rust [78]

Kind of Variable	Relevance
Climate variation (Alt. and Rainfall)	High
Soil components	Medium - low
Croppping practices	Medium
Coffee tree productivities	High

6.3 Biological Intervention

Several authors stated the importance of the relationship between living beings sharing the same environment. One of them is Haddad et al. [79], who in 2009, proposed a study to determine if seven selected isolated bacterias under greenhouse conditions would efficiently detect and control CLR. For the development of this research, they inoculated these bacterias: (six *Bacillus* sp. -B10, B25, B157, B175, B205, and B281 and one of *Pseudomonas* sp.-P286), which help to detect and control Coffee Leaf Rust in early development stages, according to a preliminary result presented by Haddad et al. (2007). For the experiment, two important coffee varieties, (*Mundo Novo* and *Catuai*) were selected due to the high susceptibility of the Coffee Leaf Rust. Therefore, for three years the varieties with the disease interacted with different treatments (bacteria) to analyze the behavior evolution between them. Based on the treatments results, the isolates P286 and B157 were as efficient as the copper fungicide in controlling the rust. Hence, considering the harmful effects due to the copper fungicide, the application of biological control with the B157 isolate of *Bacillus* sp. may be a re-

liable alternative solution to CLR management. That is why, this research displayed the opportunity to successfully biocontrol the Coffee Leaf Rust, for specialty coffee growers.

Jackson et al. [80] in 2012, proposed as well, a biological detection and control based on a fungus, *Lecanicillium lecanii*. Their primary interest of crops, in general, is the analogy of the coexistence of organisms in a specific environment with defined conditions that encounters the perfect balance. In view of the above, the biological control system of the (*A. instabilis*) ants are mutualistically be associated the while halos of the fungus, see Figure 6.4 (*L. lecanii*) based on the Coffee Leaf Rust effect.

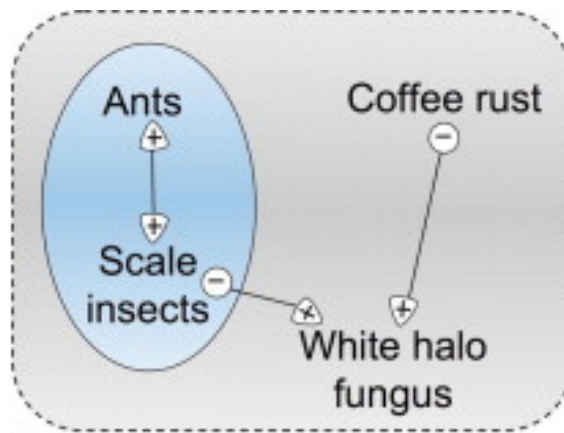


Fig. 6.4.: Diagram of co-existing organisms [80].

However, the hypothesis stated the possibility that spores from the *L. lecanii* will help to attack the *H. Vastatrix* before the rainy season. The research results demonstrated that the effect of the time delay of the *L. lecanii* in *H. Vastatrix* caused that a relationship between the two fungus and the ants could not be demonstrated, in spite of the resembling of the control experiment to the real world. In conclusion, the restriction of a biotic-factors directly affect the development of the Coffee Leaf

Rust, therefore, for future work, is important to consider the climate variation of an ecosystem to be able to predict such development [80].

6.4 Wireless Sensor Networks

Chaudhary et al. [81] emphasized in 2011 the importance of WSN in the field of PA by monitoring and controlling different critical parameters on a greenhouse through a microcontroller technology called Programmable System on Chip (PSoC). In consequence of the disproportionate rainfall dynamics, thus the need to control a suitable water distribution meeting those parameters inside the greenhouse arises. Thereby, the study tested the integration of wireless sensor nodes structured as seen in Figure 6.5, with high-bandwidth spectrum telecommunication technology. Mainly, it was proven that the integration was useful to determine an ideal irrigation plan that met the specific needs of a crop based on the nodes interaction within the greenhouse. Furthermore, the researchers recommended using reliable hardware with low current consumption for developing WSN projects, because it generates more confidence to the farmers concerning its incorporation on their crops.

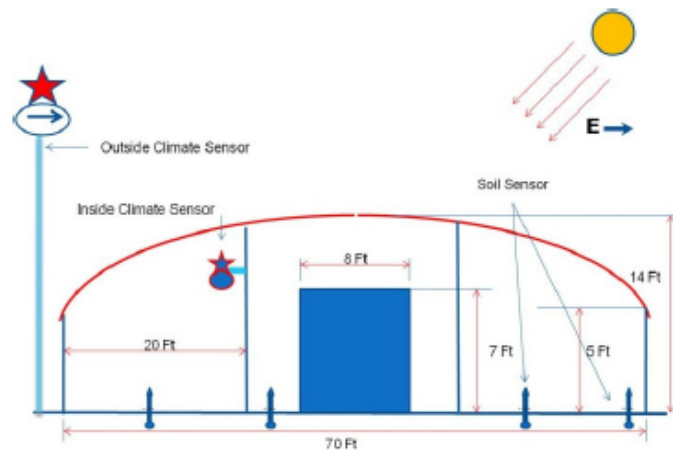


Fig. 6.5.: Structure and sensors locations of the greenhouse [81].

In addition, Piamonte et al. proposed in 2017 a WSN prototype for monitoring the Bud Rot of the African Oil Palm. With the use of pH, humidity, temperature and luminosity sensors, they aimed to measure climate variations and edaphic (related to the soil) factors to indirectly detect the presence of the fungus that causes the disease. Through Figure 6.6 it illustrates the connection scheme (two nodes) that conducted the data acquisition of edafoclimatic variables with (Node 1) and without (Node 2) the presence of bud rot. The research results demonstrated that the measurements of the mentioned abiotic factors presented slightly variations, which, according to the researchers, supplied indications about the possibility of detecting the Bud Rot of the African Oil Palm [82]. Thus, monitoring the triggering variables associated with this disease would minimize its impact on the African Oil Palm trees.

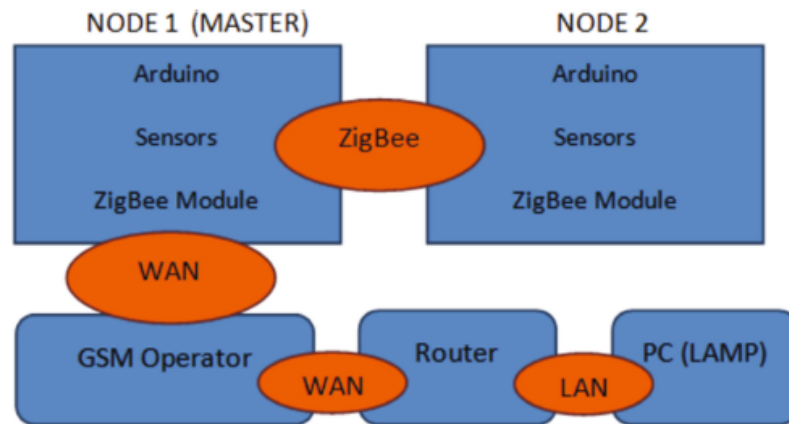


Fig. 6.6.: Connection scheme of the WSN nodes [82].

6.5 Machine Learning

Chemura et al. in 2017 [83], evaluated the potential of Sentinel-2 band in order to early detect the CLR infection levels due to its devastating rates. Through the employment of random forest (RF) and partial least squares discriminant analysis (PLS-DA) algorithms such levels could be identified for and early CLR management. The researchers employed the variety *Yellow Catuai*, which was chosen due to its

CLR susceptibility. In this matter, Chemura et al. considered only seven Sentinel-2 MSI band for its high resolution stated by previous works in biological studies. Based on the selected bands, the research results determined that the CLR reflectance was higher in NIR regions of the spectrum as can be seen in leaves from the band B4 (665 nm), B5 (705 nm) and B6 (740 nm) in Figure 6.7. These bands achieve a high overall CLR discrimination of 28.5 % and 71.4% using the RF and PLS-DA algorithms. Thus, the band and vegetation indices derived from MSI of Sentinel-2 achieved a detection of the disease and an evaluation of the CLR in early stages, avoiding unnecessary chemical protection in healthy trees.

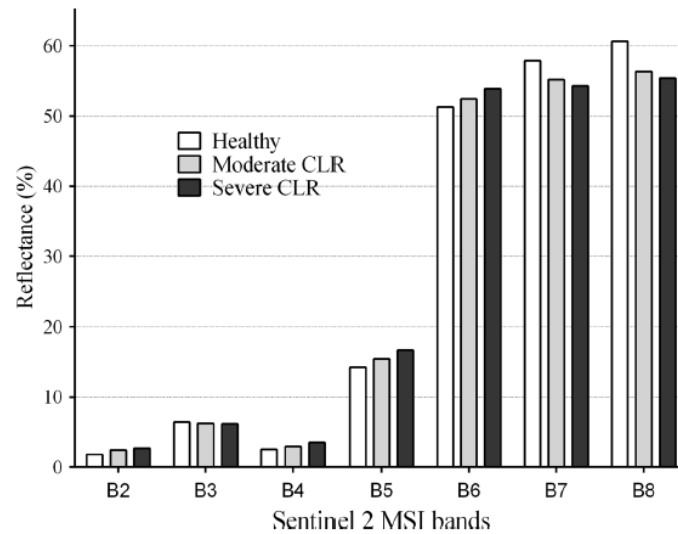


Fig. 6.7.: Mean spectral reflectance of Coffee Leaf Rust infection levels with Sentinel-2 bands [83].

As well in 2017 Chemura et al [84] studied the detection of Coffee Leaf Rust through the reflectance of the leaves' at specific electromagnetic wavelengths. The objective of their investigation was to assess the utility of the wavebands used by the Sentinel-2 MultiSpectral Imager in detection models, which were created using partial least squared regression (PLSR) and the non-linear radial basis function-partial least squares regression (RBF-PLS) machine learning algorithm. Then, both models were compared, resulting in a low accuracy prediction of the state of the disease

for the PLSR, due to its over-fitting, and a high accuracy prediction for the RBS-PLS model. Additionally, Chemura et al, through weighting the importance of the variables, found that the Blue, Red and RE1 bands had a high model correlation, as shown in Figure 6.8, but its implementation excluding the remaining four bands led to a lower accuracy in both models. On the other hand, if more than one NIR or RE band were used, the RBS-PLS model developed would over-fit, resulting in a non-transferable model. However, Chemura et al emphasize the utilization of the RBS-PLS model due to its machine learning advantage and its good adaptation to possible model over-fitting.

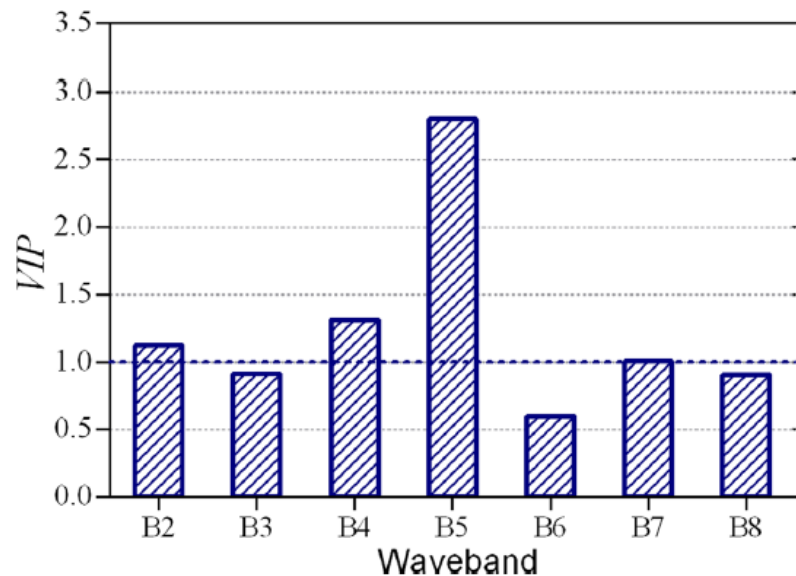


Fig. 6.8.: Correlation coefficients and significance between the spectroradiometer wave bands (B2 to B8) and coffee leaf rust severity on individual leaves [84].

6.6 Commercial technologies

In addition to the scientific advances regarding the anomalies identification techniques for disease detection in crops, some companies have also made commercial developments. These agricultural techniques assess variables such as vegetation indexes, stress by drought, nitrogen levels [85, 86] among others. Thereby, several patents

proposed diverse disease detection models based on volatile matter and light information [87, 88]. Additionally, a noninvasive and remote detection method seeks to find a substance presence based on its location and quantity via a spectral imaging system. This method comprises a spectral filter and an image capture array, which results in an identification of the symptoms of specific diseases [89].

Companies like *Parrot*, *Agrolevels*, and *Agrocloud* have addressed the detection issue as an opportunity for developing technologies for better agricultural practices through multi-spectral cameras as seen Figures 6.9 and 6.10.

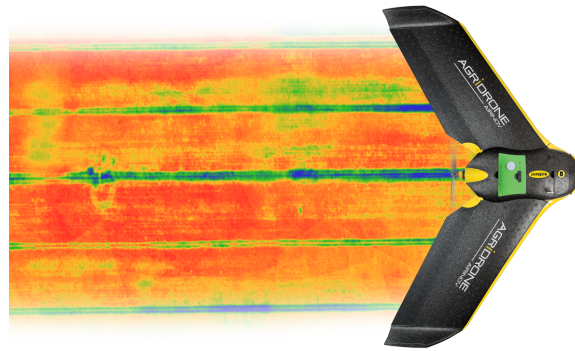


Fig. 6.9.: UAV with a parrot sequoia cameras and the results among crops [90].

An example of these cameras, are the Sequoia. This commercial components provides status of the objective crop based on the reflection bands. In this matter, the symptoms of possible diseases can be detected and treated before any repercussion affect the crop [90, 91]. For this reason, the quality of life and productivity of the analyzed crop is improved.



Fig. 6.10.: Agrocloud app for detection of disease among African Oil Palm [92].

The presented state of the art shows that several researchers and companies seek the detection of any vital element like water stress, nitrogen levels, and vegetation indexes that can lead to an improvement of production and quality in crops, which translates to an increase in profitability. In the following chapters, it will be explained how the best practices evidenced in this chapter are integrated to build a cyber-physical prototype that allows to address the research question of this Master thesis.

7. DESIGN OF EXPERIMENT

This chapter explains the design of experiment that is implemented in this Master thesis. It details how it is planned to obtain data for the analysis and presents the procedure that is used to answer the research question based on the experimental results. A suited design of experiment is fundamental because it allows estimating to what extent a factor influences the change of a response variable and helps, at the same time, by providing statistical support to the conclusions.

7.1 Completely Randomized Design

The design of experiment that is implemented in this Master thesis is called Completely Randomized Design (CRD). It is used to compare two or more treatments considering only two sources of variability: treatments and random error. The objective by using this design of experiment in this project is to analyze whether the diagnosis of the Coffee Leaf Rust development stage through the integration of RS, WSN and DL is similar to the one made with a traditional visual inspection. In that sense, the study factor is the type of inspection, which has two levels (“visual inspection” and “technological integration”), and the response variable is the development stage of the disease, which is a whole number between 0 and 4. Thereby, the fundamental hypothesis to prove, presented in Equation 7.1, helps by deciding whether the treatments **1** (“visual inspection”) and **2** (“technological integration”) are statistically equivalent with respect to their means [93].

$$\begin{aligned} H_0 : \mu_1 &= \mu_2 \\ H_A : \mu_1 &\neq \mu_2 \end{aligned} \tag{7.1}$$

The procedure for proving the mentioned hypothesis is called Analysis of Variance (ANOVA) and requires a data table containing a row for each observation and a column for each treatment indicating the measurements of the response variable. This procedure separates the variability due to the treatments from the one attributed to the random error and compares them. If the former is clearly greater than the latter, the means of the treatments are different and thus, the type of diagnosis influences the determined Coffee Leaf Rust development stage. Otherwise, the means are statistically equivalent and it is possible to conclude that the visual inspection and the technological integration are similar for diagnosing the disease. Lastly, it is important to mention that the significance level that is used for proving the hypothesis is 10 % ($\alpha = 0.1$) since the problem at hand is related to Agro, where many noise factors associated with the variation of environmental conditions are involved [93].

7.2 Data collection

For the data collection experiment, 16 six-months old, healthy coffee plants coming from Jardín, Antioquia will be used. Those plants will be stored in a *Universidad EAFIT*'s greenhouse and a Biology team will be in charge of their transplant, agronomic management (elimination of weeds, fertilization and fumigation), inoculation and supervision. For the inoculation, the Biology team will follow the process described in *Chemura et al.* [83]. It is relevant to clarify that an extra group of diseased plants will be held as a reserve in case the inoculation of the healthy plants does not take effect over time.

Furthermore, an Engineering team will be dedicated to the design and assembly of a system, in the same greenhouse, that integrates RS and WSN. It will allow building a scale crop, recording different characteristics of it on a regular basis and storing them on a remote server in order to analyze its phytosanitary status later using DL. In that way, once the plants are inoculated and the system is verified, they will be

transplanted to it so that the data collection may begin. For that purpose, the scale crop is going to be divided into four lots with certain differences in their agronomic management, which seeks to recreate various circumstances of a real coffee crop. Thereby, a greater number of scenarios will be covered and the false positive rate regarding the diagnosis will be reduced. The **lot 1** will contain four non-inoculated plants and they will be neither fertilized nor fumigated; the **lot 2** will have four non-inoculated plants and they will be fertilized but not fumigated; the **lot 3** will have four inoculated plants and they will also be fertilized but not fumigated, and the **lot 4** will have four inoculated plants and they will be neither fertilized nor fumigated.

Finally, the visual inspections for diagnosis the Coffee Leaf Rust development stage will be carried out by the Biology team for a period of three months. Once per day, one of them will examine the severity of the disease on each lot and indicate thereby the value of the response variable for each observation. For its part, the technological system will automatically record the scale crop's characteristics from each lot seven times per day at different moments (with and without sunlight, because field sensors and cameras have different illuminance requirements). After the data collection phase finishes, the diagnostic model using DL will be generated and a comparative data table for the statistical analysis will be produced, based on its predictions and the results of the visual inspections. As it is expected that a considerable amount of observations will be made, only 25 % of all collected data will be used for the statistical study. It should also be noted that, as it is recommended, the order of the table's entries will be randomized before executing the analysis.

8. CONCEPTUAL APPROACH

In order to evaluate to what extent it is possible to diagnose the Coffee Leaf Rust development stage in the Colombian *Caturra* variety through the integration of RS, WSN and DL, it is necessary to obtain empirical evidence by means of an experiment. Therefore, there is a need to build a prototype that includes a scale coffee crop which simulates different agronomic conditions and allows to capture data for diagnosing the disease.

For the construction of the mentioned prototype, the *Pahl & Beitz* methodology was followed. The results of the first three phases, which comprise everything up to the embodiment design, are presented in this chapter.

8.1 Product Design Specification

The Product Design Specification (PDS) shows the main crop, manufacturing, electronic and software requirements that should be considered for building the appropriate prototype. They were gathered from the interviews with coffee and agronomic experts and conversations with *CENICAFÉ*, the Biology team, *Universidad EAFIT*, the Heads of the Physical and Systems Engineering undergraduate programs at *Universidad EAFIT*, data processing experts and the Design team.

The PDS can be seen in Appendix A. The rows represent the categories to which the requirements belong, namely crop, manufacture, electronics and software. The crop category specifies the spatial and biological conditions for the growth of the coffee plants, the manufacture category contains materials and assembly constraints for building the prototype, the electronics category exhibits the devices to be used

for collecting data as well as their powering, safety and communication conditions, and the software category presents the requisites for collecting, storing and processing data in order to diagnose the disease and evaluate the results. For its part, the columns indicate the needs expressed by the mentioned stakeholders, the specific requirements to be fulfilled arising from those needs, their source and how to measure them (with units of measurement and range of values), their priority to be met and the way to validate/verify their compliance.

Moreover, the main requirements for each category are presented with the purpose of explaining how they are formulated. With respect to the crop requirements and according to the interviewees, it is necessary to store coffee plants from the *Caturra* variety, which can be guaranteed by acquiring them through a certified *Caturra* provider. Regarding the manufacture, the Design team determines that the materials need to be resistant to environmental conditions (temperature, humidity, soil, etc.), whose compliance can be verified checking immunization and testing the performance during the construction of the prototype. For its part, an important electronic requirement for *CENICAFÉ* and the Design team is to detect the disease by physicochemical changes in the plants, capturing RGB images per tree from the bottom. The validation of this requirement can be done with the camera's specifications and its corresponding mechanism of action. Lastly, the Design team and the data processing experts see it as necessary to implement more than five alternatives for the generation of the DL model, whose compliance can be verified through the comparison of each diagnosis.

8.2 Black box

The PDS serves as the input for creating the Black box towards the generation of a concept that covers, materializes and meets the main requirements of the stakeholders. The Black box presents the principal function that the prototype must fulfill as well as the input it receives in order to produce the desired output. The input and

output are divided into three major flows: matter, energy and signal.

As shown in Figure 8.1, the principal function is to diagnose the Coffee Leaf Rust development stage because this is the means to evaluate to what extent it is possible to equate the results to those of a traditional visual inspection. For that purpose, Coffee Leaf Rust, coffee plants, organic matter, fertilizer/fungicide, wind, water, electrical energy, human force, photovoltaic energy, agronomic management information and the results of the Biology team's visual inspections regarding the Coffee Leaf Rust development stage are used as inputs. After passing through the principal function, the adequate experimental coffee crop, dissipated energy, field sensor and general data records as well as diagnostic information of the disease are obtained.

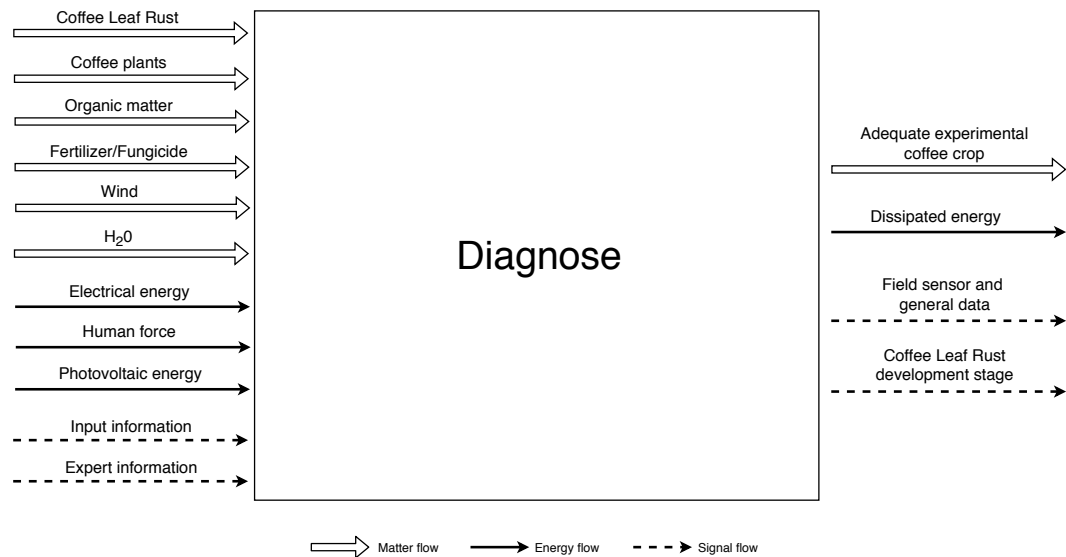


Fig. 8.1.: Black box of the prototype.

8.3 Functional structure

The Functional structure specifies which sub-functions are there and how they cooperate to achieve the prototype's principal function. In addition, it breaks down the

inputs and outputs presented in the Black box for a more detailed understanding.

The Functional structure can be seen in Appendix B. The inputs are Coffee Leaf Rust, coffee plants, organic matter, fertilizer and fungicide, wind, water, electrical energy, human force, photovoltaic energy, input information about the agronomic management and expert information about the development stage of the disease. The Functional structure is divided into two blocks, which correspond to the data collection and its further processing.

On the one hand, the coffee plants are grouped, combined with the organic matter and then divided into four lots. In addition, they are separated for inoculating the Coffee Leaf Rust in half of them and after that, the four lots are assembled again. For their agronomic management, fertilizer and fungicide are distributed and incorporated. Then, each lot is isolated from the others to make the four lots independent and the whole scale crop is combined with a rain emulation system and a wind system. Both rainfall and wind speed for the whole crop are perceived. Furthermore, using sensors in each lot, pH, illuminance, temperature, humidity and electrical conductivity are perceived, and RGB and multi-spectral images are captured. The state of each lot is integrated with the expert information to diagnose the Coffee Leaf Rust development stage and then, this information is clustered with the whole collected data. To finish the data collection process, the data is stored locally and sent to a remote server over the Internet.

On the other hand, the data that is received on the remote server is preprocessed for cleaning purposes and stored in a remote database. In order to create an adequate model for diagnosing the Coffee Leaf Rust development stage, the stored data is first divided into three sets, namely training, validation and test. The training set is processed to build the diagnostic model whereas the validation set serves to assess its intermediate performance and tune it. Once the diagnostic model is generated,

the test set is used for evaluating its final performance and a simple visualization is created to illustrate the results. The outputs of the Functional structure are the adequate experimental coffee crop and the diagnosed Coffee Leaf Rust development stage.

8.4 Morphological matrix

In the Morphological matrix, the solution proposals for implementing the different prototype's sub-functions, exposed on the Functional structure, are presented. Thereby, two candidate concepts are elaborated with the objective to evaluate both of them under different criteria and decide which is the most appropriate for the problematic at hand. Each concept is made up of a combination of the solution proposals, which is represented by a red and a blue path on the Morphological matrix. The red path indicates the components of the *Concept 1* whereas the blue path shows how the *Concept 2* was conceived. The Morphological matrix can be seen in Appendix C.

8.5 Evaluation of concepts

After defining the two candidate concepts for building the prototype through the Morphological matrix, their advantages and disadvantages are analyzed. Furthermore, possible solutions to the encountered disadvantages are also proposed. In addition, both concepts are evaluated regarding fifteen different criteria in order to select how the prototype will be constructed. The Appendix D summarizes the analysis of both concepts and the Table 8.1 presents their scoring. It is noteworthy that the numbers inside the parentheses on Appendix D correspond to the numbers of the sub-functions in the Morphological matrix.

Table 8.1.: Concepts scoring.

Value scale (score between 0 - 4)				
0 = Not satisfied, 1 = Acceptable, 2 = Sufficient, 3 = Good, 4 = Totally satisfied				
N°	Evaluation criteria	Relevance	Solutions	
			Concept 1	Concept 2
1	Functionality	11 %	4	3
2	Simplicity	5 %	3	4
3	Fulfillment of requirements	10 %	3	2
4	Robustness	3 %	4	3
5	Fabrication	7 %	3	3
6	Assembly	6 %	3	2
7	Reliability	9 %	3	3
8	Low cost	7 %	3	3
9	Expert criteria	6 %	3	3
10	Crop management	7 %	3	3
11	Maintainability	3 %	2	3
12	Performance	8 %	2	3
13	Usability	5 %	3	3
14	Testability	3 %	3	2
15	Availability	10 %	4	4
Weighted average			3.13	2.96
Total score		100 %	78 %	74 %

The fifteen evaluation criteria that are used for scoring both concepts are explained below.

- **Functionality:** how well the concept fulfills the principal function.

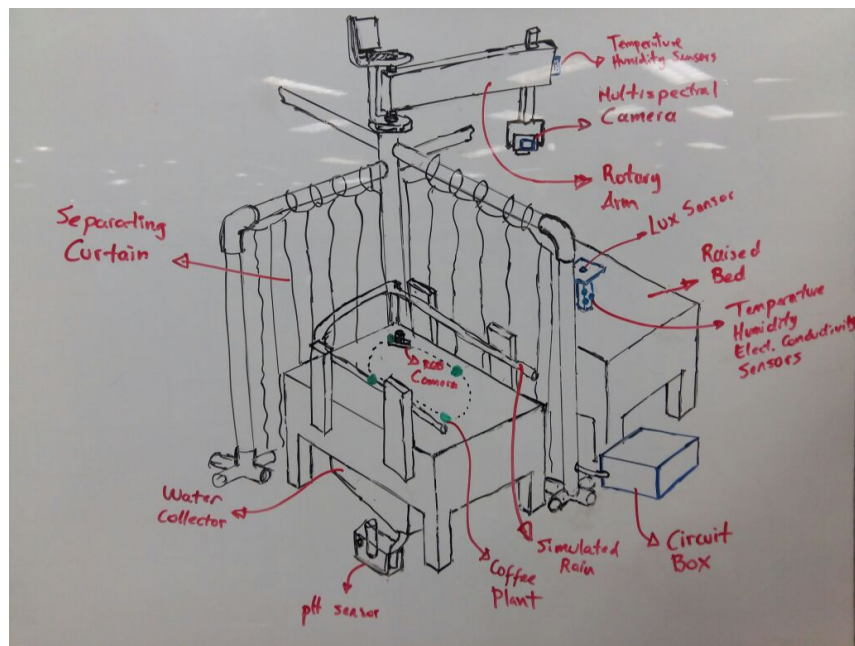
- **Simplicity:** how simple it is to elaborate and implement the concept.
- **Fulfillment of requirements:** to what extent the concept fulfills the requirements.
- **Robustness:** how well the concept behaves when failures occur.
- **Fabrication:** how simple it is to buy, build or develop the components of the concept.
- **Assembly:** how easy it is to integrate the components of the concept.
- **Reliability:** to what extent it is possible to rely on the operation, measurements and processing of the concept.
- **Low cost:** how cheap it is to construct and implement the concept.
- **Expert criteria:** what the biology, agronomy, design and research experts think about the concept.
- **Crop management:** how adequate the conditions are for the normal growth of the plants using the concept.
- **Maintainability:** how simple it is to maintain the concept and correct its possible failures.
- **Performance:** how satisfactory the concept is regarding time behavior, resource usage and capacity.
- **Usability:** how easy it is to operate and access the concept and its results.
- **Testability:** to what extent it is possible to test the different components of the concept independently.
- **Availability:** how available the concept is for data collection, processing and visualization in comparison with the total working time.

Finally, the result of analyzing and scoring the two candidate concepts determines that the winning one is the *Concept 1* (with 78 % against 74 %). It should also be noted that the encountered disadvantages and experts' comments and suggestions are taken into account to refine the concept and improve its functionality. In that sense, the Figure 8.2 shows a sketch of the refined final concept for building the prototype that will allow to evaluate the diagnosis of the Coffee Leaf Rust development stage in the Colombian *Caturra* variety by integrating RS, WSN and DL. The sketch is divided into two sections, namely the physical part (Figure 8.2(a)) and the informatic or cybernetic part (Figure 8.2(b)).

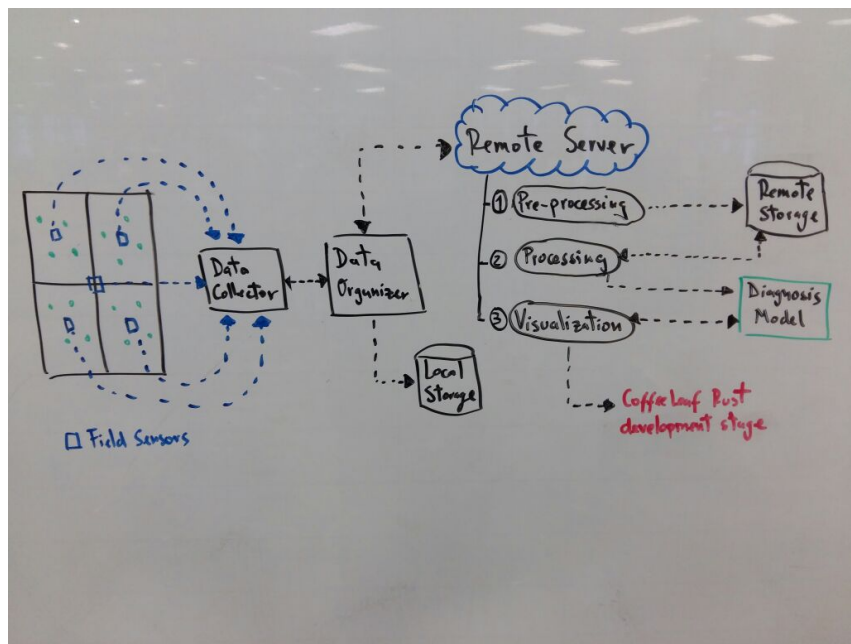
On the one hand, the physical part is composed by four raised wooden beds representing the lots and separated by four plastic curtains. Every lot contains four coffee plants and five sensors for measuring pH, soil temperature, soil moisture, illuminance and electrical conductivity, respectively. In addition, there is a drainage collector under each bed, so that the pH can be read, and a general rain system, which supplies every lot with water. Within each lot, there is an RGB camera system on the soil that goes around every plant capturing images from the bottom. Furthermore, a structure is put in the center of the four lots for supporting a rotary arm, which has a sensor on it for measuring environmental temperature/humidity and is also in charge of displacing the multi-spectral camera over the lots for capturing images from the top. Besides, there is a general circuit box that preserves the whole electrical components safe.

On the other hand, the informatic part has a collector for joining the data coming from the scale crop and an organizer for structuring it. After being properly arranged, the data is stored locally and sent to the remote server over the Internet. Then, it is cleaned (preprocessed) and stored again on the server's File system. Lastly, the processing takes place, which generates the diagnostic model as a result. That model estimates the Coffee Leaf Rust development stage and can be incorporated in the

creation of a posterior visualization that presents the results graphically.



(a) Final physical concept of the prototype.



(b) Final informatic concept of the prototype.

Fig. 8.2.: Final concept of the prototype.

9. DETAILED DESIGN APPROACH

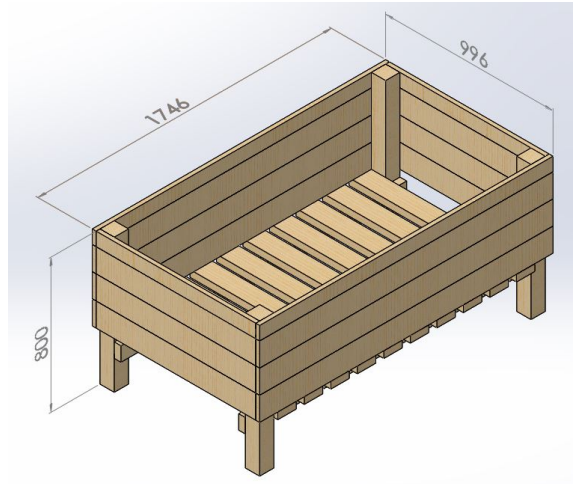
Continuing with the *Pahl & Beitz* methodology, this chapter specifies and structures the arrangement, forms, dimensions, properties, materials and technologies for the realization of the individual components that shape the final concept for data collection and processing. The information presented here is fundamental for the transformation of the mentioned concept into an integrated and functional cyber-physical prototype that allows to gather the necessary data during the experiment, create a diagnostic model and achieve the performance evaluation to answer the research question.

9.1 Mechanical Design

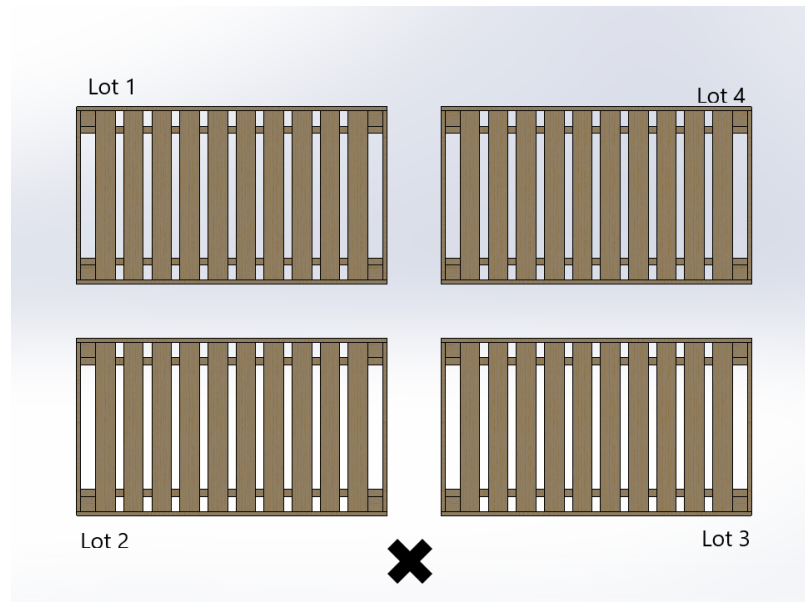
For the construction and assembly of the final prototype, the Design team developed a group of mechanical sub-assemblies. Each of them was in charge of a specific function in order to accomplish the data collection correctly. The most relevant sub-assemblies are described in detail below.

9.1.1 Coffee crop lots

For storing the coffee plants, four crop lots are designed, each one housing a set of four plants with a different agronomic management. The lots are identical in dimensions and arranged as shown in Figure 9.1, where the **X** mark in Figure 9.1(b) indicates the reference standing point. The construction drawings of the structure can be seen in Appendix E.



(a) Coffee crop lot modelling.



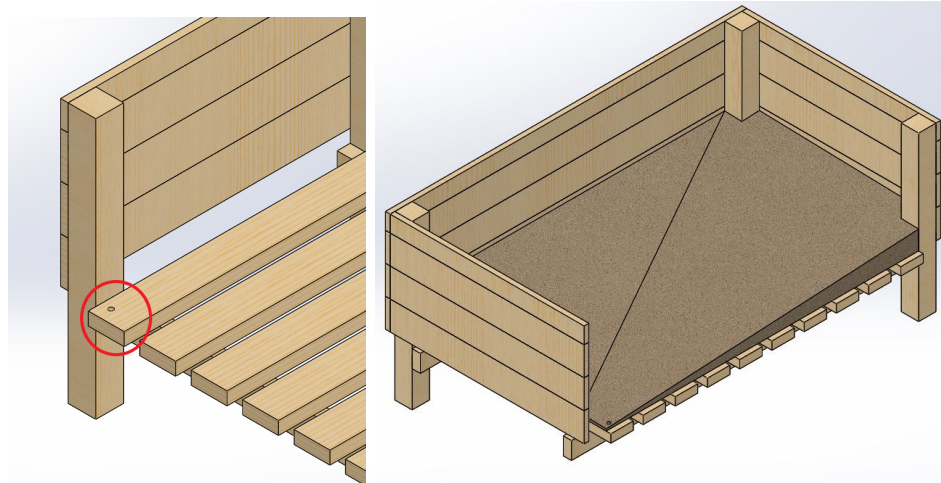
(b) Coffee crop lots arrangement.

Fig. 9.1.: Coffee crop lots.

9.1.2 pH measurement system

The pH of each lot is measured on its residual water due to the pH-meter's characteristics and limitations given by the manufacturer. This sensor is not suitable for being buried in the earth and it can be only used in liquid substances. Therefore,

each crop lot has a filtering point and a slope, as shown in Figure 9.2, that directs the water to a container. The slope is built with soil and covered with an impermeable plastic and a mesh for retaining large soil components.

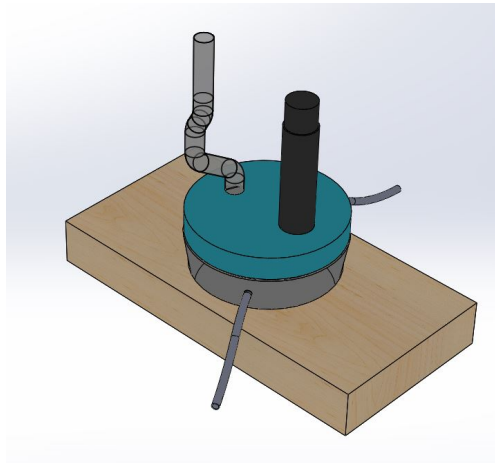


(a) Filtering point.

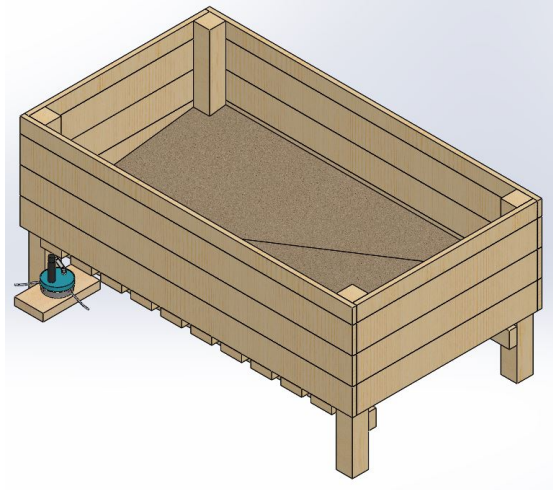
(b) Filtering slope.

Fig. 9.2.: Crop lot filtering.

At the filtering point, a hose is connected for directing the water to the container, where the pH-meter is located. It should also be noted that the container has two drain hoses on its sides for eliminating the excess and trying to keep the most recent residual water for the measurements. The Figure 9.3 illustrates the sub-assembly, whereas the construction drawing of this system can be checked in Appendix F.



(a) pH container modelling.

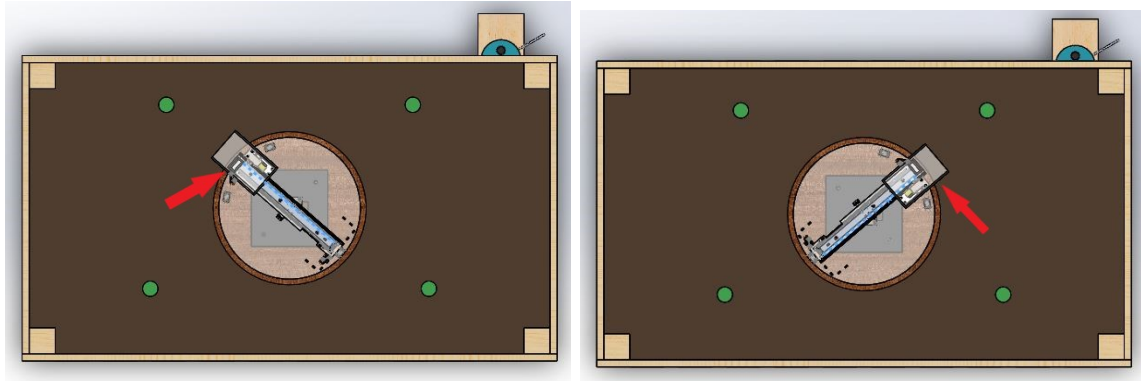


(b) pH system sub-assembly.

Fig. 9.3.: pH measurement system.

9.1.3 Rotary platform

This component rotates to all positions where a coffee plant is placed in order to allow capturing images from the bottom. The Figure 9.4, in which the green circles represent the plants, shows the rotary platform at two different stop-positions. In Figure 9.4(a), the platform stopped at position one whereas the Figure 9.4(b) illustrates the platform after stopping at position four.

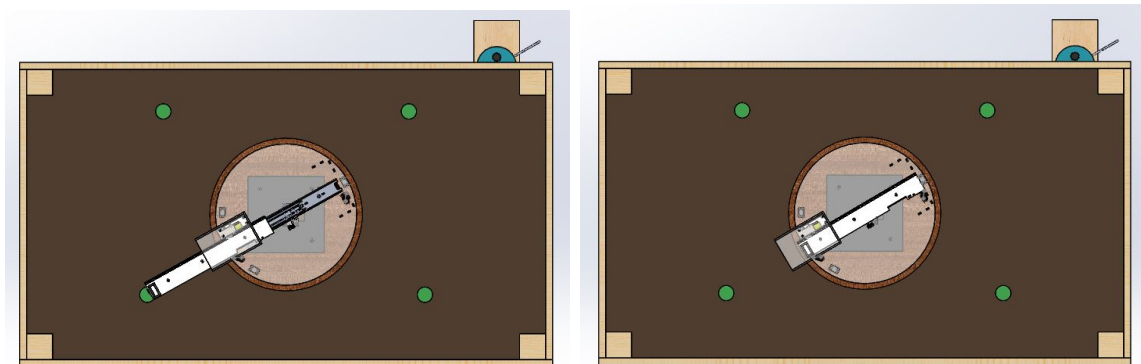


(a) Position one.

(b) Position four.

Fig. 9.4.: Rotary platform movement.

When the rotary platform is at a stop-position, a rack-pinion mechanism, called rotary platform's extension, drives a web-cam close to the corresponding plant. Thereby, the camera captures an RGB image from the bottom and, once completed, the same mechanism returns it to its initial position (aka "home"), as shown in Figure 9.5. The construction drawings of the rotary platform can be seen in Appendix G.



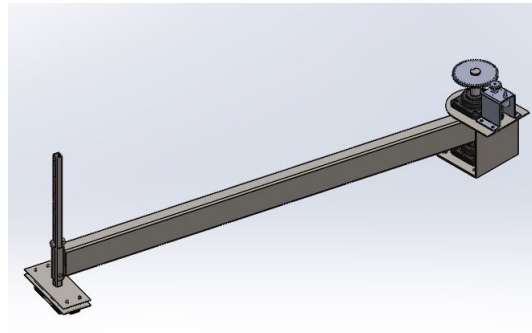
(a) Position two - Extended.

(b) Position two - Home.

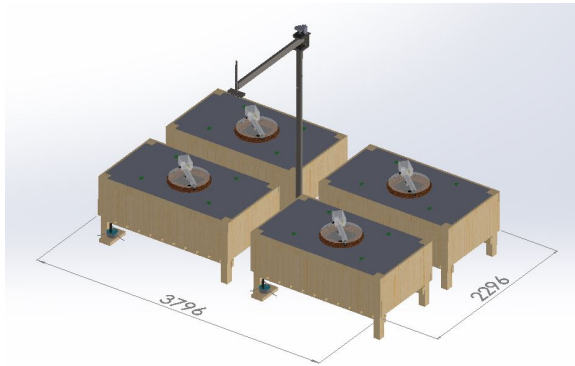
Fig. 9.5.: Rotary platform's extension (rack-pinion) mechanism.

9.1.4 Rotary arm

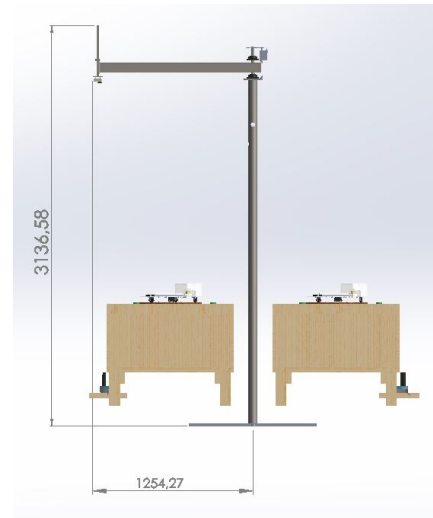
The Figure 9.6 illustrates the modelling of the rotary arm. This sub-assembly is able to rotate and position over each lot in order to allow capturing images of the plants from the top, using two multi-spectral cameras. These cameras, whose characteristics are detailed later in Section 9.2, are placed at the end of the structure (left side in Figure 9.6(a)).



(a) Rotary arm modelling with multi-spectral cameras container.



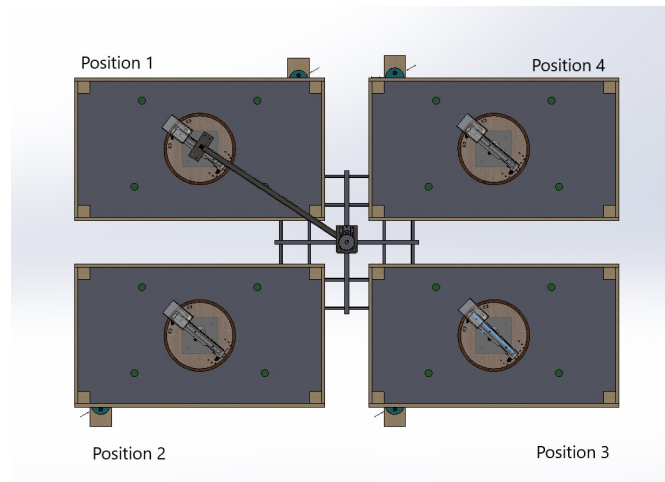
(b) Rotary arm sub-assembly - Isometric view.



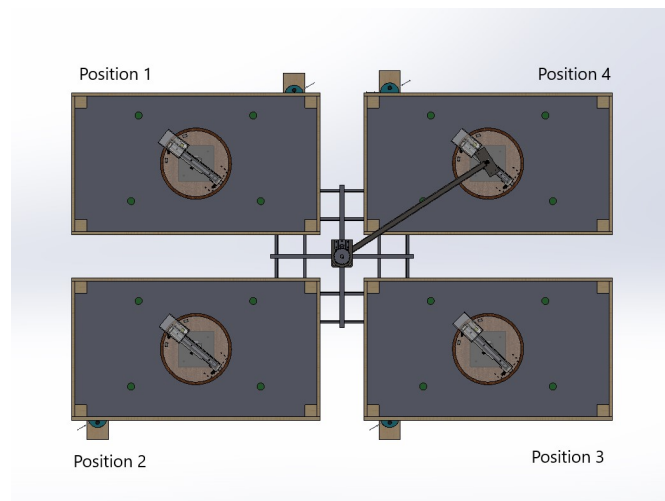
(c) Rotary arm sub-assembly - Side view.

Fig. 9.6.: Rotary arm.

Just like the rotary platform, the rotary arm has four stop-positions, each of them corresponding to each lot. The arm's initial position is always over the first lot (position one) and it can turn counterclockwise covering all lots until position four. Finally, once the multi-spectral cameras captured all necessary images, it returns to its initial position. The Figure 9.7 exemplifies the movement of the rotary arm. Its construction drawings can be seen in Appendices H and I.



(a) Position one.

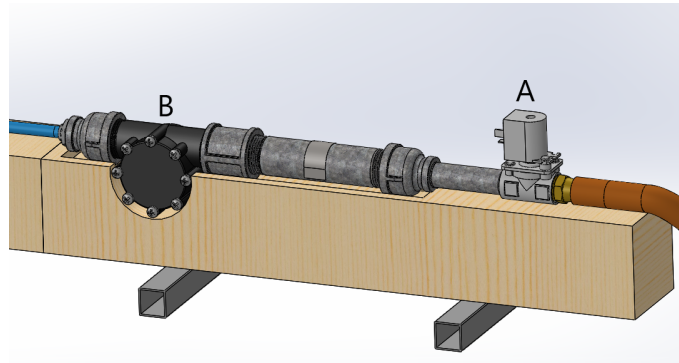


(b) Position four.

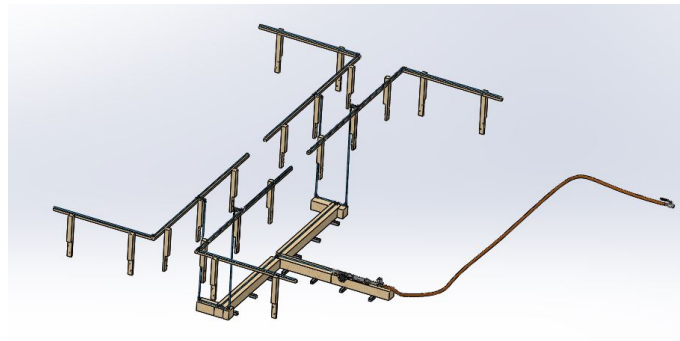
Fig. 9.7.: Rotary arm movement.

9.1.5 Rain system

This system is designed in order to enable the irrigation by emulating the rain over the scale coffee crop, which is crucial for the plants' growth. In that sense, a solenoid valve is used for controlling the activation of the system (enable/disable the passage of water) while a flow sensor helps to keep in track how much water is being used. Then, a hose system divides the flow and directs water to each lot. Lastly, there is a wooden L-structure on each lot that supports and elevates three perforated hoses, which allow to produce the rain effect. The Figure 9.8 shows the modelling of the rain system, whereas the drawings for its construction can be checked in Appendix J.



(a) Rain system's primary hose connected to the solenoid valve (A) and the flow sensor (B).



(b) Rain system sub-assembly.

Fig. 9.8.: Rain system.

The Figure 9.9 illustrates the final modelling of the prototype's mechanical part after assembling all described components into a single larger system.

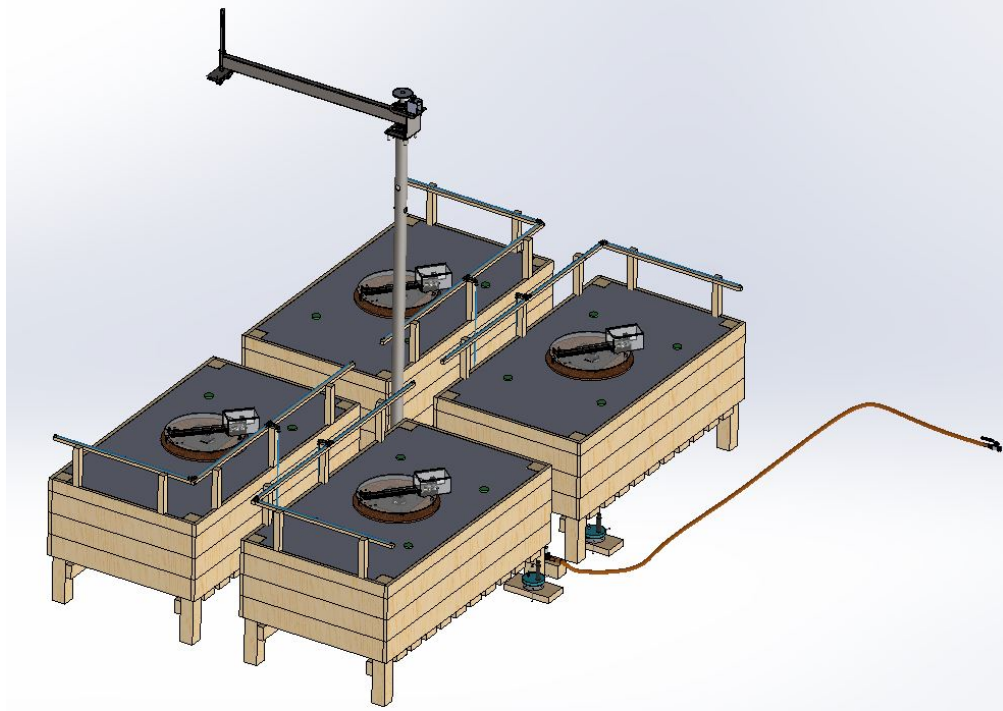


Fig. 9.9.: Final modelling of the mechanical part of the prototype.

9.2 Electronic Design

The Electronic Design details how the electronic components, such as sensors, actuators, interfaces, power supplies, micro-controllers, among others, are connected and integrated with each other in order to receive commands from the prototype's cybernetic component, execute them on the scale coffee crop and send a response back. Therefore, it is relevant to explain in detail the different conditioning interfaces, along with the proper power supply for each sensor, the drivers that control the actuators, the electronic connections between the components, the communication with the cybernetic part of the prototype, as well as the distribution of all parts inside an electrical cabinet.

9.2.1 Arduino 1: Lot data collection

To acquire data from the different sensors and actuators on each lot, an Arduino® Mega 2560 is used. This micro-controller (named Arduino 1 in the context of this research) is necessary for connecting the electronic components and executing the communication with them during the data collection process. With the purpose of sensing and controlling the lots independently, a pH-meter, a lux sensor, a soil temperature sensor, a soil moisture sensor, three micro-switches, a mini-DC motor and a digital servo are installed on each of them. It is noteworthy that the last three components are necessary for the operation of the rotary platform. In addition, an environmental temperature/humidity sensor is considered for the whole scale crop whereas the electrical conductivity is excluded from the study because the potential sensors are not suitable. It should also be noted that, in order to ensure a *plug-and-play* connectivity, the majority of components are commercial ones.

To guarantee the reliability and speed of the measurements during the data collection process, the sensors and actuators should have minimum range of response and high accuracy. The selected electronic components for Arduino 1, including their specifications, are presented below.

For measuring pH, the analog pH-meter Pro from *DFRobot*® is selected (see Figure 9.10). This sensor is suitable for long term online monitoring and is considered to be robust, because its sensitive glass membrane provides a measuring range of 0–14 with an accuracy of ± 0.1 (at 25°C), as well as a response time less than or equal to one minute. The sensor's interface, shown in Figure 9.10(b), is only produced by the manufacturer. Additionally, the pH-meter Pro requires a supply voltage from 3.3 to 5V [94].

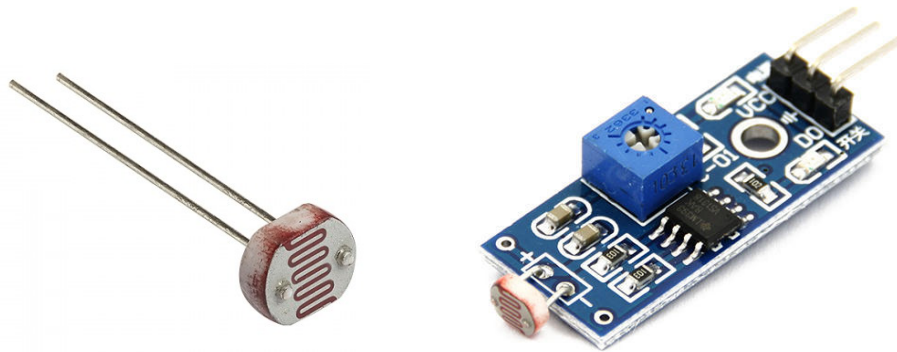


(a) pH electrode sensor (with BNC connector).

(b) pH sensor circuit board.

Fig. 9.10.: Analog pH sensor/meter pro kit for Arduino® [94].

For its part, a 5V-powered Light-dependent Resistor (LDR) is used to measure the illuminance in each lot (see Figure 9.11). This sensor is chosen due to its low cost and flexibility in the calibration of the physical variable (resistance value). The electronic component acts like a variable resistance, which changes according to the perceived light intensity. The Figure 9.11(b) shows the necessary circuit board for reading the analog signal [95].



(a) Light-dependent resistor (LDR).

(b) Interface for the LDR.

Fig. 9.11.: Lux sensor [95].

Measuring variables on the soil becomes a challenge in terms of the protection requirements for the electronic components. These requirements involve water resistance, durability and reliability. In order to collect soil temperature and moisture, two sensors from *DFRobot*[®], which require a voltage of 5V, are selected. The DS18B20, shown in Figure 9.12, is a waterproof temperature sensor with measurements from -10°C to 85°C and an accuracy of $\pm 0.5^{\circ}\text{C}$. The required interface for properly acquiring and conditioning the analog signal is also presented in Figure 9.12(b) [96].

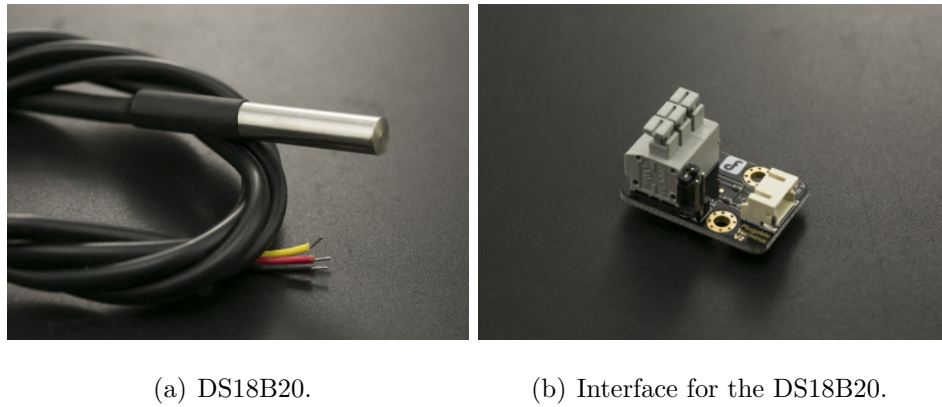


Fig. 9.12.: Waterproof soil temperature sensor [96].

The sensor used for soil moisture readings was a capacitive soil moisture sensor (see Figure 9.13). It is analog, resistant to corrosion, requires manual calibration and has a response time less than or equals to one minute [97].



Fig. 9.13.: Analog capacitive soil moisture sensor [97].

It is also considered as important to measure the micro-climate of each lot i.e. its environmental temperature and humidity. To this end, a digital temperature and humidity sensor called DHT22 is chosen. This sensor, shown in Figure 9.14, is provided by *DFRobot*[®] and will be placed on the rotary arm. The DHT22 is a capacitive digital temperature and humidity module, which ensures a high reliability and excellent long-term stability due to its sensing technology. Additionally, it has a temperature range of -40°C to 80°C with a 1°C resolution and a maximum error of $\pm 0.5^{\circ}\text{C}$. For its part, the humidity range is between 0 % and 100 %, with a 0.1 % resolution and an error less than or equal to 2 % [98].

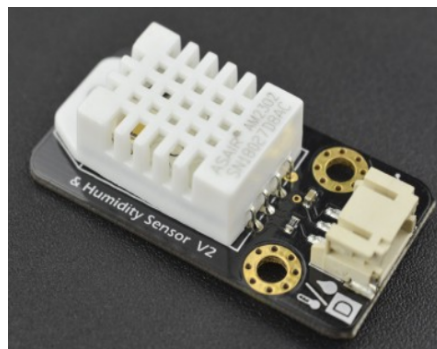


Fig. 9.14.: Environmental temperature/humidity sensor [98].

In addition to the sensors connected to the Arduino 1, there are two types of actuators that allow the operation of the rotary platform. These actuators are essential to move the web-cam close to each coffee plant in order to capture the necessary images. One of them is the Dynamixel AX-12A digital servo from *ROBOTIS*[®] (see Figure 9.15), which is used to guarantee the precise rotation of the platform so that the camera faces each plant on the lot [99].



Fig. 9.15.: Dynamixel AX-12A actuator [99].

This advanced actuator, which has become the standard for robotic projects, operates with 12V/1.5A, has a stall torque of 15.3 kg·cm, a 0°– 300°(59 rpm) resolution and communicates through an Universal Asynchronous Receiver-Transmitter (UART) protocol. Additionally, it is possible to track its speed, temperature, position, torque and load. Moreover, it has a built-in micro-controller, whose control algorithm allows to precisely preserve the shaft's position [99]. However, an extra compatible digital servo shield from *DFRobot*[®] (see Figure 9.16) is required for the interoperability with Arduino[®] [100].

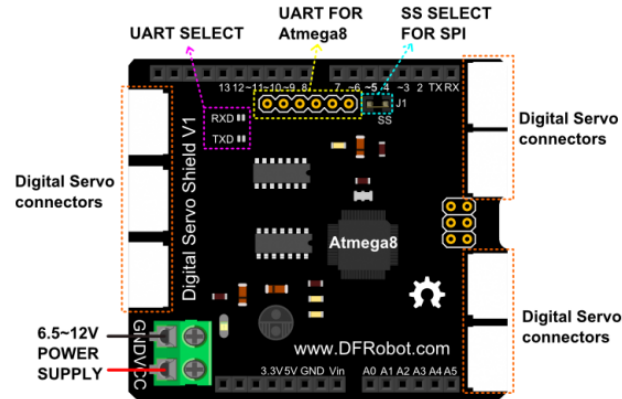


Fig. 9.16.: Digital servo shield for Arduino[®] [100].

In addition to the rotation of the platform, another actuator is required to perform the extension using the rack-pinion mechanism. For that purpose, a motor-reducer (aka mini-DC motor), powered by 5V and shown in Figure 9.17(a), is used. The specifications of the selected electronic component are sufficient for the weight that the mechanism needs to move (~ 230 g). It has a speed-ratio of 1:120, a current consumption without load less than or equal to 350mA and an approximate rotational speed of 110 rpm [101]. Furthermore, a H-bridge, called L298N, is incorporated in order to control the motor-reducer. The L298N (see Figure 9.17(b)) is a driver capable of handling the speed and direction of at most two DC-motors by virtue of two transistors integrated to it. This module works with a voltage range of 3V – 34.5V, a maximum current of 2A and a logic control voltage of 5V [102]. Since the H-bridge is able to control two DC-motors simultaneously, only two of them are necessary for controlling the required four motor-reducers.

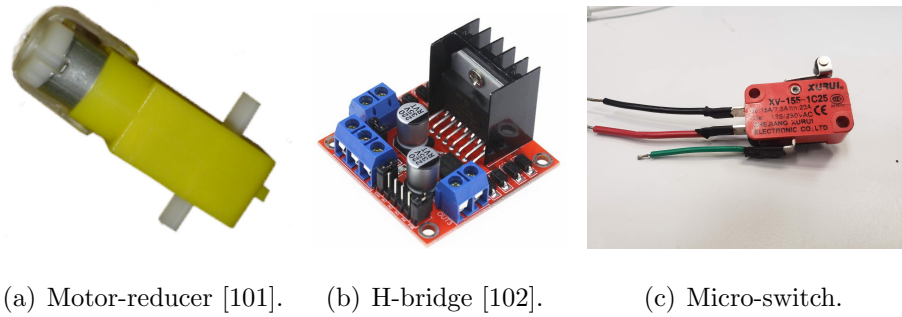


Fig. 9.17.: Electronic components of the rotary platform's extension [101, 102].

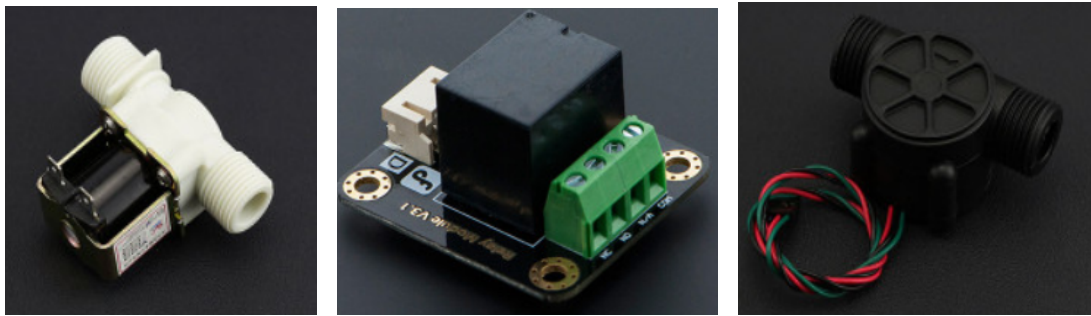
Besides, three micro-switches, like the one on Figure 9.17(c), are necessary to send a digital signal to the Arduino 1 as stop criterion for the rotary platform's extension. Taking into account that the coffee plants are not placed at the same distance from the platform, as seen in Figure 9.5, and that it is necessary to return the mechanism to the “home” position, the micro-switches should be installed in three different positions. Those positions are: **0** (no extension for “home”), **1** (partial extension for near plants) and **2** (total extension for distant plants).

The Appendix K contains the detailed schema of the connections between each of the presented electronic components and Arduino 1. The corresponding pins and necessary power supplies are also exposed in it.

9.2.2 Arduino 2: General data collection

Another Arduino[®] Mega 2560, called Arduino 2 in the context of this research, is necessary for the collection of general data and the plants' growth. It is in charge of controlling the rain system, retrieving the wind sensor's readings and moving the rotary arm for capturing the necessary multi-spectral images. The selected electronic components are explained below.

For the rain system, three essential components are needed: i) a solenoid valve, ii) a relay module and iii) a flow rate sensor. The solenoid valve, shown in Figure 9.18(a), is used to enable/disable the flow of water through the system. It has a rated voltage of 12V, a current of 0.4A, a working pressure of 0.02 MPa – 0.8 MPa and a withstand static pressure of 2.0 MPa [103]. This actuator cannot be remotely activated without the proper circuit board. Therefore, a relay module (see Figure 9.18(b)) from *DFRobot*[®] has to be connected to the valve. This module provides a digital connection for Arduino[®] that allows its activation/deactivation and therefore, the change on the solenoid valve's state. The relay module has a switching current of 10A in NO and 5A in NC and a maximum switching voltage of 150VAC / 24VDC [104].



(a) Solenoid valve.

(b) Relay module.

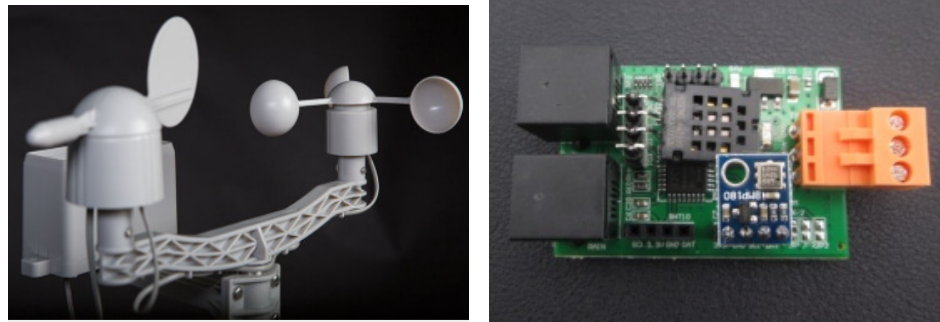
(c) Flow rate sensor.

Fig. 9.18.: Electronic components of the rain system [103–105].

Additionally, a flow rate sensor, which is provided by the same manufacturer and shown in Figure 9.18(c), is necessary for getting feedback about the volume of water being used. Its specifications include a voltage range of 3.5V – 12V, an operating current of 15mA and an accuracy of $\pm 5\%$ for readings between 2 L/min and 30 L/min [105].

For its part, a weather station kit with anemometer/wind vane/rain bucket from *DFRobot*[®] is required for measuring the wind variable. The kit is shown in Figure

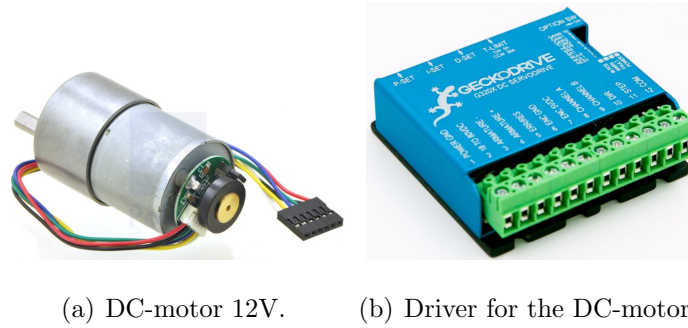
9.19(a). For this research, only the anemometer (wind sensor) is used since the only variable to be recorded is the wind speed. Additionally, this kit needs the circuit board showed in Figure 9.19(b) in order to retrieve the readings through a serial communication. The circuit board requires a 5V power supply with a current consumption of 10mA [106].



(a) Anemometer/wind vane/rain bucket. (b) Circuit board for weather station.

Fig. 9.19.: Weather station kit [106].

Considering the weight of the rotary arm and its length, a DC-motor with high torque is required for moving it over the lots. Therefore, the actuator to be used is the *Pololu*[®] (1443) 30:1 12V, shown in Figure 9.20(a). It is a brushed DC motor with a 30:1 metal gearbox and an integrated quadrature encoder that provides a resolution of 64 counts per revolution of the motor shaft. Additionally, at 12V, it reaches a no-load speed of 350 rpm with a current consumption of 300 mA, and a stall torque of 110*in with a current of 5000mA [107]. To control the motor's voltage, current, speed and direction, a robust and precise driver is required. This driver, shown in Figure 9.20(b), is the G320X. It supports a power supply range between +18VDC and +80VDC, which allows a maximum motor's current of 20A, has a quiescent current of 30mA, a 5VDC power supply for the encoder and channels for reading signals with a switching frequency of 20 kHz [108].



(a) DC-motor 12V. (b) Driver for the DC-motor.

Fig. 9.20.: Electronic components of the rotary arm [107, 108].

Taking the minimum operating voltage of the G320X into account, a buck-boost switching regulator called LTC3780 is also required. This component, shown in Figure 9.21, guarantees an output voltage of +18VDC from a 12VDC power supply due to its input voltage range of 5V – 32V and an adjustable output range of 1V – 30V. Its maximum output current is 10A.



Fig. 9.21.: Buck-boost switching regulator [109].

The Appendix L contains the detailed schema of the connections between each of the presented electronic components and Arduino 2. The corresponding pins and necessary power supplies are also exposed in it.

9.2.3 Raspberry Pi®

Both Arduino 1 and Arduino 2 operate according to certain commands that they receive through a Serial connection. For sending those commands and being able to manage the data collection process, a Raspberry Pi® 3 B+ is selected. This single-board computer, shown in Figure 9.22, is responsible for orchestrating the sequence of commands which are necessary to execute a data collection routine. To this end, not only both Arduinos but also all cameras are connected to it.



Fig. 9.22.: Raspberry Pi® [110].

The Raspberry Pi® is powered by the Broadcom BCM2837B0 communication processor (Cortex-A53 (ARMv8) 64-bit SoC @ 1.4 GHz), it has a 1GB LPDDR2 SDRAM, an extended 40-pin GPIO header and four USB 2.0 ports. The required power supply is 5V/2.5ADC [110]. Considering that two Arduinos, four web-cams (one per lot) and two multi-spectral cameras will be used for the data collection, it is necessary to incorporate an external powered USB-hub in order to have enough available ports. Thus, the USB-hub shown in Figure 9.23 is selected. This electronic component has seven USB 3.0 ports and requires a power supply of 5V/2A.



Fig. 9.23.: USB 3.0 super speed hub.

Regarding the web-cams, the Logitech® camera, shown in Figure 9.24, is chosen for capturing RGB images from the bottom of the plants. This web-cam is a portable, affordable, has a compact size and a flexible support that allows to set it in different angles. Additionally, this component has a current consumption of 250mA and a resolution of 720p [111].



Fig. 9.24.: C270 Logitech HD webcam [111].

With respect to the multi-spectral cameras, which are used to capture the reflected radiation of wavelengths that are not perceptible for the human eye, two cameras from *MAPIR*®, called Survey3 and shown in Figure 9.25, are used. Based on the information of Chapter 6, the Red+Green+NIR (RGN) and RedEdge (RE) camera-filters are chosen for being suitable to identify crop diseases, including the Coffee Leaf

Rust. Thus, one camera centered in the wavelengths 660 nm – 550 nm – 850 nm and another one centered in the 735 nm wavelength are selected for capturing images from the top of the plants. The Survey3 incorporates a Sony[®] Exmor R IMX117 12MP sensor and a sharp non-fish-eye lens for perceiving light in specific wavelengths. It is able to capture images with a default timer-trigger or when receiving a Pulse-width modulation (PWM) signal through its HDMI port. Besides, it includes an USB-transfer-mode for accessing the media, which can be entered or exited using also a PWM signal [112,113].



Fig. 9.25.: Survey3 multi-spectral camera [112,113].

The detailed electronic schema of the Raspberry Pi[®] can be seen in Appendix M. Its electronic connections with the Arduinos, RGB and multi-spectral cameras are shown in it.

9.2.4 Electrical cabinet

In order to protect, store and structure the components from the Arduino 1, Arduino 2 and Raspberry Pi[®], an electrical cabinet is required. This cabinet must have at least an IP5 environmental protection, due to the harsh conditions to which the system will be exposed inside the greenhouse. In addition, this cabinet must have electrical protection to ensure the integrity of the electronic components. To this

end, a current security breaker is necessary to avoid damages due to a peak current of 10A, considering that the calculated prototype's total current consumption is approximately 7A. This security is also attached to the 110VAC power supply to avoid possible inconsistencies. Furthermore, two independent 12V power supplies are used for the solenoid valve, the digital servos and the DC-motor, along with an individual protection fuse of 4A. For its part, an ATX 500W (20A) power source, with a protection fuse of 2A, is in charge of distributing 5V to all sensors and actuators that require it.

The Appendix N summarizes the distribution of the electrical/electronic components, including the whole power supplies. This cabinet is structured into three principal sections: i) a bottom section for the sensors circuit boards, 5V-GND power supply and the security breaker, ii) a middle section for Arduino 1 and Raspberry Pi[®], and iii) a top section for Arduino 2 and the USB-hub.

It is noteworthy that two extra 9V/1A AC/DC adapters are necessary to power each micro-controller in order to prevent them from consuming USB-hub's current. Consequently, two extension cords are required to be able to provide 110VAC to all mentioned electrical devices. In that sense, the AC/DC power adapters for the Raspberry Pi[®] and both Arduinos, the USB-hub, the ATX power source, the 12V power supplies, as well as a WiFi router for data transfer, are connected to those cords.

Lastly, it should also be noted that the cabinet's input/output points for the cables corresponding to the power sources, AC/DC adapters, RGB and multi-spectral cameras, sensors and actuators are displayed in Appendix N. Approximately 800 m of 22 AWG cable, which can handle a current of 5A, are needed for signals and 5V-GND supply, whereas about 100 m of 24 AWG cable are required for supplying 12V.

9.3 Informatic Design

The Informatic Design details how the physical components are communicated with the cybernetic part of the prototype in order to orchestrate the data collection, transfer and posterior processing properly. Thus, it is important to explain the functions, commands, components, architectures, content specifications and platforms, which are thought as necessary for managing the incoming and outgoing data flow and finally determining the Coffee Leaf Rust development stage.

9.3.1 Arduinos

To execute the necessary actions according to the commands sent from the Raspberry Pi® to the Arduinos, different functions are defined for the communication with each sensor and actuator. In that sense, the Figure 9.26 shows the functional schema of Arduino 1.

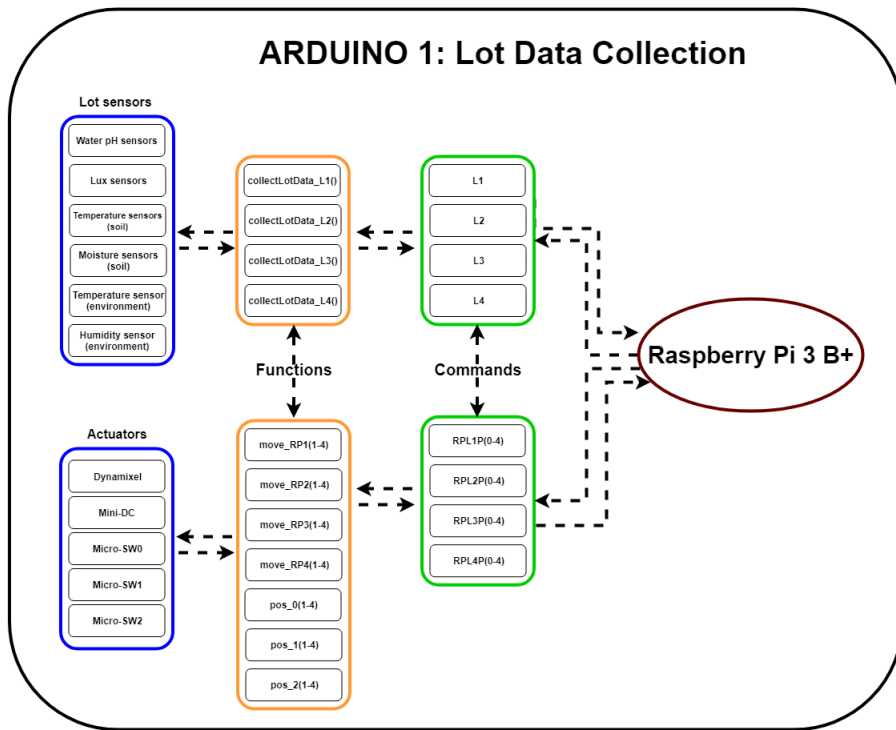


Fig. 9.26.: Functional scheme of Arduino 1.

For its part, the functional schema of Arduino 2 is presented in Figure 9.27.

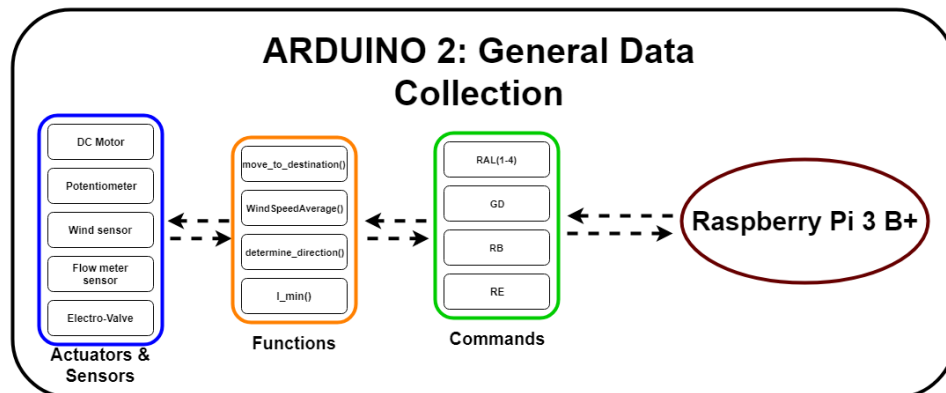


Fig. 9.27.: Functional scheme of Arduino 2.

In addition, the Table 9.1 describes the functions for the acquisition, conditioning and storage of the values of each sensor and actuator of the Arduino 1, whereas the Table 9.2 presents the functions that are determined for managing the components wired to the Arduino 2. The result after executing each function is also explained.

Table 9.1.: Functions of Arduino 1.

Function	Description	Result
collectData_L1()	Collect, condition and store readings acquired from lot 1 field sensors	Global stored values of pH, soil temperature/moisture, illuminance and environmental temperature/humidity from lot 1
collectData_L2()	Collect, condition and store readings acquired from lot 2 field sensors	Global stored values of pH, soil temperature/moisture, illuminance and environmental temperature/humidity from lot 2
collectData_L3()	Collect, condition and store readings acquired from lot 3 field sensors	Global stored values of pH, soil temperature/moisture, illuminance and environmental temperature/humidity from lot 3
collectData_L4()	Collect, condition and store readings acquired from lot 4 field sensors	Global stored values of pH, soil temperature/moisture, illuminance and environmental temperature/humidity from lot 4
move.RP1(x)	Rotate dynamixel of lot 1 to the position x (where $x \in [0, 1, 2, 3, 4]$)	Global stored value of lot 1 dynamixel position (0-300°)
move.RP2(x)	Rotate dynamixel of lot 2 to the position x (where $x \in [0, 1, 2, 3, 4]$)	Global stored value of lot 2 dynamixel position (0-300°)
move.RP3(x)	Rotate dynamixel of lot 3 to the position x (where $x \in [0, 1, 2, 3, 4]$)	Global stored value of lot 3 dynamixel position (0-300°)
move.RP4(x)	Rotate dynamixel of lot 4 to the position x (where $x \in [0, 1, 2, 3, 4]$)	Global stored value of lot 4 dynamixel position (0-300°)
pos.0(x)	Move RGB camera arm extension's to position 0 from plant x (where $x \in [1, 2, 3, 4]$)	Global stored value of RGB camera arm extension's state (0)
pos.1(x)	Move RGB camera arm extension's to position 1 from plant x (where $x \in [1, 2, 3, 4]$)	Global stored value of RGB camera arm extension's state (1)
pos.2(x)	Move RGB camera arm extension's to position 2 from plant x (where $x \in [1, 2, 3, 4]$)	Global stored value of RGB camera arm extension's state (2)

Table 9.2.: Functions of Arduino 2.

Function	Description	Result
<code>l_min()</code>	Collect, condition and store readings acquired from flow-rate sensor when solenoid is enable	Global stored value of the flow rate (l/min)
<code>determine_direction()</code>	Determine direction of the DC motor based on position	Pin state for DC motor direction
<code>WindSpeedAverage()</code>	Read, condition and store readings from the serial buffer	Global stored value of the wind speed (m/s)
<code>move_to_destination()</code>	Enable PWM pin to activate DC motor until destination	DC motor move to the given destination

9.3.2 Communication commands Raspberry Pi[®]-Arduinos

In order to collect data from the field sensors and capture the multi-spectral and RGB images, a series of commands are defined for the communication between the Raspberry Pi[®] and both Arduinos. Raspberry Pi[®] (the master) is in charge of managing the whole data collection process. For that purpose, it sends specific commands to the Arduinos (the workers) through a serial connection, so that they execute the necessary actions accordingly and return the result. The Table 9.3 explains the commands for the communication with Arduino 1 while the Table 9.4 presents the communication commands with Arduino 2.

Table 9.3.: Communication commands Raspberry Pi®-Arduino 1.

Command	Description	Result
L1\n	Get all field sensor readings and rotary platform's state (0/1) from lot 1	Readings for pH, soil temperature/moisture, illuminance, environmental temperature/humidity, rotary platform's angle/extension's state (0/1) from lot 1
L2\n	Get all field sensor readings and rotary platform's state (0/1) from lot 2	Readings for pH, soil temperature/moisture, illuminance, environmental temperature/humidity, rotary platform's angle/extension's state (0/1) from lot 2
L3\n	Get all field sensor readings and rotary platform's state (0/1) from lot 3	Readings for pH, soil temperature/moisture, illuminance, environmental temperature/humidity, rotary platform's angle/extension's state (0/1) from lot 3
L4\n	Get all field sensor readings and rotary platform's state (0/1) from lot 4	Readings for pH, soil temperature/moisture, illuminance, environmental temperature/humidity, rotary platform's angle/extension's state (0/1) from lot 4
RPL1P0\n	Move lot 1's rotary platform to the position one - Home	The lot 1's rotary platform's angle/extension's state (0/1)
RPL1P $\mathbf{x}$ \n	Move lot 1's rotary platform to the position $\mathbf{x}$ (where $\mathbf{x} \in [1, 2, 3, 4]$) - Extended	The lot 1's rotary platform's angle/extension's state (0/1)
RPL2P0\n	Move lot 2's rotary platform to the position one - Home	The lot 2's rotary platform's angle/extension's state (0/1)
RPL2P $\mathbf{x}$ \n	Move lot 2's rotary platform to the position $\mathbf{x}$ (where $\mathbf{x} \in [1, 2, 3, 4]$) - Extended	The lot 2's rotary platform's angle/extension's state (0/1)
RPL3P0\n	Move lot 3's rotary platform to the position one - Home	The lot 3's rotary platform's angle/extension's state (0/1)
RPL3P $\mathbf{x}$ \n	Move lot 3's rotary platform to the position $\mathbf{x}$ (where $\mathbf{x} \in [1, 2, 3, 4]$) - Extended	The lot 3's rotary platform's angle/extension's state (0/1)
RPL4P0\n	Move lot 4's rotary platform to the position one - Home	The lot 4's rotary platform's angle/extension's state (0/1)
RPL4P $\mathbf{x}$ \n	Move lot 4's rotary platform to the position $\mathbf{x}$ (where $\mathbf{x} \in [1, 2, 3, 4]$) - Extended	The lot 4's rotary platform's angle/extension's state (0/1)

Table 9.4.: Communication commands Raspberry Pi®-Arduino 2.

Command	Description	Result
GD\n	Get all general sensor readings, solenoid valve's state (0/1) and rotary arm's angle	Readings for flow rate, wind speed, solenoid valve's state (0/1) and rotary arm's angle
RAL1\n	Move rotary arm to lot 1	The rotary arm's angle
RAL2\n	Move rotary arm to lot 2	The rotary arm's angle
RAL3\n	Move rotary arm to lot 3	The rotary arm's angle
RAL4\n	Move rotary arm to lot 4	The rotary arm's angle
RB\n	Open the solenoid valve to enable water flow	The solenoid valve's state (0/1)
RE\n	Close the solenoid valve to disable water flow	The solenoid valve's state (0/1)

9.3.3 Raspberry Pi®

The Raspberry Pi® orchestrates every step during the routine that collects data from the scale coffee crop and executes this process seven times during the day automatically (with and without sunlight with the purpose of covering all sensors). It is also in charge of triggering when the solenoid valve's state has to change in order to activate/deactivate the rain system according to a pre-set calendar. The Raspberry Pi® communicates with both Arduinos, receives and organizes data, reports current state during the process to *Thingworx*® (IoT platform), triggers the cameras to capture the RGB and multi-spectral images, stores everything locally in a structured way and transfers the collected data to the Academic Data Center (ADC) (remote server) over the Internet using a WiFi connection. It is also relevant to mention that the software to be used for the implementation is *Python 3.5.3*, *OpenCV 3.4.1* [114], *RPi.GPIO 0.6.3* [115] and *MongoDB*® [116]

For the Raspberry Pi® to work properly, it is necessary to develop a set of components that fulfill different functions. After that, they can be integrated to achieve the automatic data collection. The components are explained below.

- **Data subdirectory creator:** creates a subdirectory and its internal structure on the main data directory each time a data collection routine begins.

- **Arduino[®] communicator:** establishes the communication with each Arduino[®], sends the commands and groups the returned results.
- **Data structurer:** organizes and transforms the data coming from both Arduinos in order to structure which value belongs to which reading.
- **Document local writer:** writes a document locally containing the collected field and general sensor data.
- **Thingworx[®] communicator:** reports the current state during the data collection process to Thingworx[®], so that the whole system can be remotely monitored in real-time.
- **Image capturer:** captures RGB images and saves them locally. It is also in charge of managing the rotation and resize operations on the multispectral images.
- **Pulse generator:** communicates with both multispectral cameras and generates the different pulses necessary to trigger them. It makes the cameras capture images and also puts them in transfer-mode in order to allow accessing the captures.
- **Data copier:** finds the most recent captured images on the multispectral cameras and copies them to the local storage.
- **MongoDB[®] communicator:** inserts the documents containing the field and general sensor data into a remote MongoDB[®]. MongoDB[®] is a scalable and flexible document-oriented database, which stores data in JavaScript Object Notation (JSON)-like documents [116].
- **VPN connector:** checks whether the Virtual Private Network (VPN) connection, which allows the communication with the ADC over the Internet, is established because it is mandatory for storing the data remotely. It is also in charge of verifying the Internet connection status.

- **Data uploader:** given a path to a local directory containing the collected data of a specific routine, creates the same directory structure in the ADC and uploads every single file to its corresponding place. It replicates the local storage remotely.

In addition, it is relevant to illustrate how the data is structured and specify the content of each data collection routine. When it is time to collect data, a directory with the format *ddmmYYYYHHMMss* is automatically created, where *dd*, *mm*, *YYYY*, *HH*, *MM* and *ss* represent day, month, year, hour, minutes and seconds, respectively. For instance, if the process occurs on December 20th 2018 at 16:30:04 hours, the created directory would be 20122018163004. The internal structure of the mentioned directory is composed by five subdirectories: `lot_1`, `lot_2`, `lot_3`, `lot_4` and `general_data`. Each `lot_x` subdirectory contains a JSON file with the corresponding field sensor data, four RGB images (one per plant on the lot) and two multispectral images (one per camera), while the `general_data` subdirectory has a single JSON file inside with the general sensor data. The JSON files' keys are detailed below.

- **ph:** the pH sensor reading (field data).
- **soil_temperature:** the soil temperature sensor reading (field data).
- **soil_moisture:** the soil moisture sensor reading (field data).
- **illuminance:** the illuminance sensor reading (field data).
- **env_temperature:** the environmental temperature sensor reading (field data).
- **env_humidity:** the environmental humidity sensor reading (field data).
- **rotary_platforms_angle:** the current rotary platform's angle (field data).
- **rotary_platform_extensions_state:** the current rotary platform extension's state (0/1) (field data).
- **lot_number:** the lot number to which the data belongs to (field data).

- **plant_1:** the path to the image of the plant in position 1 (field data).
- **plant_2:** the path to the image of the plant in position 2 (field data).
- **plant_3:** the path to the image of the plant in position 3 (field data).
- **plant_4:** the path to the image of the plant in position 4 (field data).
- **rgn:** the path to the Red+Green+NIR (RGN) multispectral image (field data).
- **re:** the path to the RedEdge (RE) multispectral image (field data).
- **flow_rate:** the flow rate sensor reading (general data).
- **wind_speed:** the wind speed sensor reading (general data).
- **solenoid_valves_state:** the current solenoid valve's state (0/1) (general data).
- **rotary_arms_angle:** the current rotary arm's angle (general data).
- **timestamp:** the timestamp of the data collection in Coordinated Universal Time *UTC* format (field and general data).

In order to clarify how a data collection routine works, the Figure 9.28 details the whole pipeline from the sensor readings and image captures until the remote storage. The data from internal and external sensors/actuators is gathered and smoothed by Arduino 1 and Arduino 2 respectively, whilst RGB and multispectral images are captured by the cameras. The totality of the data is collected by the Raspberry Pi[®], which constantly notifies the progress to *Thingworx*[®] while it creates a single data package organized under the mentioned structure. The package containing the documents with the lots and general data, as well as the images, is stored locally. In addition, the documents are inserted into the remote *MongoDB*[®], which resides in the ADC, and the entire data package is uploaded via Secure File Transfer Protocol (SFTP) to the ADC's File system. At that point, the data collection routine finishes.

Finally, it is also relevant to mention how the collected data can be reviewed so that the process can be verified. Using a personal computer, *Thingworx*[®], the Raspberry Pi[®] and the ADC can be accessed over the Internet. The access to the IoT platform just requires a web browser, while the Raspberry Pi[®] and the ADC can be remotely inspected through the graphical desktop sharing system Virtual Network Computing (VNC) or the cryptographic network protocol Secure Shell (SSH).

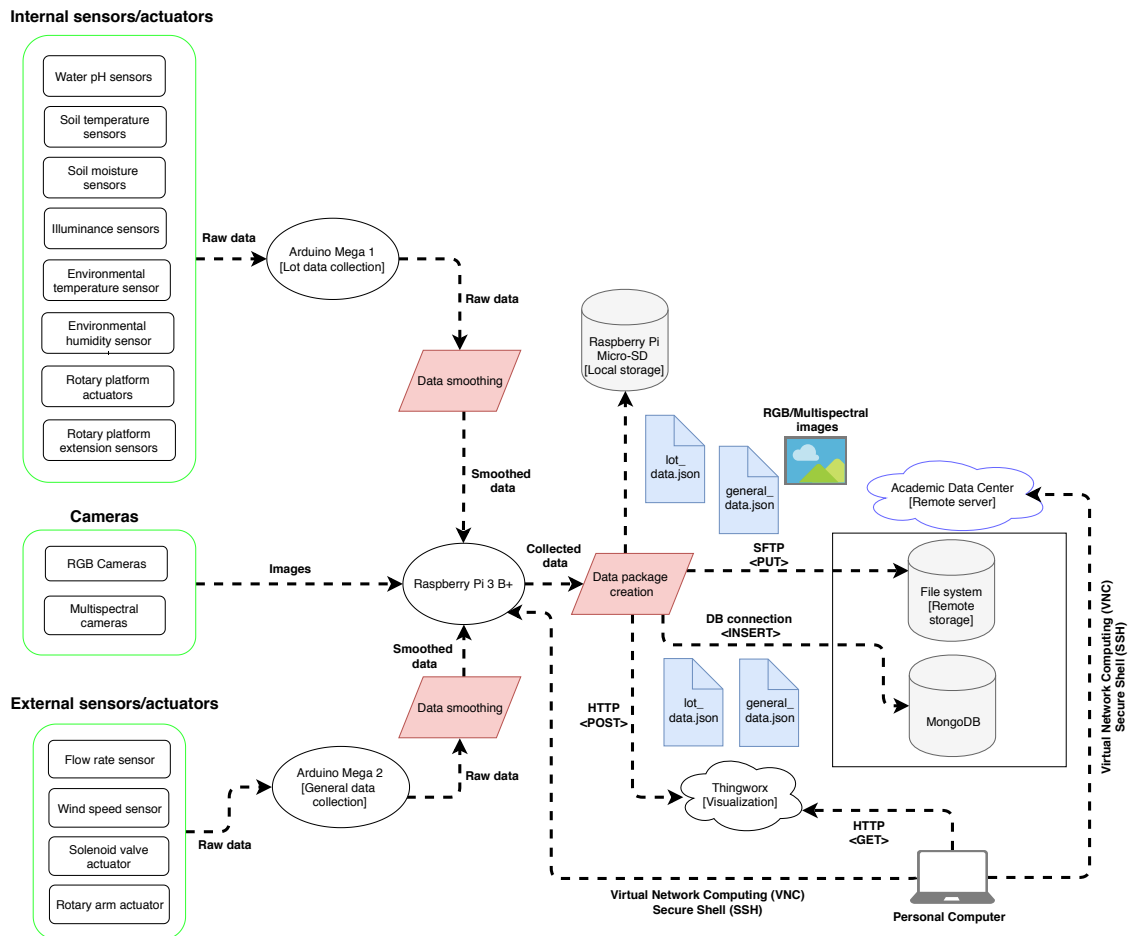


Fig. 9.28.: Data collection pipeline.

9.3.4 Thingworx®

Thingworx® is a complete technology platform designed for the industrial IoT [117]. Within the framework of this project, it is used to visualize the field and general data in real-time while the data collection routine is executed. Thanks to it, it is possible to remotely and timely check whether a sensor/actuator is reporting anomalous readings or the process is not behaving as it should and take corrective actions accordingly.

The Figure 9.29 presents the overview of the data visualization dashboard. It is composed by five blocks, four on the top for each lot and another one on the bottom for the general information. The widgets that are incorporated to create the dashboard are LED displays, Shapes and Labels. The LED displays show in blue numbers the values for pH, soil temperature/moisture, illuminance, environmental temperature/humidity, rotary platform's angle, flow rate, wind speed and rotary arm's angle. The Shapes make it possible to partially or totally fill their background with a specific color according to a particular value. Thus, they are useful to indicate the rotary platform extension's state and the solenoid valve's state. The Labels are plain text and they are used to clarify which widget belongs to which property as well as to show the timestamp of the last received reading. Additionally, it is important to mention that the Raspberry Pi® reports data to *Thingworx*® through Hypertext Transfer Protocol (HTTP) requests to the services that the platform exposes in order to update every single block on the visualization dashboard.



Fig. 9.29.: Data visualization dashboard.

9.3.5 Academic Data Center

The ADC is the remote server that the Computer Science Department of *Universidad EAFIT* provides to its students to potentiate their academic and research projects. It is composed by a powerful set of machines with high processing and storing capabilities (64 Cores, 512 GB RAM, 4 TB Storage, 1 GPU Tesla K80 and Ubuntu Operating System) and can be accessed over the Internet through a VPN connection and an username-password authentication.

Within the framework of this project, the ADC is used to remotely store the data collected on the physical part of the prototype. Both the *MongoDB*[®] instance in it as well as its File system make the replication of Raspberry Pi's local storage possible and facilitate the ubiquitous access to that information. Furthermore, the ADC is the place where the data preprocessing, model generation and Coffee Leaf Rust development stage diagnosis occur. It is also relevant to mention that the soft-

ware to be used for the implementation is *Python 3.6.0*, *NumPy 1.16.0* [118], *Pandas 0.24.0* [119], *Scikit-learn 0.20.2* [120] and *Keras 2.2.4* [121] running on top of *TensorFlow 1.12* [122].

The Appendix O illustrates how the collected data is manipulated in order to extract the model that will be used to diagnose the development stage of the disease in question. Five sub-directories ranging from 0 to 4 are created in order to classify the lot data directories according to their corresponding label (i.e. development stage). For that purpose, the Biology team determines the labels carrying out visual inspections on field on all plants once a day during the whole data collection phase. In that sense, it assigns a whole number between 0 and 4 to each plant on each lot evaluating the plant leaves' severity level and calculates the specific lot's label as the rounded average of its four plants' disease development stages. In order to avoid visiting the scale coffee crop several times per day, all data directories between the penultimate and last visual inspection corresponding to the same lot are labelled with the same value, which was determined on the most recent checkup. In that way, the visual inspections and the lot data classification in the mentioned five sub-directories happen simultaneously.

Additionally, it is crucial to mention that an extra **rgb_images** directory containing five sub-directories (ranging from 0 to 4) is created within every classified lot data directory that has RGB images. This is done to assign every image of this type to the correct label, considering that it may differ from the lot's one to which the image belongs. For its part, the JSON files and multi-spectral captures have the same label as the lot's one in which they reside. Furthermore, in the case of images in general, they are visually checked one by one in order to keep only the ones with valuable content (focus, brightness level, etc.) and remove the others. Besides this, a script is executed to eliminate the irrelevant JSON files (those with missing values and outliers) as well as the sub-directories that end up with no content. At the end, there will be

five sub-directories containing the data from the whole lots properly labelled. Those sub-directories are the ones that will be used for the generation and final evaluation of the diagnostic model, taking into account that the diagnosis will be at the lot level.

Once all data is correctly distributed, the content of each of the five sub-directories is virtually shuffled and the elements per file type are counted for every sub-directory. Then, per label sub-directory, the minimum of those values is found, its 25 % is calculated and the file type associated with that minimum is determined. The resulting numbers indicate the number of files per the respective determined file type and per label that can be used, at most, for testing the diagnostic model. Taking those threshold numbers into account, the shuffled lot data directories within each label subdirectory are individually analyzed with the purpose of splitting them into two groups, namely train and test sets. If a particular lot data directory is considered as complete (i.e. it has a JSON file, the two multi-spectral images and at least one RGB image) and supposing that using its files for testing does not exceed the corresponding threshold, then it is copied under the same structure to another location in order to feed the test set. Otherwise, the lot data directory is also copied but to grow the train set. Thereby, the train set ($\sim 75\%$ of all data) will be used to train and tweak the model whilst the test set ($\sim 25\%$ of all data) will be only incorporated at the time of the diagnosis evaluation.

After the two sets are properly obtained, one sub-model is generated for each file type i.e. JSON, RGB, RGN and RE. For the JSON files, MLP is used, whereas CNNs are implemented to classify the RGB, RGN and RE images. For that purpose, the data on the train set is first divided into four subdirectories according to the file type, while preserving the same structure. Then, each of them is preprocessed so that the noise is removed from the images and the irrelevant keys in the documents are also identified and eliminated. After that, the corresponding data is loaded within each sub-model's generation, divided into feature data (the files themselves) and label data (the names

of the label subdirectories that contain the files) and permuted. Thereby, the data is randomly mixed while it is still possible to know each feature’s respective label unequivocally. Then, if applicable, the data is normalized and structured in order to scale the input and format it as recommended when using deep ANNs. On having the data prepared, different architectures and hyperparameter values are tried out to train the sub-model with the purpose of tuning it for reaching higher performance values on the predictions.

The technique for tuning the sub-model is called Grid search with cross-validation. It consists in executing an exhaustive search over specified hyperparameter values for an estimator in order to find out which combination achieves the best performance, which is by default the higher accuracy but different metrics can be chosen. Using this technique, one candidate estimator for each combination of hyperparameters is built and evaluated, so that the best estimator, its attributes and its average performance can be extracted once the search is completed [123]. For its part, the procedure for measuring the average performance of each candidate estimator during the generation of the sub-model is called *k-fold* cross-validation, where *k* separate learning experiments are run on the same estimator in order to calculate *k* performance values and average them. To achieve this, the feature and label data are split at the beginning into *k* nonoverlapping subsets (aka “folds”), so that on every experiment one different fold is kept for measuring the performance, whereas the remaining *k* - 1 are put together to form the train set for fitting the estimator [124]. Finally, when the Grid search processes conclude, the four sub-models are extracted and saved for the definitive diagnosis about the Coffee Leaf Rust development stage.

For selecting the best estimator during the Grid search with cross-validation, the chosen metric is the F_1 -score, which, in the multi-label case, is the weighted average of the labels’ F_1 -scores. The F_1 -score of the label *L* is a value in the $[0, 1]$ range and it is calculated as the harmonic mean of the estimator’s precision and recall with respect

to L (see Equation 9.1). The precision with respect to L is the ratio of the number of times that L was correctly predicted to the overall number of times that L was predicted. For its part, the recall with respect to L is the ratio of the number of times that L was correctly predicted to the overall number of times that L should have been predicted. Thereby, the general F_1 -score reaches its best value at 1, indicating that the estimator perfectly matches the reality, and its worst at 0, showing that the estimator never coincides with the reality [125].

$$F_1\text{-score}_{\mathbf{L}} = \frac{2 * precision_{\mathbf{L}} * recall_{\mathbf{L}}}{precision_{\mathbf{L}} + recall_{\mathbf{L}}} \quad (9.1)$$

At this point, the data that was kept to be only incorporated at the time of the diagnosis evaluation is brought up. First, the sub-models are loaded. Then, each lot data directory contained in the test set is submitted to the following process. At the beginning, its data is divided according to the file type. After that, each type is sent to its corresponding sub-model, where it is first cleaned, normalized and structured applying the same particular procedures that were used to prepare the data for the sub-model generation. Subsequently, the sub-model makes his own prediction based on it. It is also relevant to mention that, considering that the diagnosis is made at the lot level, the RGB sub-model could be used up to four times per lot data directory before retrieving its result (which is the rounded average of its predictions). The final step consists in joining the outcomes of the four sub-models and calculating their rounded weighted average, taking the respective F_1 -scores into account to determine the weights. Thereby, the definitive lot's Coffee Leaf Rust diagnosis is obtained and it is recorded among with the processed lot data directory's label. Once the whole test set is covered, a table showing comparative results is generated for the statistical analysis and the performance reached by the composite model is assessed with the calculation of the F_1 -score.

10. PROTOTYPING AND RESULTS

After designing the cyber-physical prototype in detail, the next step consisted of materializing it. In that sense, this chapter shows the building process of the final prototype, the main difficulties that were faced, and the respective modifications that were made on all disciplines in order to achieve its correct integrated functioning and to address the research question. Additionally, the scope and limitations are described. Finally, the obtained results, including the diagnostic model, the comparative table for the statistical analysis and the performance reached, are also presented.

10.1 Mechanical Prototyping

The Mechanical Prototyping shows how the sub-assemblies were constructed and how they were joined to form the larger system. Here, the different steps involved in this process are explained and illustrated below.

10.1.1 Rotary arm

It began with the construction of the support structure for the rotary arm. First, the central tube was cut according to the required dimensions, and then, two passing holes were drilled for the installation of a set of curtains, as shown in in Figure 10.1, which were planned to work as a separation between the four crop lots.

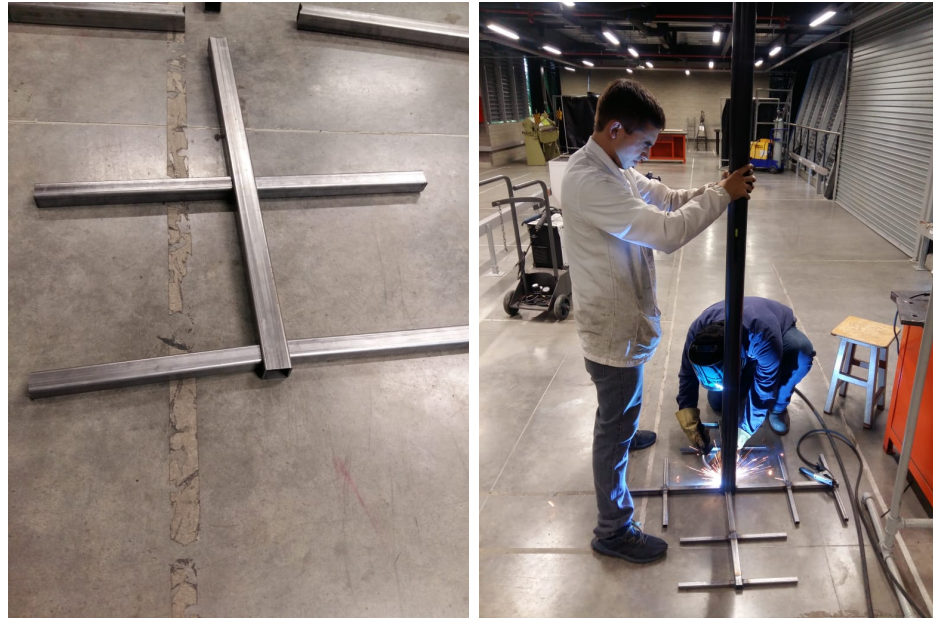


(a) Central tube drill mounting.

(b) Central tube being drilled.

Fig. 10.1.: Rotary arm's support manufacture.

Next, some steel square profiles were cut and welded meeting the design dimensions in order to shape a base structure. Then, they were welded to the central tube's end which was not drilled, as shown in Figure 10.2, with the purpose of serving as its support.



(a) Base structure.

(b) Welding base structure to the central tube.

Fig. 10.2.: Central tube's support manufacture.

At the other end of the central tube, a steel plate was welded for placing the rotary arm sub-assembly. After that, the whole structure was painted to protect it against corrosion. Figure 10.3 shows the rotary arm's finished support structure.



Fig. 10.3.: Finished rotary arm's support structure.

Referring to the arm as such, the construction began with the welding of a set of steel plates in a C-shape, which would be the base for all the components of the system. In addition, a bent plate was assembled for anchoring the drive motor in charge of the arm's rotation.

Then, two flange bearings and a steel shaft were mounted on the main base. Lastly, two spur gears were assembled, one on the arm's shaft and the other on the drive motor's shaft, in order to allow the arm's movement. It should also be noted that the arm's shaft was manufactured within the *Universidad EAFIT*'s workshops, while the bent plate and gears were made outside the campus. Figure 10.4 illustrates the mentioned structure.

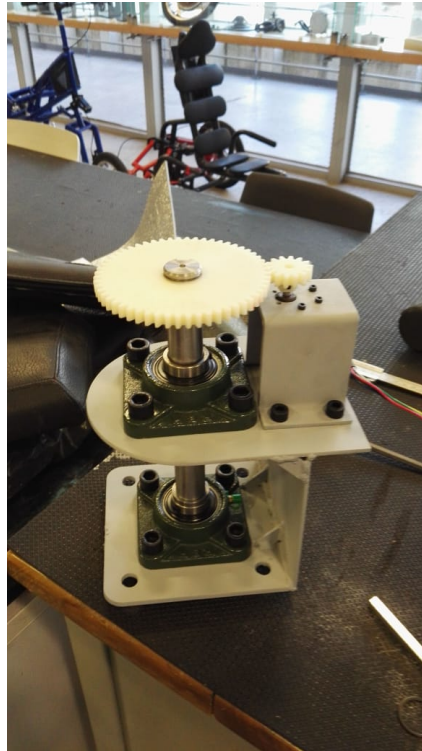


Fig. 10.4.: Rotary arm's main base.

In addition, a long steel square profile was welded to the arm's shaft, which would be in charge of holding the multispectral cameras over the lots. On the arm's end, the cameras had their mounting point. Figure 10.5 shows the assembled rotary arm.



(a) Arm's profile on the main base.



(b) Mounting point of the arm's profile for the multispectral cameras.

Fig. 10.5.: Assembled rotary arm.

Once the arm and its support structure were constructed, they were installed in a *Universidad EAFIT*'s greenhouse. The process and result are illustrated in Figure

10.6.



(a) Terrain leveling.

(b) Arm support structure installation.



(c) Rotary arm fully installed.

Fig. 10.6.: Installation of the rotary arm.

For this part, a support structure for the multi-spectral cameras was built, which would be placed on the previously mentioned mounting point. It consisted of a steel plate with an aluminum bent plate, in which the cameras (protected inside their water proof cases) were installed, and a steel square profile with multiple passing holes welded to that steel plate. Those passing holes were planned to allow height adjustment. The Figure 10.7 shows the rotary arm after installing the multispectral cameras.

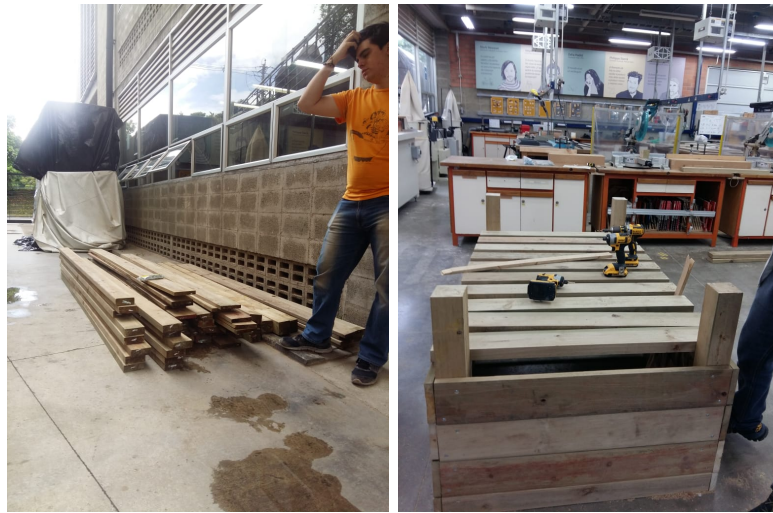


Fig. 10.7.: Installation of the multispectral cameras.

10.1.2 Coffee crop lots

Due to the emulated rain conditions that the prototype would be submitted to, immunized wood was chosen for the construction of the coffee crop lots since it is resistant to moisture. The wood was cut and assembled within the *Universidad EAFIT*'s workshops in order to form the lots. At the bottom of each of them, a mesh in conjunction

with a plastic tarp was installed to contain the soil. The Figure 10.8 illustrates the building process.



(a) Immunized wood.

(b) Crop lot assembly.



(c) Mesh and plastic tarp installation.

Fig. 10.8.: Construction of the crop lots.

After the four lots were built, they were taken to the greenhouse and installed as shown in Figure 10.9. The positions described in Section 9.1.1 guided the installation.



Fig. 10.9.: Crop lots on their respective positions inside the greenhouse.

10.1.3 Rain system

Near the greenhouse, in which the final cyber-physical prototype was assembled, there was a water source. Thus, a hose was connected to it and lead to the crop lots with the purpose of feeding the rain system. The Figure 10.10 presents the result.



(a) Water source.

(b) Hose for feeding the rain system.

Fig. 10.10.: Connection to the water source.

All accessories that conformed the pipeline for the incoming water were installed over immunized wooden blocks. In addition, steel profiles were screwed below the blocks and then, the whole wooden base was buried in order to keep the structured fixed. Finally, as shown in Figure 10.11, five supports were also installed on two adjacent sides of each lot for mounting a wooden L-structure over them. That structure was in charge of serving as a base for three additional hoses with small perforations, which would distribute the water within the lot producing the rain effect.



Fig. 10.11.: Rain system's L-structure.

10.1.4 pH measurement system

Within each lot, a slope was built for the residual water to flow towards a filtering point. The slope was made with soil and an impermeable plastic in conjunction with a plastic mesh was placed on top of it. The Figure 10.12 shows the building process.



Fig. 10.12.: Slope towards the filtering point.

At the filtering point, a hose was connected in order to lead the residual water to the container for pH measurement. The latter was installed on a wooden base and two holes were made on its lid, one for the mentioned hose and the other for the pH-meter. Furthermore, two additional perforations were also made on the container sides for installing the drain hoses. The Figure 10.13 illustrates the assembly of the pH measurement system.



Fig. 10.13.: Assembly of the pH measurement system.

10.1.5 Rotary platform

For the assembly of this system, several acrylic pieces were cut, including the rotary platform, rack, pinion, support for the rack, base and protection for the web-cam, supports for the mini-DC and DC motors and stop for the rotary platform's extension. The Figure 10.14 shows the assembly of both motors to the platform, along with three micro-switches for indicating the extension's position and a set of wheels for smoothing the rotation. In addition, it is noteworthy that a bent aluminum plate was used for keeping the mini-DC motor fixed to its support.



Fig. 10.14.: Motors, micro-switches and wheels on the rotary platform.

For its part, a drawer slide was installed below the rotary platform's extension allowing its movement and therefore, the camera displacement towards the coffee plants, as shown in Figure 10.15. Besides this, the base for the web-cam was placed at one end of the extension and a protective case was built around it using four acrylic walls with the aim of avoiding the entry of water from the rain system. Thereby, not only the camera but also the mini-DC motor would be safe.



(a) Front view of the the web-
cam assembly.

(b) Side view of the the web-
cam assembly.

Fig. 10.15.: Web-cam assembly.

Lastly, for placing the rotary platforms on the lots, steel plates were assembled to the supports of the digital servos. Additionally, four levellers were screwed to the corners of each plate in order to put them underground for fixing the whole structure. As illustrated in Figure 10.16, the rotary platform was finally placed on top of a wooden base to mitigate the terrain's instability.



Fig. 10.16.: Rotary platform placed on a lot.

After following the mentioned procedures to accomplish the subsystems assembly and integrating them into a single system, the general operation was tested and some drawbacks occurred. For instance, the unstable and irregular terrain, on which the rotary platform relied, as seen in Figure 10.16, caused the acrylic surface not to be perpendicular to the Dynamixel's shaft. As a consequence, an over-torque was applied to the latter sometimes, which blocked it for receiving further commands with the purpose of preventing its internal damage. The solution was to put the levelers of the rotary platform on wooden boards in order to give more stability to the structure and try to ensure its rotation. Additionally, the proposed design for the pH measurement system made it difficult to filter the water effectively and prevent soil particles from entering the pH container. Although this system was not modified during the construction of the prototype, a better solution for future work could consist in installing a plastic cover below each lot forming a funnel structure towards

the pH container. Then, before reaching it, it would be useful to put a porous plastic fabric for filtering the residual water from the soil.

10.2 Electronic Prototyping

The Electronic prototyping details how each sensor and actuator was tested, calibrated and integrated with the micro-controllers composing the data collection system. Due to the harsh conditions where the rain factor is continuously present, the components had to have individual connectors for their protection and easy maintenance, as seen in Figure 10.17. Then, in order to provide a hermetic seal, a thick silicon layer was applied to their metal terminals.

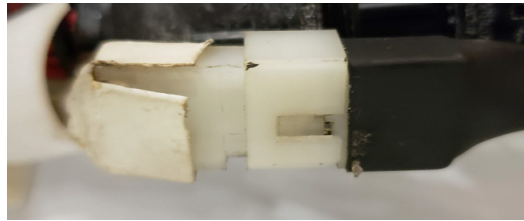


Fig. 10.17.: Example of individual connector assembly.

Then, using Appendix P, two types of signals (analog and digital) and the required circuit board that is adequate by each component could be identified. On the other hand, Table 10.1 aims to test the correct functioning of the interaction between both Arduinos, Raspberry Pi® and the cameras connected to the latter. The purpose of the table was to ensure the integration of all the components related to the data collection system, dividing the analysis into function names, command response and observations from the Arduinos with the single board computer.

Table 10.1.: Checklist to test Raspberry Pi's components.

<i>Commands</i>	<i>Observation</i>	<i>Component</i>
L1		Arduino 1
L2		
L3		
L4		
GD		Arduino 2
RGB_1		Raspberry Pi®
RGB_2		
RGB_3		
RGB_4		
RGN		
RE		

On the other hand, Appendix R shows the protocol applied to test sensors and actuators on the greenhouse. This table is divided into columns of continuity tests for signal and voltage, functioning, and observations based on the operation of each component. Finally, Appendix T demonstrates the final checklist based on the integration's response between both Arduinos and the Raspberry Pi® to commands from the corresponding component (Arduino, RGB or multi-spectral) on the greenhouse. Additionally, it has also an observation column if any anomaly shows up. In the following subsections, the individual steps are listed in order to accomplish the integration of each Arduino and the Raspberry Pi® inside the electrical cabinet.

10.2.1 Electrical cabinet

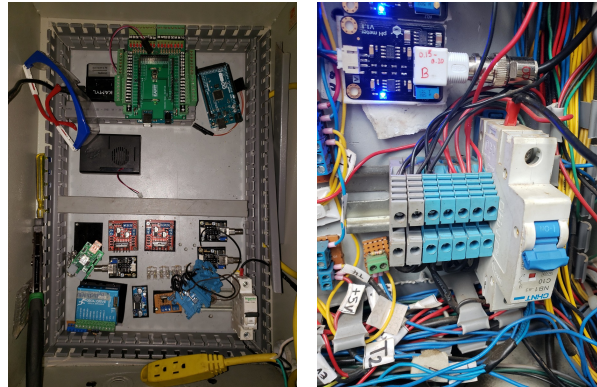
The electrical cabinet is in charge of containing, protecting and distributing the essential components of the data collection system. Due to harsh conditions it is subjected

to, it involves the protection of these components from water, dust and high temperatures. Figure 10.18 shows the final location of the two cabinets aside from the data collection system.



Fig. 10.18.: Electrical cabinets' location.

The first step was dividing the electrical cabinet with plastic rails as stated in Subsection 9.2.4, and as seen in Figure 10.19(a). Additionally, Figure 10.19(b) shows the location of the breaker protection and the block terminals. Throughout these terminals, the 5V and GND are distributed all along the cabinet. In this way, this power supply may distribute the demand voltage to the required components.



(a) First divisions assembly of the electrical cabinet. (b) Detailed view of terminal blocks in the cabinet.

Fig. 10.19.: First steps of the electrical cabinet.

Due to the number of elements and the sole cabinet for the project, a new cabinet was required to store and protect the ATX power supply, as seen in Figure 10.20. As a result, this supply feeds the junction boxes and the corresponding sensors with 5V and GND.



Fig. 10.20.: ATX 5V power source.

10.2.2 Arduino 1

The lot data collection Arduino is the most demanding micro-controller of the two Arduinos, as seen in Subsection 9.2.1. Thus, testing each sensor and actuator for this micro-controller was the first step of the prototyping process. As seen in Figure 10.21, each component was tested using Appendix P, in order to check its correct functioning before being integrated and installed in the greenhouse.

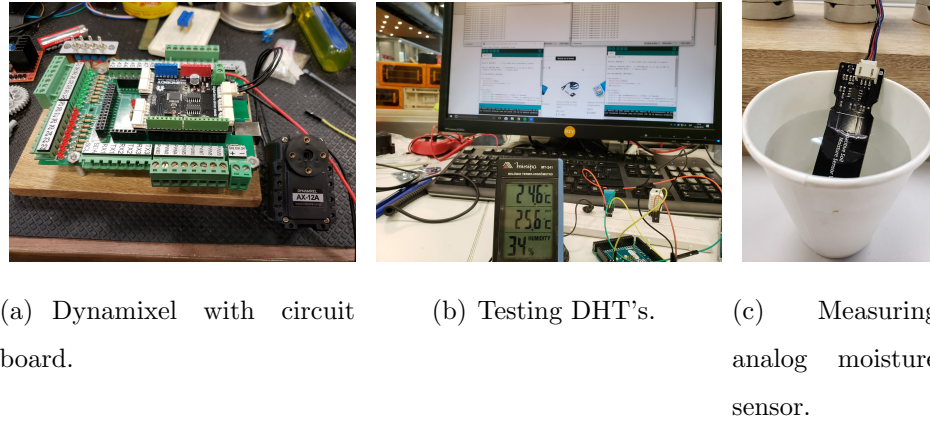


Fig. 10.21.: Individual testing of sensors and actuators from Lot data collection Arduino.

Additionally, Appendix Q details the results of the tests done in the laboratory for all the components handled by the Arduino 1. Afterwards, the 22 AWG cable was cut and then mounted in a sealed connector for each sensor and actuator that would go towards the electronic enclosure (cabinet). However, the sole element that did not need a connector was the Dynamixel. This was due to the fact that the cables have their own connector which was compatible with the circuit board and the Dynamixel. Consequently, each component was placed and put on their corresponding position within the lots. For example, Figure 10.22(b) and Figure 10.22(a) show the final location of the LDR and soil moisture sensors.



(a) Soil moisture sensor. (b) LDR on the rotatory platform.

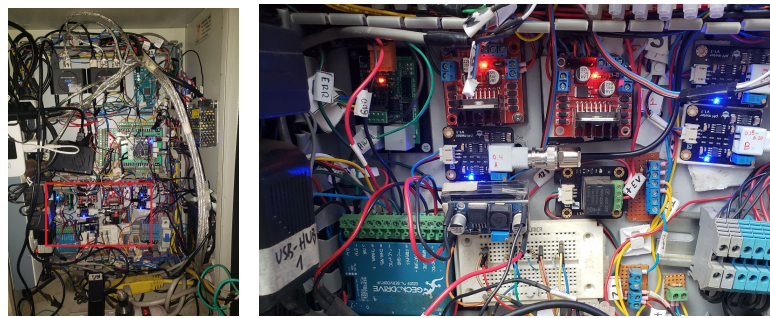
Fig. 10.22.: Sensors in position on the greenhouse.

Subsequently, the cables from all the sensors and actuators from the four lots were guided through an impermeable PVC tube. Then, the 5v and GND power supply cables were connected to some terminal blocks within the junction box, installed aside of the Lot 2 and 3 as seen in Figure 10.23.



Fig. 10.23.: Power supply junction box.

It is noteworthy to clarify that these junction boxes were introduced to avoid the excessive use of cable, so the 5V and GND from the electrical cabinet was distributed to most of the sensors and actuators. Subsequently, the signal cables continued their path through the plastic PVC tubes towards the electrical cabinet, which distributed the incoming cables to their corresponding circuit boards, as seen in Figure 10.24. Moreover, these boards were connected to the Arduino 1 pins as stated before, and as seen in Figure 10.25.

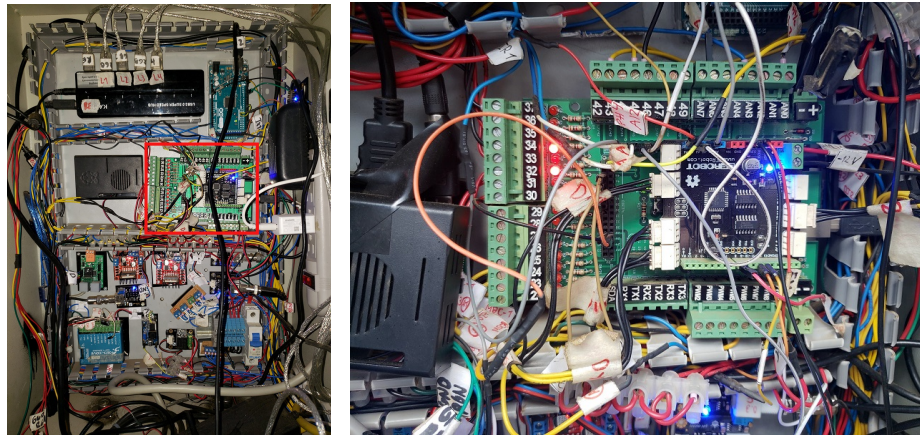


(a) Circuit boards (b) Detailed view of circuit boards for Arduino 1. for Arduino 1. in the cabinet.

Fig. 10.24.: Circuit boards from Arduino 1.

After coupling all sensors and actuators to the corresponding pins of the Arduino, the continuity, voltage, and ground values were checked for each lot component could be reported through the greenhouse checklist from Appendix R. Based on this report, some elements presented a lack of continuity on their signal cables but not on their supply ones. Among these elements of interest, the LDR and soil temperatures sensors were the most important during the final check. Due to the lack of contact within the connector, the LDR connectors were removed and replaced by only two individual wires. This was not the case for the soil temperature ones, because instead of the lack of continuity, the sole circuit board dedicated for the four of them generated a loss of data for two of these sensors, which was due to the lack of continuity of the circuit

board from the signal wire. As a result, it was decided to join the sensors from lot one and two into a single circuit board, and the others from lot three and four into another one. This method of verification (checklist table) let recognize and solve the inconsistency data issue. Although, to prevent future any inconsistent contacts among the Arduino connection, a terminal block shield was coupled to it ensuring connectivity between the sensors and itself, as shown in Figure 10.25(b). Finally, the functions from this Arduino, which can be seen in Table 9.1, were tested with a satisfying result.



(a) Arduino 1 installed on the electrical cabinet. (b) Detailed view of Arduino 1 with terminal blocks on the cabinet.

Fig. 10.25.: Arduino 1 installed on the electrical cabinet.

10.2.3 Arduino 2

The General data collection micro-controller acquires and controls the external variables involving the greenhouse. These components are in charge of controlling and sensing the rain system, rotatory arm, and wind sensor, which were established in Subsection 9.2.2.

The rotatory arm system was tested without any load by controlling the number of pulses given by a specific command. Figure 10.26 demonstrates an example of such testing process. On the other hand, all components were tested and assembled with a sealed connector and the 22 AWG corresponding wire length according to their position in the greenhouse.

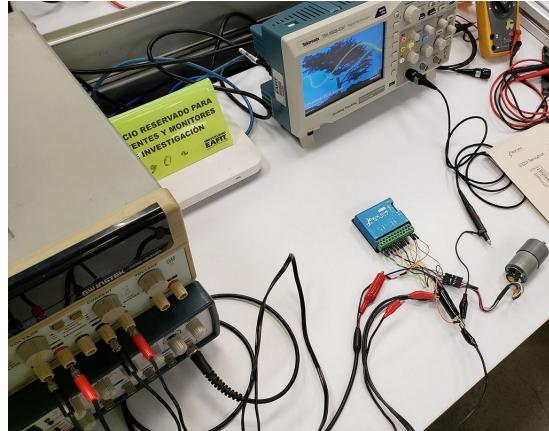
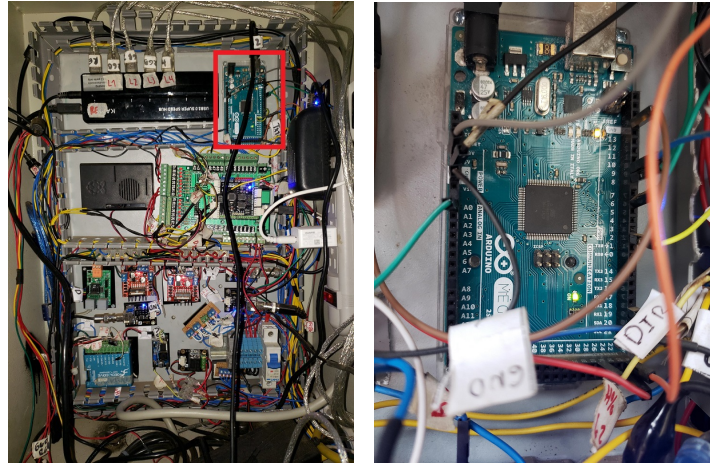


Fig. 10.26.: Testing Pololu (DC-motor) with GECKO driver.

Since all the elements were integrated into the Arduino 2, the laboratory checklist based on Appendix Q determined successfully its correct functioning. As the components of the Arduino 2 passed the table checklist, the next step was moving and installing them on the greenhouse to their established location depending on the subsystem they belong.

The cables from the Pololu[®] (DC-motor), Flow-rate, and Solenoid-valve were guided through a single PVC tube towards the cabinet. Inside the cabinet, the wires, which reached their final destination, were guided and positioned to their corresponding circuit boards as instructed on the Subsection 9.2.2, and as the connections seen in Figure 10.27.

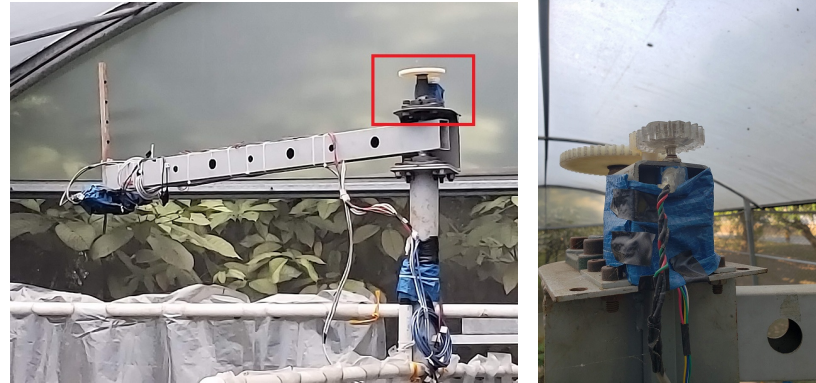


(a) Arduino 2 installed on the electrical cabinet. (b) Detailed view of Arduino 2 in the cabinet.

Fig. 10.27.: Arduino 2 installed on the electrical cabinet.

After connecting the general data collection wires to their respective circuit boards and then the latter to the Arduinos, the cables continuity were checked alongside with voltage 5V, 12V, and GND. The checklist ensured the correct voltage and continuity of the cables coming off the components and its power supply. Based on this result, the next step was testing the functions of the Arduino on the greenhouse, as seen in Table 9.2.

During the testing process, a lack of accuracy of the rotatory arm's precise position was detected, whose origin was due to the given pulses. To address this problem, it was required to integrate a gear with a ten turns potentiometer to the Pololu as seen in Figure 10.28(b). This addition provided a precise feedback position of the rotatory arm for the micro-controller. Besides that, the electromagnetic noise was also present, leading to a replacement of the Pololu®'s AWG cables and the potentiometer normal cables to shielded ones in order to insulate the signals from the noise.



(a) 10 turns potentiometer installed on the (b) Detailed view of 10
rotary arm. turns potentiometer on
the rotatory arm.

Fig. 10.28.: Position of the implemented potentiometer with the rotatory arm.

Finally, the Arduino 2's blueprint was changed due to the potentiometer's addition and its corresponding circuit board. Those changes are reflected in Appendix S.

10.2.4 Raspberry Pi®

Using Appendix T, it was possible to easily identify the issues that any component presented. Employing the commands from Tables 9.3 and 9.4 on the Raspberry Pi®, the data collection system was tested as a whole. However, the data collection routines' test presented problems from PWM signal sent by the Raspberry Pi® in order to change the multi-spectral camera's mode. Thanks to the checklist table, it was determined that the origin of the problem was due to the electromagnetic noise caused by the 110V power surge as stated before. Therefore, it was decided to implement and replace the AWG cables with shielded cables as well, ensuring a insulated signal from the single board computer to the multi-spectral cameras.

Another faced problem during the electronic prototyping was activating the RGB cameras correctly. This problem was caused mainly by the quantity of elements

connected to the USB hub, and that only two of the seven available ports were powered (2A). Hence, two USB hubs had to be added with four constant energized USB ports as seen in Figure 10.29(b).



(a) Overview of the USB hubs' distribution on the electrical cabinet. (b) Detailed view of the four USB 3.0 port hubs.

Fig. 10.29.: Electrical cabinet new distribution with the USB hub's.

Subsequently, these hubs were installed on the electrical cabinet, and then distribute the RGB cameras with the Arduinos on the USB ports from both hubs. Nevertheless, the Raspberry Pi[®] started to block as the test routines continued due to the overheating of the electrical cabinet. Thus, to solve this overheating, a fan towards the CPU was implemented, relieving the high temperature. Given these points, the electronic blueprint of the Raspberry Pi[®] changed to reflect the provided solutions. This final version is detailed in Appendix U.

The successful integration was accomplished as a result of the feedback from the checklist tables in order to overcome the series of drawbacks that put on risk the electronic section. For future work, it is recommended to use shielded cables instead of AWG for all sensor and actuator signals, preventing electromagnetic noise. Like-

wise, it could be considered to implement the PRO version of the pH sensor, which allows to use it directly measure the pH in the ground and not only in liquid, avoiding the construction of a water filter system. As a consequence of the amount of water caused by the rain system, the LDR sensitive layer is affected by corrosion. Hence, it is recommended to implement an encapsulated LDR as seen in Figure 10.30, with the intention of preventing this damage.

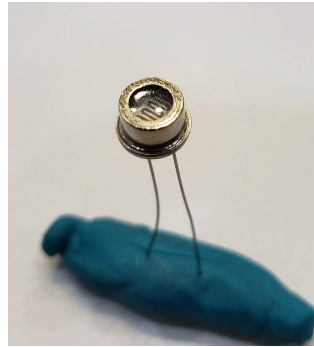


Fig. 10.30.: Encapsulated LDR.

Alternately, to prevent the use of so much wiring, it could be considered to apply a serial communication with the RS485 chip. This component would allow to divide the lot acquisition by nodes, making it scalable and with fewer cables integrated only into one Arduino Mega. However, the most crucial consideration is separating the 100 VAC power supply from the electronic components on a different cabinet from the same as where the micro-controllers and circuit boards are located.

Two factors that caused the high temperature inside of the cabinet were the overheating of the Raspberry Pi[®], and the heat exchanged from the external environment to the metal plate (cabinet). A possible solution to this problem could be applying a fan to extract the cabinet heat to the exterior, but due to the hermetic seal protection from the cabinet, entering air/water particles inside the cabinet was not possible.

Therefore, it is proposed to implement liquid refrigeration on the Raspberry Pi[®], for future work.

10.3 Informatic Prototyping

The Informatic Prototyping shows the implementation process of the final prototype's cybernetic component. It includes the communication between the Raspberry Pi[®] and both Arduinos, the interaction with the cameras, the structuring and transfer of the collected data, the integration with *Thingworx*[®] and the local/remote storage. Furthermore, the preprocessing of the data as well as the generation of the DL model for diagnosing the Coffee Leaf Rust development stage are presented below.

10.3.1 Communication commands Raspberry Pi[®] - Arduinos

The first step in the informatic prototyping consisted of connecting the Raspberry Pi[®] (the master) with the Arduinos (the workers). For that purpose, the *Arduino*[®] *communicator* component was implemented and the whole commands presented in Tables 9.3 and 9.4 were duly tested. The mentioned component was able to establish a Serial connection with both Arduinos, send any command to each of them separately and group the responses of the micro-controllers into single programming objects for their posterior structuring and storage. It is noteworthy that it was necessary to add a delay between the establishment of the connection and the sending of commands since no response was received on the Raspberry Pi[®] otherwise.

In addition, the *Data structurer* component was also implemented for organizing the responses coming from the Arduinos and giving meaning to them i.e. structuring which value belonged to which sensor's or actuator's reading. Thanks to the integration of the mentioned components, it was possible to effectively send commands from the Raspberry Pi[®] to each Arduino[®], trigger its appropriate functions and consolidate the results in an understandable way. It should also be noted that the

integration was tested executing a script that incorporated both components and allowed to manually change which command to send.

10.3.2 Raspberry Pi®

Once the bidirectional communication between the master and the workers was successfully established, several components were implemented and tested to enable the general operation of the data collection system. Then, the Raspberry Pi® was able to use these components in order to orchestrate the routines.

Next, the *Data subdirectory creator* component was implemented for creating a subdirectory within the main data directory using the timestamp at which a data collection routine began. Additionally, it was in charge of generating the proper internal structure, which consisted in one subdirectory for each lot and another one for the general data. This component was tested at different times on different days, resulting in a satisfactory outcome. As an example, Figure 10.31 shows one of these created subdirectories.

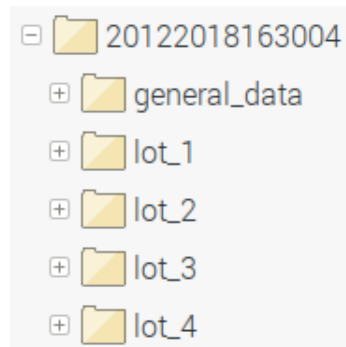


Fig. 10.31.: Internal structure of an automatically created subdirectory.

The next developed component was the *Thingworx® communicator*, which allowed to send data to the IoT platform to update the visualization dashboard. Using the

output of the *Data structurer* component, the *Thingworx*[®] *communicator* was able to add the necessary credentials and format them properly in order to send an update request to the corresponding end-point in *Thingworx*[®] (the available end-points are presented in the next section). As a result, it was proven that it was possible to collect data from the Arduinos and report it immediately to the IoT platform, so the readings could be monitored remotely and in real-time.

The next step consisted in implementing the *Image capturer* component. As the name implies, it was in charge of triggering the correct RGB camera for capturing images of 640x480 pixels on a specified lot and saving the resulting files locally. Besides this, considering that the images of the multi-spectral cameras were captured at different positions according to the lot and that they were also very large (3000x4000 pixels), the mentioned component was able to rotate them to emulate that all of them had been captured on different lots but preserving the position as well as convert them to a more manageable size (640x853 pixels). Then, the *Image capturer* was duly tested. In this regard, Figure 10.32 shows an example of two RGB images captured in the lot 3 whereas the Figure 10.33 illustrates the rotation applied to multi-spectral images from lots 1 and 3.

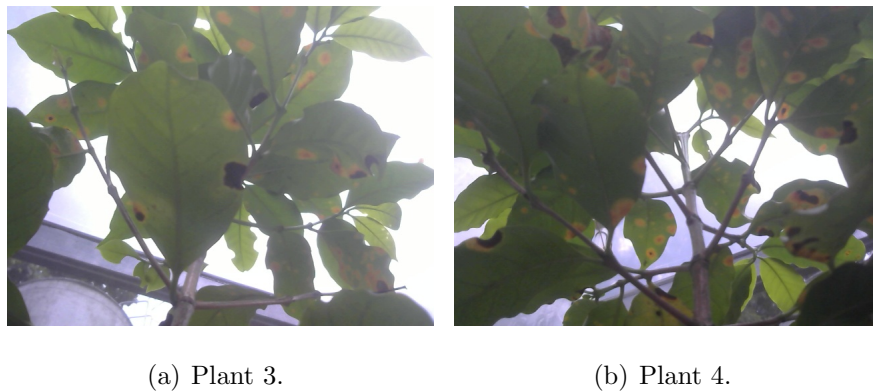


Fig. 10.32.: RGB images captured in the lot 3.

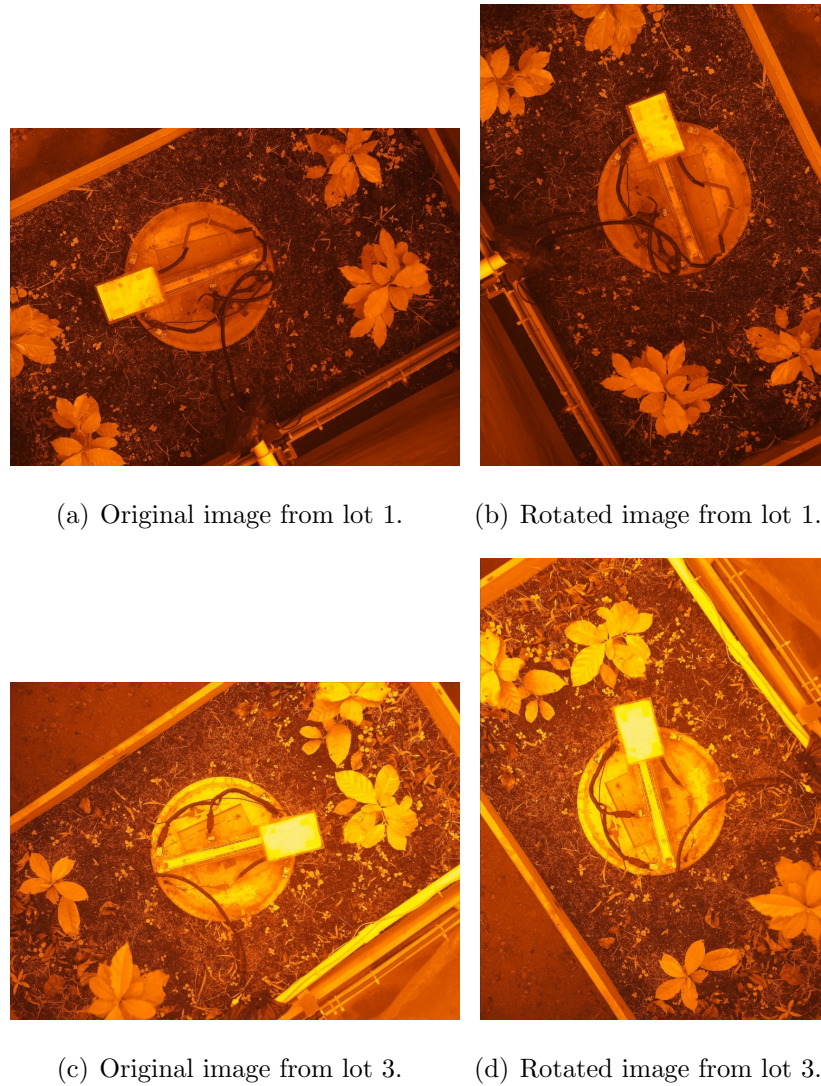


Fig. 10.33.: Original and rotated multi-spectral images from lots 1 and 3.

When the *Image capturer* component was tested, it was found that it was necessary to add a delay between the camera selection and its triggering in order to avoid errors when capturing the images. It is important to mention that the testing was carried out executing a script that incorporated this component and allowed to manually change which camera to trigger.

For the interaction with the multi-spectral cameras, two components were developed, namely *Pulse generator* and *Data copier*. On the one hand, the *Pulse generator* was in charge of generating the necessary pulses for changing the behavior of the cameras i.e. constantly sending pulses to them in order to maintain the communication and to modulate the pulse width anytime for capturing images, or setting the mode in which the cameras were required to be (capture-mode or transfer-mode). Due to possible electromagnetic noise, the component contained also a safety-check to verify the current mode of each camera before taking any action on them. On the other hand, considering that the captured images were saved on the internal storage of each multi-spectral camera, the *Data copier* component was implemented for finding the most recent files on them and copying those files to the Raspberry Pi's local storage for their further transfer. The integration of both components along with the *Image capturer* one allowed to successfully trigger each camera to capture an image, set them to transfer-mode, copy the correct files to the Raspberry Pi®, rotate and resize them properly and finally set the cameras to capture-mode again. Figure 10.34 shows an example of two multi-spectral images on the Raspberry Pi®.



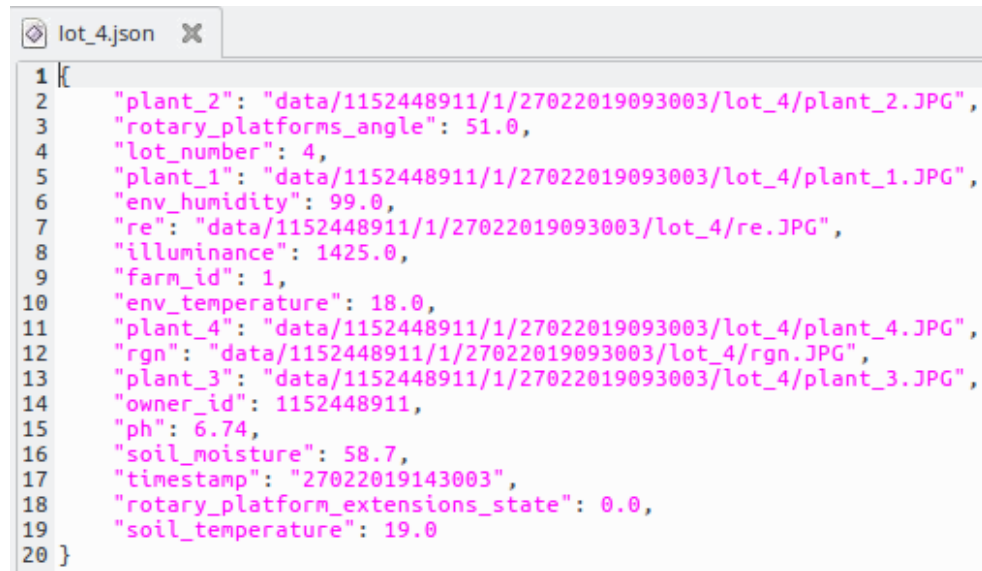
(a) RGN image from lot 1.

(b) RE image from lot 1.

Fig. 10.34.: Multi-spectral images on the Raspberry Pi®.

The mentioned integration was tested executing a script that incorporated the three (3) mentioned components and allowed them to interact with each camera separately. It should also be noted that it was required to add delays between the changes of the camera's mode and also between the capture/transfer actions in order to avoid unexpected behaviors and data loss.

The next step consisted of implementing the *Document local writer* component. Its goal was to add the respective timestamp to the output of the *Data structurer* component and store it as a JSON file (aka document) on a specified location within the Raspberry Pi's local storage. The resulting document could contain either the collected field sensor data or the general sensor data from the scale coffee crop depending on which command was sent through the *Arduino*[®] *communicator* component. Whenever the document contained field sensor data, if the corresponding data collection routine was executed with sunlight, it included also the paths to the captured images. The *Document local writer*'s behavior was then successfully proven, where the Figure 10.35 illustrates an example of two documents belonging to the same routine stored in the Raspberry Pi[®].

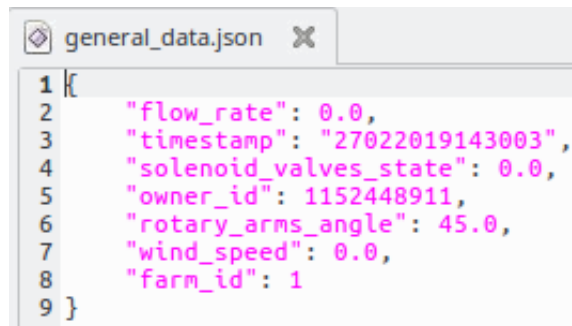


```

1 {
2   "plant_2": "data/1152448911/1/27022019093003/lot_4/plant_2.JPG",
3   "rotary_platforms_angle": 51.0,
4   "lot_number": 4,
5   "plant_1": "data/1152448911/1/27022019093003/lot_4/plant_1.JPG",
6   "env_humidity": 99.0,
7   "re": "data/1152448911/1/27022019093003/lot_4/re.JPG",
8   "illuminance": 1425.0,
9   "farm_id": 1,
10  "env_temperature": 18.0,
11  "plant_4": "data/1152448911/1/27022019093003/lot_4/plant_4.JPG",
12  "rgn": "data/1152448911/1/27022019093003/lot_4/rgn.JPG",
13  "plant_3": "data/1152448911/1/27022019093003/lot_4/plant_3.JPG",
14  "owner_id": 1152448911,
15  "ph": 6.74,
16  "soil_moisture": 58.7,
17  "timestamp": "27022019143003",
18  "rotary_platform_extensions_state": 0.0,
19  "soil_temperature": 19.0
20 }

```

(a) Field sensor data and paths to the images from lot 4.



```

1 {
2   "flow_rate": 0.0,
3   "timestamp": "27022019143003",
4   "solenoid_valves_state": 0.0,
5   "owner_id": 1152448911,
6   "rotary_arms_angle": 45.0,
7   "wind_speed": 0.0,
8   "farm_id": 1
9 }

```

(b) General data.

Fig. 10.35.: Documents from the same routine stored in the Raspberry Pi®.

Additionally, an important developed component was the *VPN connector*, which was in charge of checking the status of the VPN connection before trying to store the data remotely, since its establishment was mandatory for reaching the ADC over the Internet. Moreover, it was also used to verify the status of the Internet connection every time the interaction with the *Thingworx® communicator* occurred or once the data was ready to be transferred to the remote server. Subsequently, the *VPN connector* was duly tested and the result was its capability of validate the connectivity

requirements in a satisfactory manner.

The last two components that were implemented for completing a data collection routine were the *MongoDB[®] communicator* and the *Data uploader* components. The former was able to insert the documents containing the collected data into a *MongoDB[®]* instance within the ADC as well as managing the respective connection. For its part, the latter was in charge of replicating each data subdirectory on the Raspberry Pi's local storage into the ADC's File system i.e. creating the same subdirectory's internal structure and uploading every single file to its corresponding place via SFTP. It is relevant to mention that both components had additional logic for retrying the data transfer whenever an error occurred during the upload and also a timeout for avoiding an infinite transaction. The *MongoDB[®] communicator* as well as the *Data uploader* were then tested under an intermittent VPN connection using a manually created subdirectory, proving that both of them made their best effort within the timeout restriction to upload the data to the remote server i.e. they followed the retry protocol as expected and transferred as many files as possible.

Finally, a master script called `data_collection.py` was implemented in order to incorporate the whole mentioned components and orchestrate the steps to complete a data collection routine. This script had one schedule for collecting data and another one for the rain days and times. Furthermore, it was composed of five functions (described in Table 10.2), which were in charge of invoking each component, whenever it was required during the process, in order to cause the desired behavior on the cyber-physical system.

Table 10.2.: Functions of the `data_collection.py` script.

Function	Description
<code>collect_data_for_visualization</code>	Collects field and general sensor data (without using the cameras) and sends it to Thingworx for real-time monitoring and visualization
<code>collect_partial_data_for_storage</code>	Collects field and general sensor data (without using the cameras), stores it locally and transfers it to the remote server
<code>collect_total_data_for_storage</code>	Collects field and general sensor data (also using the cameras), stores it locally and transfers it to the remote server
<code>begin_rain</code>	Starts the rain on the scale coffee crop
<code>end_rain</code>	Finishes the rain on the scale coffee crop

In that sense, the `data_collection.py` script was duly tested and the result showed that it permanently scanned over its schedules and successfully executed the appropriate functions at the right time. It should also be noted that a complete routine had to generate twenty nine files (one JSON file for the general data, one JSON file per lot for the field sensor data, two multi-spectral images per lot and four RGB images per lot). For instance, Figure 10.36 presents an example of the lot 3 directory's content on the Raspberry Pi® after one data collection routine concluded.



Fig. 10.36.: Lot 3 directory's content after a data collection routine.

10.3.3 Thingworx®

The data visualization dashboard was developed using *Thingworx*® according to the presented design. Its appearance was preserved and it was accessible through an URL

with user-password authentication. Every used widget had an unique identifier to target each of them separately, for updating their values. In that way, it was proven that one or more readings of the scale coffee crop could be sent through a HTTP-POST request from the Raspberry Pi[®] to *Thingworx*[®] and that only the expected values got modified on the visualization dashboard. Furthermore, each time a request was received, the respective timestamp was logged with the purpose of showing it on the visualization and looking to enhance the monitoring. Besides this, the whole dashboard was automatically updated every second with the last recorded value for each variable.

In order to correctly update the data, five end-points were created and exposed. Each of them was in charge of updating the values within a specific block on the visualization dashboard. The Table 10.3 describes the input and output of the five mentioned end-points.

Table 10.3.: End-points for updating the data in the visualization dashboard.

End-point	Description	Input	Output
update_lot_1	Updates the values corresponding to the lot 1 block	JSON object with the values for the update	JSON object with the updated values
update_lot_2	Updates the values corresponding to the lot 2 block	JSON object with the values for the update	JSON object with the updated values
update_lot_3	Updates the values corresponding to the lot 3 block	JSON object with the values for the update	JSON object with the updated values
update_lot_4	Updates the values corresponding to the lot 4 block	JSON object with the values for the update	JSON object with the updated values
update_general_data	Updates the values corresponding to the general data block	JSON object with the values for the update	JSON object with the updated values

10.3.4 ADC

Once a data collection routine finished on the Raspberry Pi[®], the procedure to verify that the result was satisfactory consisted of checking the ADC's File system and looking for the corresponding data subdirectory. Thus, the file number inside of it had to match the number of uploaded files by the *Data uploader* component and those files should be located in exactly the same place, as if they were in the local subdirectory. Additionally, the files had to be identical with respect to their name and content. The mentioned verification was carried out during several days after each data collection routine and it was proven that the data transfer to the remote server was successful. Figure 10.37 illustrates a comparison example between the content of a local subdirectory of the lot 3 and its remote replica.



(a) Content of the local subdirectory.



(b) Content of the remote subdirectory.

Fig. 10.37.: Content comparison of the same subdirectory locally and remotely.

Regarding the data processing phase, several informatic components and scripts were implemented in order to generate the diagnostic model based on the collected data. In that sense, the ADC's File system was first examined, then all directories containing lot data were extracted and manually assigned to their respective label directory. In total, 1,580 directories were classified by the Biology team with the *label 0*, no directories with the *label 1*, 121 directories with the *label 2*, 195 directories with the *label 3* and 248 directories with the *label 4*. Besides, an extra **rgb_images** directory containing five sub-directories (ranging from 0 to 4) was created within every classified lot data directory that had RGB images. Thereby, those images could be moved to their corresponding label within each lot data directory. At this point, all images were also visually depurated in order to remove non-valuable content.

Additionally, a script called **cleanStructure.py** was developed and executed in order to eliminate the irrelevant JSON files (those with missing values and outliers) as well as the lot data directories that ended up with no content following that elimination. After the execution of the script, the *label 0* directory contained 860 items, the *label 1* directory 0 items, the *label 2* directory 67 items, the *label 3* directory 108 items and the *label 4* directory 135 items.

The resulting label directories were used for splitting the data into train and test sets. To achieve this, a script named **dataSplitter.py** was used, which allowed to shuffle the lot data directories within each label directory, count the elements per file type, calculate the number of files of a specific type that should be included in the test set and, according to this number, copy a fraction of the lot data directories to another location for evaluation purposes. The execution of the mentioned script successfully created the train and test sets, which contained approximately 75 % and 25 % of the whole data respectively. It is noteworthy that the *label 1* directory was excluded from both sets since it was empty. However, it was kept inside the **rgb_images** directories

because some RGB files were classified under that label.

Subsequently, the content of the lot data directories within the train set was divided according to the file type. A script called `createStructure.py` was developed and executed in order to achieve that. As a result, four different directories were generated under the names `json_data`, `rgb_data`, `rgn_data` and `re_data`, each of them storing all files of a specific type. In total, the `json_data` directory was composed by 672 files, the `rgb_data` directory by 2,192 files, the `rgn_data` directory by 603 files and the `re_data` directory by 641 files. Then, the content of those directories was preprocessed using a script called `preProcessing.py`. It allowed to clean the files by locating the plants and removing the background, in the case of images, and eliminating unnecessary keys, in the case of documents. Figure 10.38 illustrates an example of four preprocessed files.

```

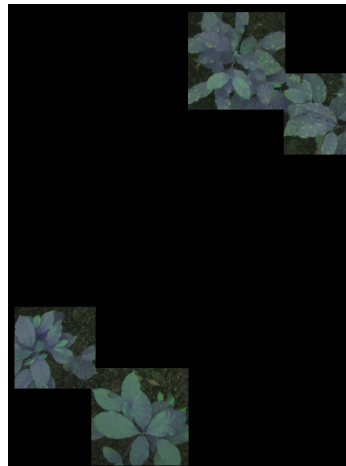
1 [{
2   "ph": 6.81,
3   "soil_temperature": 21.5,
4   "soil_moisture": 61.59,
5   "illuminance": 1217.0,
6   "env_temperature": 24.0,
7   "env_humidity": 99.0
8 }]

```

(a) Preprocessed JSON file.



(b) Preprocessed RGB file.



(c) Preprocessed RGN file.



(d) Preprocessed RE file.

Fig. 10.38.: Example of preprocessed JSON, RGB, RGN and RE files.

After the mentioned directories were duly preprocessed, a script called `data_structuring.py`, along with four components named *JSON data structurer*, *RGB data structurer*, *RGN data structurer* and *RE data structurer*, was used to produce one file per directory containing the data for the generation of the respective sub-model. The files were produced with the help of the *pickle* module, which implements binary protocols for serializing and de-serializing a *Python* programming object. In that sense, the `data_structuring.py` script was in charge of taking each directory, loading its raw files into memory, resizing them to reduce the spatial complexity (in the case of images), structuring them and their labels into two separate suitable pro-

gramming objects (*feature_data* and *label_data*), permuting both of them, normalizing the *feature_data* and, finally, generating a *pickle* file containing the resulting objects. Thereby, the input for each sub-model was successfully created.

Once the input data was structured, a script called `sub_model_generator.py` was implemented in order to generate one sub-model for each file type. To this end and considering that the sub-models were meant to be created individually, four components named *JSON sub-model generator*, *RGB sub-model generator*, *RGN sub-model generator* and *RE sub-model generator* were also developed. Each of those components allowed to load the programming objects contained in the respective *pickle* files into memory, calculate the class weights (i.e. a weight for each label according to its occurrence on the data to account for label imbalance during training), set the sub-model’s architecture, create a hyperparameters grid and find the best estimator, including the employed hyperparameters and the performance reached, using Grid search with cross-validation. Lastly, the corresponding component saved the generated sub-model (best estimator) into a *h5* file, which is a format designed to store and organize large amounts of data. Appendix V summarizes the architectures of the sub-models whereas the Table 10.4 shows the selected hyperparameters and obtained F_1 -score for each of them.

Table 10.4.: Hyperparameters and F_1 -score for each generated sub-model.

Sub-model	Batch size	Epochs	Kernel initializer	Activation	Rate	Optimizer	F_1 -score
JSON	16	20	normal	relu	0.4	adam	0.651319
RGB	16	6	glorot_uniform	relu	0.4	adam	0.948832
RGN	32	9	glorot_uniform	elu	0.3	adam	0.927606
RE	16	6	normal	relu	0.4	adam	0.878447

The last step to complete the informatic prototyping consisted of integrating the four presented sub-models and evaluating the composite model i.e. diagnosing the Cof-

fee Leaf Rust development stage through it, creating a comparative table with the results and calculating the model's performance. For that purpose, a script called `model_evaluator.py` was implemented. It was in charge of loading the sub-models into memory, iterating over the whole test set, taking each lot data directory within it, dividing the contained files according to their type and preprocessing them, resizing them to reduce the spatial complexity (in the case of images), normalizing and structuring each file according to the sub-models' expected input and sending them to its corresponding sub-model for getting a prediction. Besides, the script allowed to gather the four predicted labels and calculate their rounded weighted average, since the generated sub-models presented different performances for diagnosing the Coffee Leaf Rust development stage. The Table 10.5 shows the weights for the predictions of each sub-model, which were determined as the proportion of each F_1 -score in Table 10.4 with respect to the sum of all F_1 -scores.

Table 10.5.: Weights for the predictions of each sub-model.

Sub-model	Weight for predictions
JSON	0.191215
RGB	0.278559
RGN	0.272328
RE	0.257896

In addition, the `model_evaluator.py` script was also capable to record the predicted label and the actual development stage for each lot data directory during its execution. Finally, it successfully generated a Comma-separated values (CSV) file containing the comparative results for the statistical analysis and calculated the F_1 -scores reached by the individual sub-models as well as the composite model on the test set.

The main drawbacks during the informatic prototyping, with respect to the data collection phase, were related with timing and they were solved adding delays between

the actions triggered by the Raspberry Pi® as well as to the operation of the peripheral devices. It is also important to mention that the Raspberry Pi® froze or even restarted on several occasions when it reached temperatures over 56°C, mainly due to the environmental conditions inside the greenhouse. As a result, its normal behavior was altered and some data was lost or not transferred. Thus, it was sometimes necessary to go to the greenhouse and capture missing RGB images again or extract missing multi-spectral images manually from the cameras' internal storage. For future work, it is recommended to use liquid cooling on the electrical cabinet in order to mitigate unexpected behaviors. In addition, it is noteworthy that, for managing the communication, it is advantageous not to use the name of the port to which a peripheral device is connected but the device's unique identifier. The reason is that the same device can be assigned to a different port when the Raspberry Pi® overheats whereas its unique identifier is permanent.

On the other hand, some problems arose during the informatic prototyping regarding the data processing phase. At the beginning, the proposed hyperparameters grid for each sub-model was very wide, which resulted in more than 470 possible combinations i.e. more than 470 estimators needed to be fitted for the generation of a single sub-model. The solution was to reduce the grid's size in order to include just the most relevant values for the most significant hyperparameters. Therefore, the time and spatial complexity were decreased in a satisfactory manner. Furthermore, it was necessary to clear the *Keras* session after fitting each estimator to free up GPU-memory, because the process for generating the sub-models was being unexpectedly killed due to an Out-of-Memory (OOM) error. For future work, it is recommended to allow the data collection phase to last longer considering that this implies creating a larger dataset, which is always a desirable condition for DL models.

10.4 Integration

For diagnosing the Coffee Leaf Rust development stage, it was necessary to collect data from the scale coffee crop in the first instance and then, process it. In that sense, the integration of the mechanical, electronic and part of the informatic prototyping allowed to conclude the data collection phase with the construction of the cyber-physical functional prototype, as seen in Figure 10.39. At the moment, it collects data automatically at certain times during the day and transfers it to the ADC in order to feed the crop data-set continuously. However, it works automatically to 90 % of its full capacity, considering that it is necessary to reconnect the Dynamixels manually when they are blocked as well as restart the system when the Raspberry Pi[®] reaches very high temperatures and freezes.



Fig. 10.39.: Built cyber-physical prototype for data collection.

For its part, during the data collection period, three pH sensors had to be replaced because the residual water in which they reside damaged the sensitive membrane of

their internal electrode. Besides this, three LDRs and four moisture sensors had to be replaced as well due to the difficult conditions the sensors were in. This could have been partly caused due to the lack of proper insulation, which could have negatively affected their internal circuits and sensitive surfaces.

Moreover, the height of the rotary arm's support structure played an essential role since it caused the elimination of one micro-switch per rotary platform (the furthest from the "home" position). It was found that, using those micro-switches, the plants would be too far from the lot's center, causing the height of the rotary arm's support structure to be insufficient for capturing images of the entire lot with the multi-spectral cameras. Thus, it was decided to bring the plants closer and not to use the mentioned micro-switches. For future work, it is essential to first determine the necessary height, from which the cameras can cover the entire lot, with the aim of not having to affect the position of the plants. As a final point, when the electronic components and the power sources are defined, it is recommended to obtain an electrical cabinet of the right size and divide the power from the signals.

The next section concerns the data processing phase. The obtained results in terms of relevant modifications, employed data for training and evaluating, comparative table for the statistical analysis and performance reached by the diagnostic model are exposed in it.

10.5 Results

On the one hand, it is essential to mention that, during the data collection phase, the Biology team had to replace 12 coffee plants of the scale crop with external diseased ones because the inoculation was having no effect after two months (all plants stayed in the 0 development stage). Thus, for example, it was not possible to classify any lot with the *label 1* since only one of the new plants presented the Coffee Leaf Rust on

that development stage. In addition, this justifies the dominant presence of lot data directories inside the *label 0* directory.

On the other hand, the train set used for fitting the sub-models was composed by 968 directories. In total, they contained 672 JSON files, 2,192 RGB files, 603 RGN files and 641 RE files. Besides, the test set employed for the composite model evaluation comprised 202 lot data directories. Finally, after evaluating the diagnosis of the Coffee Leaf Rust development stage in the Colombian *Caturra* variety by means of the created DL model, a comparative table, among with a performance table, was successfully generated. The Appendix W shows the comparative table for the statistical analysis whereas the Table 10.6 presents the definitive F_1 -score reached by each sub-model and the composite model.

Table 10.6.: F_1 -score reached by the individual sub-models and the composite model.

Model	F_1 -score
JSON	0.569573
RGB	0.920034
RGN	0.945806
RE	0.944209
Composite	0.775278

11. ANALYSIS OF RESULTS

This chapter presents the statistical analysis of the results regarding the performance evaluation of the diagnostic model. It was carried out using the comparative table generated by the `model_evaluator.py` script and the software *R 3.5.2*. The purpose of the analysis was to determine whether there was a significant difference in the mean Coffee Leaf Rust development stage diagnosed with a visual inspection and by means of the proposed technological integration. The outcome was relevant to get the necessary statistical support for answering the research question.

The comparative table contained 202 observations with the corresponding diagnosed development stage for both treatments. The Figure 11.1 shows the box-plot chart describing the measurements. The x-axis contains the two treatments (“visual inspection” and “technological integration”) whereas the Coffee Leaf Rust development stage is presented on the y-axis. The graphical similarity on the data distribution of each treatment suggests a possible similarity on the means of the response variable. For assessing this condition and making a decision with respect to the hypothesis, an ANOVA was executed.

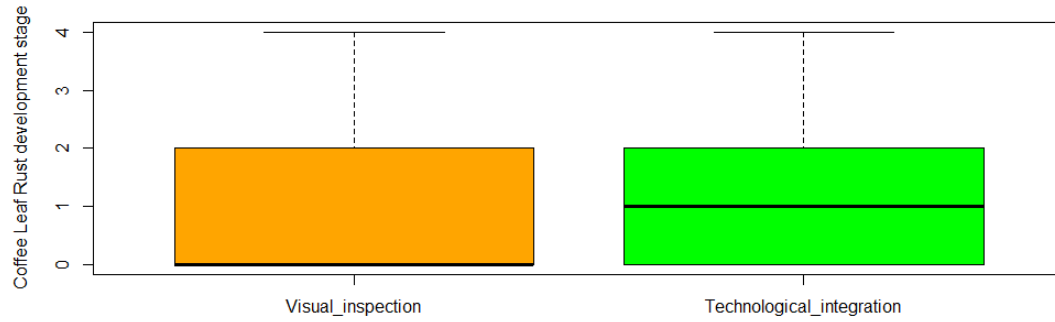


Fig. 11.1.: Data distribution of the observations for both treatments.

The results of the ANOVA can be seen in Table 11.1. The obtained *p-value* for the treatments factor was 0.231. This value is greater than the set significance ($\alpha = 0.1$), which means that there is not sufficient evidence for rejecting the null hypothesis. Thus, it is concluded, with 90 % confidence, that there is no statistical significant difference between the diagnosis of the Coffee Leaf Rust development stage made through the visual inspection and by means of the technological integration. This indicates that both methods are significantly similar for diagnosing the disease.

Table 11.1.: ANOVA table of the statistical analysis.

	Df	Sum Sq	Mean Sq	F value	<i>Pr(>F)</i>
Treatments	1	2.7	2.696	1.437	0.231
Residuals	402	753.9	1.875		

This study demonstrated the possibility of diagnosing the Coffee Leaf Rust development stage in the Colombian *Caturra* variety, with significant performance, through the integration of RS, WSN and DL. The analysis of results allowed to obtain statistical evidence for supporting the research hypothesis. In that sense, the outcome suggests that a technological integration could contribute to the protection of the

phytosanitary status of coffee crops since it showed potential to complement the traditional visual inspections towards the diagnosis of the most economically limiting disease for the Colombian coffee production.

12. CONCLUSIONS

This section analyses the fulfillment of the general and specific objectives of the research project, explains its contribution and compares the research hypothesis with the actual obtained results. The most relevant conclusions are presented below.

The integration of RS, WSN and DL within the framework of this study successfully allowed to evaluate to what extent it is possible to diagnose the Coffee Leaf Rust development stage in the Colombian *Caturra* variety. To this end, the most relevant information obtained in ten interviews with coffee experts was consolidated, the knowledge about the study context and the Coffee Leaf Rust was detailed and the repercussions of the disease in the Colombian coffee growing industry was identified. In addition, the theoretical and conceptual framework as well as the state of the art were reviewed, creative design sessions were carried out for defining the most useful technological integration of RS and WSN, and a functional prototype, that automatically collects data in the field and transfers it to a remote server over the Internet, was built. Besides, a diagnostic model using DL was implemented based on part of the stored data and it successfully allowed to evaluate the Coffee Leaf Rust development stage in unknown field data.

It should also be noted that the time constraints did not allow to create a simple user interface for visualizing the results of the Coffee Leaf Rust development stage diagnosis using the generated DL model. Moreover, this restriction explains why the performance evaluation could not be made on a real coffee crop. Nevertheless, it was possible to compare the results of the diagnostic model with those of a traditional visual inspection using data from the same scale coffee crop. As a result, the model's performance, in terms of the calculated development stage of the disease, could be

measured in a satisfactory manner.

The motivation of this research project was to contribute to the rural development through technological innovation in order to strengthen the profitability of Colombian coffee growers. Considering that the country has the potential, in terms of environmental conditions and diverse ecosystems, to generate a giant portfolio of exotic products that would be better valued in the specialty coffee market, this Master thesis evaluated, with empirical evidence, a technological approach that pretends to facilitate the diagnosis and mitigate the risks of one of the most economically limiting diseases for coffee production. In that sense, the proposed technological integration could positively impact the rural sector since those innovations promote investments in infrastructure, which are crucial to empower the rural community and improve their living standards and activities concerning progress, productivity and income generation.

The obtained *p-value* in the analysis of results was 0.231, which helps to determine, with 90 % confidence, that the visual inspection and the technological integration did not present a statistically significant difference regarding their diagnosis of the Coffee Leaf Rust development stage. Thus, it can be said that the assessment of the disease leads to a similar outcome using any of both methods, which suggests that the obtained results support the research hypothesis. Finally, it can be asserted that through the integration of RS, WSN and DL, it is possible to diagnose the Coffee Leaf Rust development stage in the Colombian *Caturra* variety with a F_1 -score of 0.775278. This value indicates that, on average, the diagnostic model is very good in terms of the certainty and usefulness of its diagnosis.

13. FUTURE WORK

During the physical integration process of the mechanical and electronic sub-systems, as well as while testing the data collection routines, some aspects were noticed, which could be considered for future work. For avoiding the blockage of the digital servos, a solid base could be built for fixing the rotary platform. Additionally, in order to better distribute each electronic component and mitigate errors related to electromagnetic noise, it is necessary to use a larger cabinet in which even the power supplies fit, separate control signals from power, add a line filter and incorporate an Uninterruptible Power Supply (UPS) against blackouts. Besides, it is recommended to wire all electronic components using shielded cables. A liquid cooling system for the Raspberry Pi[®] could be also considered to prevent it from freezing and blocking its execution due to overheating. Furthermore, the use of individual field sensors per plant may be more valuable, in terms of data, for the generation of a diagnostic model that reaches a higher performance.

Regarding the data processing phase, a further extension of this research project could include the implementation of a simple user interface for visualizing the diagnosis of the Coffee Leaf Rust development stage through the generated DL model in order to better illustrate the results to a coffee grower. In addition, the technological integration could be scaled to a real context, taking into account the major challenges that this adaptation implies, with the objective of increasing the validity of the study. Finally, the reached F_1 -scores on the test set, which showed that the sub-models based on images presented a higher performance than the JSON sub-model, suggest to reconsider the composite model for future work and focus all efforts on improving the collection and processing of just multi-spectral data.

REFERENCES

REFERENCES

- [1] National Coffee Association USA, “The Influence of Coffee Around the World,” 2015. [Online]. Available: <https://nationalcoffeeblog.org/2015/06/15/the-influence-of-coffee-around-the-world/>
- [2] SCAA, “SCAA Protocols Cupping Specialty Coffee,” *Specialty Coffee Association of America*, 2015. [Online]. Available: <http://www.scaa.org/?page=resources&d=coffee-protocols>
- [3] CropLife Latin America, “Roya del cafeto.” [Online]. Available: <https://www.croplifela.org/es/plagas/listado-de-plagas/roya-del-cafeto>
- [4] The Observatory of Economic Complexity (OEC), “Colombia (COL) Exports, Imports, and Trade Partners,” 2016. [Online]. Available: <https://atlas.media.mit.edu/en/profile/country/col/>
- [5] Federación Nacional de Cafeteros, “Estadísticas Historicas.” [Online]. Available: https://www.federaciondecafeteros.org/caficultores/es/quienes_somos/119_estadisticas_historicas/
- [6] Conpes 3850, “FONDO COLOMBIA EN PAZ,” Conpes, Bogotá, Tech. Rep., 2015.
- [7] C. Rivillas, C. Serna, M. Cristancho, and A. Gaitan, “La roya del cafeto en colombia: Impacto manejo y costos del control,” *Avances Tecnicos Cenicafe*, 2011.
- [8] K. Mollazade, M. Omid, F. A. Tab, and S. S. Mohtasebi, “Principles and applications of light backscattering imaging in quality evaluation of agro-food products: a review,” *Food and Bioprocess Technology*, vol. 5, no. 5, pp. 1465–1485, Jul 2012. [Online]. Available: <https://doi.org/10.1007/s11947-012-0821-x>
- [9] Revista Dinero, “Oportunidades para el café de Colombia,” 2014. [Online]. Available: <https://www.dinero.com/pais/articulo/oportunidades-para-cafe-colombia/201878>
- [10] Colciencias, “CONVOCATORIA ECOSISTEMA CIENTÍFICO PARA LA CONFORMACIÓN DE UN BANCO DE PROGRAMAS DE I+D+i ELEGIBLES QUE CONTRIBUYAN AL MEJORAMIENTO DE LA CALIDAD DE LAS INSTITUCIONES DE EDUCACIÓN SUPERIOR COLOMBIANAS - 2017.” Colciencias, Bogotá, Tech. Rep., 2017.
- [11] Departamento Administrativo Nacional de Estadísticas de Colombia (DANE), “Censo Nacional Agropecuario Décima Entrega Resultados - 2014,” *Censo Nacional Agropecuario 2014*, pp. 1–43, 2014.

- [12] Colciencias, “CONVOCATORIA NACIONAL JÓVENES INVESTIGADORES E INNOVADORES POR LA PAZ 2017,” Colciencias, Bogotá, Tech. Rep., 2017.
- [13] A. Bárcena, S. Ejecutiva, A. Prado, and S. E. Adjunto, *Agenda 2030 y los Objetivos de Desarrollo Sostenible Una oportunidad para América Latina y el Caribe*. Santiago de Chile: Publicación de las Naciones Unidas, 2016. [Online]. Available: <http://www.sela.org/media/2262361/agenda-2030-y-los-objetivos-de-desarrollo-sostenible.pdf>
- [14] H. Imre, “Comparison of three methodological approaches of design research,” *Guidelines for a Decision Support Method Adapted to NPD Processes*, 2007.
- [15] G. Pahl, K. Wallace, L. Blessing, W. Beitz, and F. Bauert, *Engineering Design: A Systematic Approach*. Springer London, 2013. [Online]. Available: <https://books.google.com.co/books?id=4uvSBwAAQBAJ>
- [16] “Rainforest alliance certified coffee,” Sep 2014. [Online]. Available: <https://www.rainforest-alliance.org/articles/rainforest-alliance-certified-coffee>
- [17] U. R. Alliance, “The label and program for sustainable farming,” 2017. [Online]. Available: <https://utz.org/>
- [18] Comité Nacional de Cafeteros, “Resolución Número 5 de 2002,” 2002. [Online]. Available: <http://www.cafedecolombia.com/static/files/Resolucion5de2002%28CalidadesExportacion%29.pdf>
- [19] S. I. Mussatto, E. M. Machado, S. Martins, and J. A. Teixeira, “Production, composition, and application of coffee and its industrial residues,” *Food and Bioprocess Technology*, vol. 4, no. 5, p. 661, 2011.
- [20] H. Etienne, “Somatic embryogenesis protocol: coffee (*coffea arabica* l. and *c. canephora* p.),” in *Protocol for somatic embryogenesis in woody plants*. Springer, 2005, pp. 167–179.
- [21] C. F. du Café, “Café a la decouverte du cafe,” *Adexquation Publicite, Paris*, 1997.
- [22] J. Arcila, F. FARFAN, A. Moreno, L. Salazar, and E. Hincapié, “Sistemas de producción de café en colombia,” 2007.
- [23] D. Barjolle, X. F. Quiñones-Ruiz, M. Bagal, and H. Comoé, “The role of the state for geographical indications of coffee: Case studies from colombia and kenya,” *World Development*, vol. 98, pp. 105–119, 2017.
- [24] X. Quiñones-Ruiz, M. Penker, C. Vogl, and L. Samper-Gartner, “Can origin labels re-shape relationships along international supply chains?—the case of café de colombia,” *International Journal of the Commons*, vol. 9, no. 1, 2015.
- [25] J. Arcila, F. Farfan, A. Moreno, L. Salazar, and E. Hincapié, “Crecimiento y desarrollo de la planta del cafe. sistemas de producción de café en colombia,” pp. 22–60, 2007.
- [26] J. Arcila, F. FARFAN, A. Moreno, L. Salazar, and E. Hincapié, “Variedades del cafe. sistemas de producción de café en colombia,” pp. 169–204, 2007.

- [27] G. Puerta, "Calidad en taza de las variedades de coffea arabica l cultivadas en colombia," *Biblioteca CENICAFE*, 1998.
- [28] A. Hicks, "Post-harvest processing and quality assurance for speciality/organic coffee products," *AU Journal of Technology*, vol. 5, no. 3, 2002.
- [29] A. for Coffee Excellence, "Cup of Excellence," 2019. [Online]. Available: <https://allianceforcoffeexcellence.org/cup-of-excellence/>
- [30] J. M. Waller, M. Bigger, and R. J. Hillocks, *Coffee pests, diseases and their management*. CABI, 2007.
- [31] R. Gómez-Quiroga and C. Baeza-Aragón, "Control químico del volcamiento en germinadores de café (rhizoctonia solani, coffea arabica)," *Avances Técnicos Cenicafe (Colombia)*., no. 85, 1979.
- [32] S. López Duque and O. Fernández Borrero, "Epidemiología de la mancha de hierro del caféto cercospora coffeicola berk y cook," *Cenicafe (Colombia) v. 20 (1) p. 3-19*, 1969.
- [33] F. Guharay, J. Monterrey, D. Monterroso, C. Staver *et al.*, "Manejo integrado de plagas en el cultivo del café," *Manual técnico*, no. 44, 2000.
- [34] J. Zuluaga, G. Valencia, and J. Gonzalez, "Contribución al estudio de la naturaleza de la resistencia del caféto a ceratocystis fimbriata ell halst hunt," *Biblioteca CENICAFE*, 1971.
- [35] A. Rena, M. Maestri, A. Rena, E. Malavolta, M. Rocha, and T. Yamada, "Fisiologia do cafeeiro," *Informe agropecuario*, vol. 11, no. 126, 1986.
- [36] A. Wang, B. Bertrand, and B. Rapidel, "El ojo de gallo del caféto (mycena citricolor) desafíos de la caficultura en centroamérica," IICA, San José (Costa Rica). PROMECAFE CIRAD, París (Francia), Tech. Rep., 1999.
- [37] L. V. Cartagena, E. Sánchez, M. Vargas, A. Solórzano, F. Hernández, H. Iwasawa, and E. Freer, "Presencia de bacterias en el xilema de cafétos (rubiaceae: Coffea arabica) afectados por la enfermedad conocida como crespere," *Revista de biología tropical*, vol. 50, no. 1, pp. 45–48, 2002.
- [38] J. Waller, P. Bridge, R. Black, and G. Hakiza, "Characterization of the coffee berry disease pathogen, colletotrichum kahawae sp. nov." *Mycological Research*, vol. 97, no. 8, pp. 989–994, 1993.
- [39] GSMA Association, "Understanding the Internet of Things (IoT)," *Gsma Connected Living*, p. 15, 2014. [Online]. Available: http://www.gsma.com/connectedliving/wp-content/uploads/2014/08/cl_iot_wp_07_14.pdf
- [40] R. Kamal, *Internet of Things*:. McGraw Hill Education, 2017. [Online]. Available: <https://books.google.com.co/books?id=uS1HDwAAQBAJ>
- [41] K. A. M. Zeinab and S. A. A. Elmustafa, "Internet of things applications, challenges and related future technologies," *World Scientific News*, vol. 67, no. 2, pp. 126–148, 2017.

- [42] Joint Research Centre (JRC) of the European Commission, “Precision Agriculture: an Opportunity for Eu Farmers- Potential Support With the Cap 2014 - 2020,” *European Union*, p. 56, 2014. [Online]. Available: [http://www.europarl.europa.eu/RegData/etudes/note/join/2014/529049/IPOL-AGRI-NT\(2014\)529049_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/note/join/2014/529049/IPOL-AGRI-NT(2014)529049_EN.pdf)
- [43] J.-J. Jeong, H.-J. Ju, and J. Noh, “A Review of Detection Methods for the Plant Viruses,” *Research in Plant Disease*, vol. 20, no. 3, pp. 173–181, 2014. [Online]. Available: <http://koreascience.or.kr/journal/view.jsp?kj=SMBRCU&py=2014&vnc=v20n3&sp=173>
- [44] A. K. Q. Nascimento, J. A. A. Lima, and G. d. S. Barbosa, “A simple kit of plate-trapped antigen enzyme-linked immunosorbent assay for identification of plant viruses,” *Revista Ciência Agronômica*, vol. 48, no. 1, pp. 216–220, 2017.
- [45] F. Martinelli, R. Scalenghe, S. Davino, S. Panno, G. Scuderi, P. Ruisi, P. Villa, D. Stroppiana, M. Boschetti, L. R. Goulart, C. E. Davis, and A. M. Dandekar, “Advanced methods of plant disease detection. a review,” *Agronomy for Sustainable Development*, vol. 35, no. 1, pp. 1–25, Jan 2015. [Online]. Available: <https://doi.org/10.1007/s13593-014-0246-1>
- [46] K. Makkouk and S. Kumari, “Molecular diagnosis of plant viruses,” *Arab Journal of Plant Protection*, vol. 24, no. 1, pp. 135–138, 2006. [Online]. Available: [http://www.asplantprotection.org/PDF/AJPP/24-2{-}](http://www.asplantprotection.org/PDF/AJPP/24-2{-}2006/135-138.pdf)
- [47] O. Guzmán, E. Gómez, C. Rivillas, and C. Oliveros, “Utilización del procesamiento de imágenes para determinar la severidad de La Mancha de Hierro, en hojas de café,” *Cenicafé*, vol. 54, no. 3, pp. 258–265, 2003. [Online]. Available: <http://biblioteca.cenicafe.org/bitstream/10778/244/1/arc054%2803%29258-265.pdf>
- [48] Agriculturers, “Instalan datalogger en cafetales para monitorizar roya,” 2015. [Online]. Available: <http://agriculturers.com/instalan-datalogger-en-cafetales-para-monitorizar-roya/>
- [49] S. Sankaran, A. Mishra, R. Ehsani, and C. Davis, “A review of advanced techniques for detecting plant diseases,” *Computers and Electronics in Agriculture*, vol. 72, no. 1, pp. 1–13, 2010. [Online]. Available: <http://dx.doi.org/10.1016/j.compag.2010.02.007>
- [50] B. D. Lampson, A. Khalilian, J. K. Greene, Y. J. Han, and D. C. Degenhardt, “Development of a portable electronic nose for detection of cotton damaged by nezara viridula (hemiptera: Pentatomidae),” *Journal of Insects*, vol. 2014, 2014.
- [51] D. J. Mulla, “Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps,” *Biosystems Engineering*, vol. 114, no. 4, pp. 358 – 371, 2013, special Issue: Sensing Technologies for Sustainable Agriculture. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1537511012001419>
- [52] L. Araque and A. Jiménez, “Caracterización de firma espectral a partir de sensores remotos para el manejo de sanidad vegetal en el cultivo de palma de aceite,” *Revista Palmas*, vol. 30, no. 3, pp. 63–79, 2009.

- [53] M. H. Penner, *Basic Principles of Spectroscopy*. Cham: Springer International Publishing, 2017, pp. 79–88. [Online]. Available: https://doi.org/10.1007/978-3-319-45776-5_6
- [54] G. Elmasry, M. Kamruzzaman, D.-W. Sun, and P. Allen, “Principles and applications of hyperspectral imaging in quality evaluation of agro-food products: A review,” *Critical Reviews in Food Science and Nutrition*, vol. 52, no. 11, pp. 999–1023, 2012, pMID: 22823348. [Online]. Available: <https://doi.org/10.1080/10408398.2010.543495>
- [55] S. Azfar, A. Nadeem, and A. Basit, “Pest detection and control techniques using wireless sensor network: A review,” *Journal of Entomology and Zoology Studies*, vol. 3, no. 2, pp. 92–99, 2015.
- [56] W. Dargie and C. Poellabauer, *Fundamentals of Wireless Sensor Networks: Theory and Practice*, ser. Wireless Communications and Mobile Computing. Wiley, 2010. [Online]. Available: <https://books.google.com.co/books?id=8c6k0EVr6rMC>
- [57] N. Meratnia, B. J. Van der Zwaag, H. W. Van Dijk, D. J. Bijwaard, and P. J. Havinga, “Sensor networks in the low lands,” *Sensors*, vol. 10, no. 9, pp. 8504–8525, 2010. [Online]. Available: <http://www.mdpi.com/1424-8220/10/9/8504>
- [58] R. Jansen, J. Wildt, I. Kappers, H. Bouwmeester, J. Hofstee, and E. van Henten, “Detection of diseased plants by analysis of volatile organic compound emission,” *Annual Review of Phytopathology*, vol. 49, no. 1, pp. 157–174, 2011, pMID: 21663436. [Online]. Available: <https://doi.org/10.1146/annurev-phyto-072910-095227>
- [59] Y. Fang and R. P. Ramasamy, “Current and prospective methods for plant disease detection,” *Biosensors*, vol. 5, no. 3, pp. 537–561, 2015. [Online]. Available: <http://www.mdpi.com/2079-6374/5/3/537>
- [60] F. Banica, *Chemical Sensors and Biosensors: Fundamentals and Applications*. Wiley, 2012. [Online]. Available: <https://books.google.com.co/books?id=3m3qaVQdK3sC>
- [61] A. Trafton, “Bionic plants — MIT News,” 2014. [Online]. Available: <http://news.mit.edu/2014/bionic-plants>
- [62] R. Szeliski, *Computer Vision: Algorithms and Applications*, ser. Texts in Computer Science. Springer London, 2010. [Online]. Available: <https://books.google.com.co/books?id=bXzAlkODwa8C>
- [63] G. Bradski and A. Kaehler, *Learning OpenCV: Computer Vision with the OpenCV Library*. O’Reilly Media, 2008. [Online]. Available: <https://books.google.es/books?id=seAgiOfu2EIC>
- [64] E. Brown, “Nvidia aims octa-core, 256-GPU, Tegra X1 SoC at smart cars,” 2015. [Online]. Available: <http://linuxgizmos.com/nvidia-aims-octa-core-256-gpu-tegra-x1-soc-at-smart-cars/>
- [65] A. Bhardwaj, W. Di, and J. Wei, *Deep Learning Essentials: Your hands-on guide to the fundamentals of deep learning and neural network modeling*. Packt Publishing, 2018. [Online]. Available: <https://books.google.com.co/books?id=ISBKDwAAQBAJ>

- [66] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, ser. Adaptive computation and machine learning. MIT Press, 2016. [Online]. Available: <https://books.google.com.co/books?id=Np9SDQAAQBAJ>
- [67] J. Patterson and A. Gibson, *Deep Learning: A Practitioner's Approach*. O'Reilly Media, 2017. [Online]. Available: <https://books.google.com.co/books?id=rLcuDwAAQBAJ>
- [68] Ved, "Feature Learning , Deep Learning and Machine learning d4datascience.wordpress.com," 2015. [Online]. Available: <https://d4datascience.wordpress.com/2016/08/25/feature-learning-deep-learning-and-machine-learning/>
- [69] 3Blue1Brown, "But what *is* a neural network? — deep learning, chapter 1," Oct 2017. [Online]. Available: <https://www.youtube.com/watch?v=aircAruvnKk>
- [70] J. J. Gutiérrez, "Qué es un framework web?" Available in: http://www.lsi.us.es/~javierj/investigacion_ficheros/Framework.pdf Accessed May, vol. 12, 2014.
- [71] V. Dmytriiev, "Django or Rails," 2018. [Online]. Available: <https://djangostars.com/blog/django-or-rails/>
- [72] B. Lobitz, L. Beck, A. Huq, B. Wood, G. Fuchs, A. Faruque, and R. Colwell, "Climate and infectious disease: use of remote sensing for detection of vibrio cholerae by indirect measurement," *Proceedings of the National Academy of Sciences*, vol. 97, no. 4, pp. 1438–1443, 2000.
- [73] N.-Y. Su, "Remote monitoring system for detecting termites," Apr. 18 2000, uS Patent 6,052,066.
- [74] M. Mirik, J. E. Norland, R. L. Crabtree, and M. E. Biondini, "Hyperspectral one-meter-resolution remote sensing in yellowstone national park, wyoming: I. forage nutritional values," *Rangeland Ecology & Management*, vol. 58, no. 5, pp. 452–458, 2005.
- [75] P. K. Goel, S. O. Prasher, J.-A. Landry, R. M. Patel, R. Bonnell, A. A. Viau, and J. Miller, "Potential of airborne hyperspectral remote sensing to detect nitrogen deficiency and weed infestation in corn," *Computers and electronics in agriculture*, vol. 38, no. 2, pp. 99–124, 2003.
- [76] J. A. Bolaños, L. Campo, and J. C. Corrales, "Characterization in the visible and infrared spectrum of agricultural crops from a multirotor air vehicle," in *International Conference of ICT for Adapting Agriculture to Climate Change*. Springer, 2017, pp. 29–43.
- [77] A. Chemura, O. Mutanga, and T. Dube, "Remote sensing leaf water stress in coffee (coffea arabica) using secondary effects of water absorption and random forests," *Physics and Chemistry of the Earth, Parts A/B/C*, vol. 100, pp. 317–324, 2017.
- [78] J. Avelino, H. Zelaya, A. Merlo, A. Pineda, M. Ordoñez, and S. Savary, "The intensity of a coffee rust epidemic is dependent on production situations," *Ecological modelling*, vol. 197, no. 3-4, pp. 431–447, 2006.

- [79] F. Haddad, L. A. Maffia, E. S. Mizubuti, and H. Teixeira, "Biological control of coffee rust by antagonistic bacteria under field conditions in brazil," *Biological Control*, vol. 49, no. 2, pp. 114–119, 2009.
- [80] D. Jackson, J. Skillman, and J. Vandermeer, "Indirect biological control of the coffee leaf rust, *hemileia vastatrix*, by the entomogenous fungus *lecanicillium lecanii* in a complex coffee agroecosystem," *Biological Control*, vol. 61, no. 1, pp. 89–97, 2012.
- [81] D. Chaudhary, S. Nayse, and L. Waghmare, "Application of wireless sensor networks for greenhouse parameter control in precision agriculture," *International Journal of Wireless & Mobile Networks (IJWMN)*, vol. 3, no. 1, pp. 140–149, 2011.
- [82] M. Piamonte, M. Huerta, R. Clotet, J. Padilla, T. Vargas, and D. Rivas, "Wsn prototype for african oil palm bud rot monitoring," in *International Conference of ICT for Adapting Agriculture to Climate Change*. Springer, 2017, pp. 170–181.
- [83] A. Chemura, O. Mutanga, and T. Dube, "Separability of coffee leaf rust infection levels with machine learning methods at sentinel-2 msi spectral resolutions," *Precision Agriculture*, vol. 18, no. 5, pp. 859–881, Oct 2017. [Online]. Available: <https://doi.org/10.1007/s11119-016-9495-0>
- [84] A. Chemura, O. Mutanga, M. Sibanda, and P. Chidoko, "Machine learning prediction of coffee rust severity on leaves using spectroradiometer data," *Tropical Plant Pathology*, vol. 43, no. 2, pp. 117–127, 2018.
- [85] H. Geospatial, "Harris geospatial solutions in-depth knowledge of advanced geospatial analytics, machine learning, and remotely sensed data to make better decisions." [Online]. Available: <https://www.harrisgeospatial.com/>
- [86] F. Website, "Un laboratorio de innovacin para." [Online]. Available: <https://www.fomin.org/es-es/PORTADA.aspx>
- [87] C. Shaoming and W. Jun, "Method for rapid diagnosis of crop disease by volatile matter," Jan 2013.
- [88] G. R. Henderson and D. A. Grafton, "Method and apparatus for measuring light reflectance absorption and or transmission," Oct. 7 1975, uS Patent 3,910,701.
- [89] G. L. McQuilkin and G. L. Engelke, "Spectral imaging system for remote and noninvasive detection of target substances using spectral filter arrays and image capture arrays."
- [90] L. Domenzain and C. Fallet, "Necessary steps for the systematic calibration of a multispectral imaging system to achieve a targetless workflow in reflectance estimation: a study of parrot sequoia for precision agriculture," in *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXIV*, vol. 10644. International Society for Optics and Photonics, 2018, p. 1064416.
- [91] P. Agarwal, "Agrocloudopen surveillance of indian agriculture via cloud," in *Information Technology (InCITE)-The Next Generation IT Summit on the Theme-Internet of Things: Connect your Worlds, International Conference on*. IEEE, 2016, pp. 182–186.

- [92] E. Lasso and J. C. Corrales, "Towards an alert system for coffee diseases and pests in a smart farming approach based on semi-supervised learning and graph similarity," in *International Conference of ICT for Adapting Agriculture to Climate Change*. Springer, 2017, pp. 111–123.
- [93] H. G. Pulido, R. De la Vara Salazar, P. G. González, C. T. Martínez, and M. d. C. T. Pérez, *Análisis y diseño de experimentos*. McGraw-Hill New York, NY, USA:, 2012.
- [94] DFrobot.com. Gravity: Analog pH Sensor / Meter Pro Kit For Arduino. (2018). [Online]. Available: <https://www.dfrobot.com/product-1110.html>
- [95] Kitronik.co.uk. How an LDR (Light Dependent Resistor) Works. (2018). [Online]. Available: <https://www.kitronik.co.uk/blog/how-an-ldr-light-dependent-resistor-works/>
- [96] Dfrobot.com. Gravity: Waterproof DS18B20 Sensor Kit. (2018). [Online]. Available: <https://www.dfrobot.com/product-1354.html>
- [97] ——. Gravity: Analog Capacitive Soil Moisture Sensor- Corrosion Resistant. (2018). [Online]. Available: <https://www.dfrobot.com/product-1385.html>
- [98] ——. DHT22 Temperature and Humidity Sensor. (2018). [Online]. Available: <https://www.dfrobot.com/product-1102.html>
- [99] Trossenrobotics.com. Dynamixel AX-12A Robot Actuator. (2018). [Online]. Available: <https://www.trossenrobotics.com/dynamixel-ax-12-robot-actuator.aspx>
- [100] Kitronik.co.uk. Digital Servo Shield for Arduino SKU:DRI0027. (2018). [Online]. Available: https://www.dfrobot.com/wiki/index.php/Digital_Servo_Shield_for_Arduino_SKU:DRI0027
- [101] Didacticaselectronicas.com. Moto reductor 110rpm. (2018). [Online]. Available: <https://www.didacticaselectronicas.com/index.php/elementos-electromecanicos/motores-y-solenoides-1/motores-dc-1/motorreductor-mtrr110rpm-a-motor-dc-110rpm-detail>
- [102] Lastminuteengineers.com. In-Depth: Interface L298N DC Motor Driver Module with Arduino. (2018). [Online]. Available: <https://lastminuteengineers.com/l298n-dc-stepper-driver-arduino-tutorial/>
- [103] Dfrobot.com. 12V Solenoid Valve - DN20. (2018). [Online]. Available: <https://www.dfrobot.com/product-1530.html>
- [104] ——. Gravity: Digital 5A Relay Module. (2018). [Online]. Available: <https://www.dfrobot.com/product-64.html>
- [105] ——. Water Flow Sensor - 1/2 SKU: SEN0217. (2018). [Online]. Available: https://www.dfrobot.com/wiki/index.php/Water_Flow_Sensor_-_1/2_SKU:_SEN0217
- [106] ——. Weather Station Kit with Anemometer/Wind Vane/Rain Bucket. (2018). [Online]. Available: <https://www.dfrobot.com/product-1308.html>

- [107] Pololu.com. Pololu - 30:1 Metal Gearmotor 37Dx52L mm with 64 CPR Encoder (No End Cap). (2018). [Online]. Available: <https://www.dfrobot.com/product-1308.html>
- [108] Geckodrive.com. G320X Digital Servo Drive. (2018). [Online]. Available: <https://www.geckodrive.com/g320x-digital-servo-drive.html>
- [109] Analog.com. LTC3780. (2018). [Online]. Available: <https://www.analog.com/en/products/ltc3780.html>
- [110] Raspberrypi.org. Raspberry Pi 3 Model B. (2018). [Online]. Available: <https://www.raspberrypi.org/products/raspberry-pi-3-model-b-plus/>
- [111] Logitech.com. C270 HD Webcam. (2018). [Online]. Available: <https://www.logitech.com/es-es/product/hd-webcam-c270>
- [112] Mapir.camera. Survey3N Camera - Red Green NIR (RGN, NDVI). (2018). [Online]. Available: <https://www.mapir.camera/collections/survey3/products/survey3n-camera-red-green-nir-rgn-ndvi>
- [113] ——. Survey3N Camera - RedEdge (RE). (2018). [Online]. Available: <https://www.mapir.camera/collections/survey3/products/survey3n-camera-rededge-re>
- [114] K. Pulli, A. Baksheev, K. Korniyakov, and V. Eruhimov, “Real-time computer vision with opencv,” *Communications of the ACM*, vol. 55, no. 6, pp. 61–69, 2012.
- [115] Z. Chaczko and R. Braun, “Learning data engineering: Creating iot apps using the node-red and the rpi technologies,” in *2017 16th International Conference on Information Technology Based Higher Education and Training (ITHET)*. IEEE, 2017, pp. 1–8.
- [116] M. Alvermann, “Introduction to mongodb,” *URL* <https://www.mongodb.com/citedon>, p. 9, 2016.
- [117] J. Mineraud, O. Mazhelis, X. Su, and S. Tarkoma, “Contemporary internet of things platforms,” *arXiv preprint arXiv:1501.07438*, 2015.
- [118] Numpy.org. NumPy. (2018). [Online]. Available: <http://www.numpy.org/>
- [119] Pandas.pydata.org. Pandas: powerful Python data analysis toolkit. (2018). [Online]. Available: <https://pandas.pydata.org/>
- [120] Scikit-learn.org. Scikit-learn. (2018). [Online]. Available: <https://scikit-learn.org/stable/>
- [121] Keras.io. Keras: The Python Deep Learning library. (2018). [Online]. Available: <https://keras.io/>
- [122] Tensorflow.org. TensorFlow. (2018). [Online]. Available: <https://www.tensorflow.org/>
- [123] scikit-learn developers, “sklearn.model_selection.gridsearchcv.” [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html

- [124] —, “3.1. cross-validation: evaluating estimator performance.” [Online]. Available: https://scikit-learn.org/stable/modules/cross_validation.html
- [125] A. Burkov, *The Hundred-Page Machine Learning Book*. Andriy Burkov, 2019. [Online]. Available: <https://books.google.com.co/books?id=0jbxwQEACAAJ>

APPENDICES

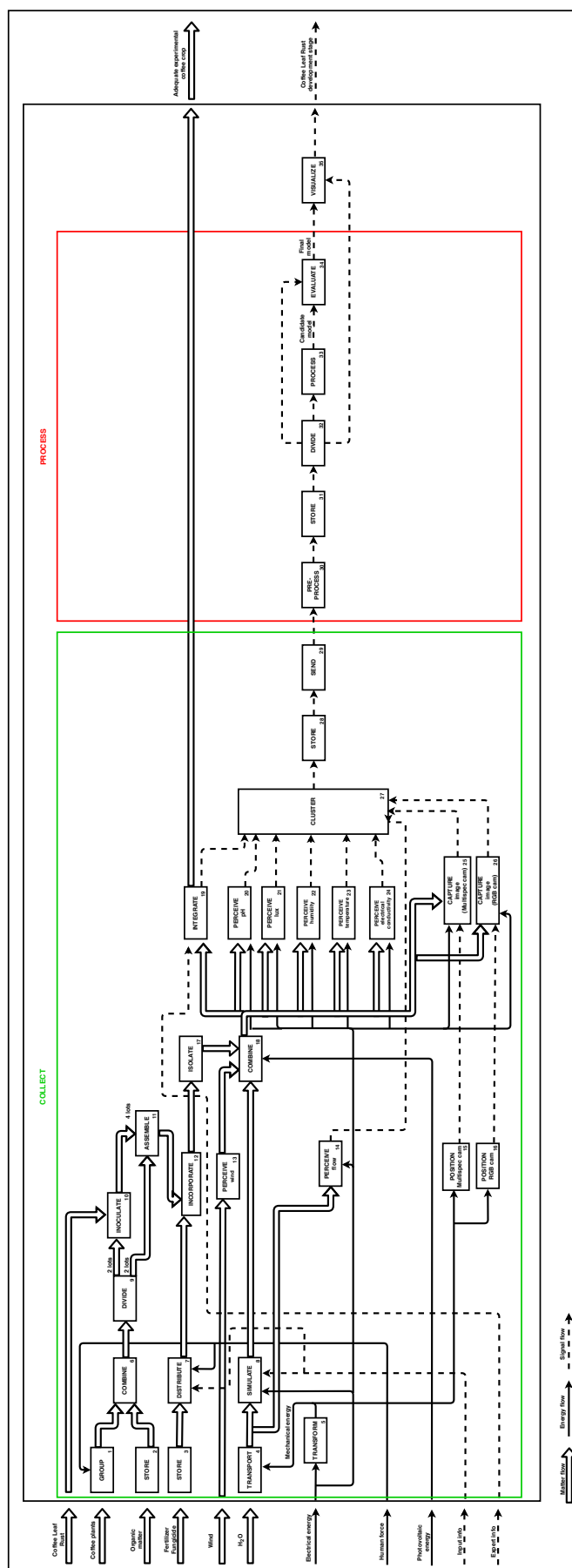
A PDS

REQUIREMENTS	NEED	SPECIFIC REQUIREMENTS TO BE FULFILLED	SOURCE	MEASURE VARIABLE	MEASURE UNIT	VALUE-RANGE	PRIORITY (0-5)	VALIDATION AND VERIFICATION
CROP REQUIREMENTS	Store coffee plants	Quantity of plants	Design team and CENICAFÉ	Quantity	Numeric	16-40	5	Number of plants in scale crop
		Varieties of coffee	Interviews	Coffee variety	Variety	Caturra	5	Variety guaranteed by a Caturra provider
		Environmental resistance	Design team and Biology team	Corrosion resistance, surface protection methods	Y/N	Y	5	Additional protection against moisture
		Water drainage from soil	Design team	Water flow	cm ³ /h	Not defined	4	Tests in construction of prototype
		Hold weight of plants and soil	Design team	Weight	Kg	<=2,000	5	Static and dynamic calculations
	Separate coffee crop lots	Different agronomic management in lots (fertilizers and fungicides)	Luisa Posada, Biology team and Interviews	Quantity of different coffee lots	Numeric	>=4	5	Commercial products in coffee lots
		Separate crop soil from rocky soil	Design team	Independent crop soil from rocky soil	Y/N	Y	4	Construction of prototype
		Separate soil between lots	Design team and Luisa Posada	Independent soil between lots	Y/N	Y	5	Construction of prototype
		Air barrier for Coffee Leaf Rust propagation problems	Design team, CENICAFÉ, WCR and Biology team	Isolated air conditions	Y/N	Y	5	Construction of prototype
		Distance between plants	CENICAFÉ	Distance	cm	>75	4	CAD Model and construction of prototype
MANUFACTURING REQUIREMENTS	Free growing space for each plant	Root growth without problems	CENICAFÉ	Distance	Radio (cm) Depth (cm)	>75 >40	4	Biological specifications for plants and CAD model
	Light control	Shade control by lot	Diego Miguel Sierra, Design team and CENICAFÉ	Percentage of shade	%	<50	3	Lux meter and commercial poly-shade
	Irrigation	Simulate rain characteristics of a specific Colombian region	Interviews and Design team	Liters	lt	Not defined	4	Pluviometer
		Long time for water disposal	Design team	Liters per hour	lt/h	Not defined	3	Flow of water
		Budget for construction	Design team and EAFIT University	General budget for materials	COP/m ³	3M-5M	4	Tests in construction of prototype
	Materials	Facility of joints between components	Design team	Plug and Play	0-5	4	3	Construction of prototype
		Manufacturability of materials	Design team	Commercial or standard components	% 0-10	>70 >7	4	Tests in construction of prototype
		Resistance of materials against environmental conditions	Design team	Immunitized materials	Y/N	Y	5	Matweb and tests in construction of prototype
		Door dimensions	Infrastructure Department at EAFIT	Height Width	cm cm	<217 <117	5	Flexometer, construction of prototype and CAD model
		Easy assembly mechanism, non-electric assembly	Design team	Assembly in field	0-5	4	4	CAD model and construction of prototype
Assembly constraints	Materials that do not affect soil nutritional status	Design team and Eduardo Duque	Materials' effects in soil	0-5	<1	5	Matweb and materials' properties	
	Dimensions of space inside the greenhouse	Infrastructure Department at EAFIT	Height Width Length	m m m	0.5 1 2	5	Flexometer and construction of prototype	

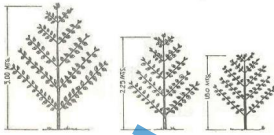
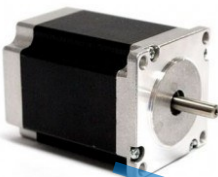
REQUIREMENTS	NEED	SPECIFIC REQUIREMENTS TO BE FULFILLED	SOURCE	MEASURE VARIABLE	MEASURE UNIT	VALUE-RANGE	PRIORITY (0-5)	VALIDATION AND VERIFICATION
ELECTRONIC REQUIREMENTS	Source of energy	Type of electricity	Design team	Type of electricity	VAC Frequency	110 @60Hz	4	Electronic Implementation
		Source of electricity	Design team and Infrastructure Department at EAFIT	Source of electricity	Normal source	Normal	4	Construction of prototype and Physical Plant Department at EAFIT
		Electric availability	Infrastructure Department at EAFIT	Distance of electricity source Availability of electric power	m Y/N	<5 Y	5	Physical Plant Department at EAFIT
	Sensors for every lot	pH sensor	Design team and Interviews	pH pH precision Direct measurement from soil	pH ± pH Y/N	0-14 0.5 Y	5	Sensor tests and specifications
		Humidity sensor	Design team and Interviews	Humidity from air Humidity from air precision	% ± %	50-100 5	5	Sensor tests and specifications
		Luminosity sensor	Design team and Interviews	Lux	lumens	Not defined	5	Sensor tests and specifications
		Temperature sensor (air and soil)	Design team and Interviews	Temperature from air and soil Temperature from air and soil precision	°C ± °C	18-30 0.1	5	Sensor tests and specifications
		Electrical conductivity sensor	Design team and Interviews	EC from soil EC from soil precision	S/m ± S/m	0-1 0.1	4	Sensor tests and specifications
		Flow rate sensor	Design team and Interviews	Flow rate from rain-system	m3/s	Not defined	4	Sensor tests and specifications
		Wind speed sensor	Design team and Interviews	Speed	m/s	0-10	5	Sensor tests and specifications
		Moisture sensor	Design team and Interviews	Moisture from soil Moisture from soil precision	% ± %	50-100 5	5	Sensor tests and specifications
		Sensors protected against liquids and heat	Design team and Biology team	IP protection	IP rank	>3	4	Tests under liquid and heat conditions
		Embedded system device for sensing and actuating	Design team	Analog pins Digital pins	Numeric Numeric	16 16	4	Electronic and software implementation
	Embedded devices	Device for Internet connectivity	Design team	WiFi connection	Y/N	Y	4	Specifications and tests of prototype
		Independent systems for data collection and rain simulation	Design team	Number of embedded devices	Numeric	>2	3	Tests of independent processes
		Embedded devices protected against liquids and heat	Design team and Biology team	IP protection	IP rank	>3	4	Tests under liquid and heat conditions
	Communication system	Type of communication	Design team	USB communication and WiFi communication	Y/N	Y	4	Tests of local and remote communication
	Detect diseases by physiological and chemical changes in plants	Remote Sensing	Design team and Alejandro Marulanda	Wavelengths [λ]	nm	500-1,000	5	Multispectral camera specifications
		Support for Remote Sensing	Design team	Type of joint Adjustability	Screw Degrees	Not defined 50	5	Multispectral camera specifications and construction of prototype
		RGB images per tree from the bottom	Design team and CENICAFE	Image per tree from the bottom	Y/N	Y	5	RGB camera specifications and construction of prototype

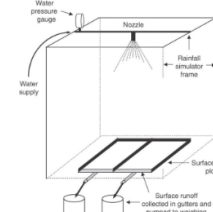
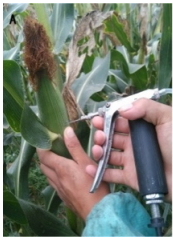
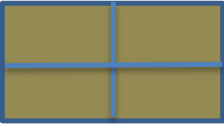
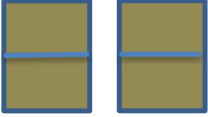
REQUIREMENTS	NEED	SPECIFIC REQUIREMENTS TO BE FULFILLED	SOURCE	MEASURE VARIABLE	MEASURE UNIT	VALUE-RANGE	PRIORITY (0-5)	VALIDATION AND VERIFICATION
SOFTWARE REQUIREMENTS	Database	Database selection for storage	Design team and Edwin Montoya	Memory space	GB	>16	4	Specifications of selected database plan
		Cost of database usage	Design team	Monthly cost	COP	<50,000	5	Specifications of selected database plan
		Redundancy	Design team	Quantity of database copies	Numeric	>2	5	Specifications of selected database plan
		Local and remote database	Design team	Local and remote storage availability	Y/N	Y	5	Tests of redundant storage, same data in both databases
	Data collection	Frequency of data collection	Design team	Daily data updates	Numeric	>2	4	Timestamp in database
		Quality of collected data	Design team	NANs per week Server acknowledgment	Numeric Status code	<4 200	5 5	Daily data review Data transfer history in embedded device
		Automatic and manual data collection	Design team	Data collection in both modes	Y/N	Y	4	Tests of software behavior
		ID for each lot	Design team and Biology team	Congruent IDs in database	Y/N	Y	5	Tests of software and database records
		Data persistence without Internet connection	Design team	Local storage of crop package before data transfer	Y/N	Y	5	Tests of elapsed time between data collection between 2 contiguous lots
				Regular time step between samples in remote database	Time between samples (h)	3 Y	4 5	Checks of timestamp in remote database
				Reconnection protocol	Y/N	Y	5	Tests of data transfer disconnecting the Internet for a while and reconnecting it
		Data collection when it is not raining	Design team	Times for data collection differ from times of raining	Y/N min	Y 20	5	Tests of rain simulation at data collection time Cross of data collection cronogram with rain simulation cronogram
	Data processing	Pre-processing of data	Design team and Sebastián Rodríguez	Average of data before transfer NANs in data Dimensionality reduction of data	Numeric Numeric Y/N	>3 0 Y	4 5 4	Tests through manual data collection Unaltered data main characteristics Adjusted R squared
		Model construction based on data	Design team, Felipe Posada, Sebastian García and Sebastian Rodríguez	Accuracy Multiple models	% Numeric	>90 >5	5 5	Confusion matrix, RMSE, Test-set accuracy Tests of models, comparison of each model diagnosis
	Results of diagnosis model	Visualization of the diagnosis on the Internet	Design team, coffee growers and interviews	Visualization of crop with characteristics and diagnosis per lot	Y/N	Y	5	Tests of visualization of results on a web page
		Accuracy of technological diagnosis based on the real Coffee Leaf Rust development stage under scale coffee crop conditions	Design team and Biology team	Error between technological diagnosis and visual diagnosis	%	<10	5	Biology team
		Accuracy of technological diagnosis based on the real Coffee Leaf Rust development stage under real farming conditions	Agronomy experts	Error between technological diagnosis and visual diagnosis	%	<15	5	Agronomy experts on coffee

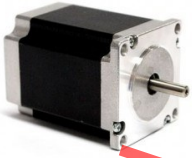
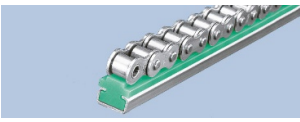
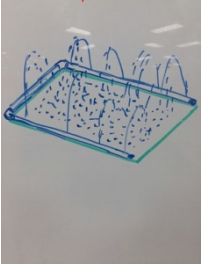
B FUNCTIONAL STRUCTURE

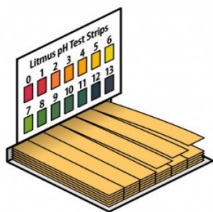
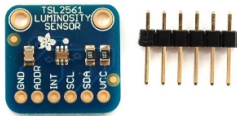
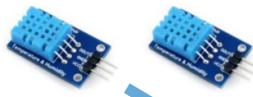
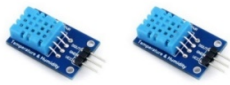


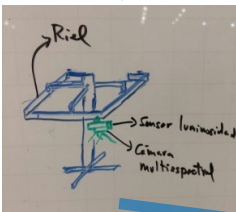
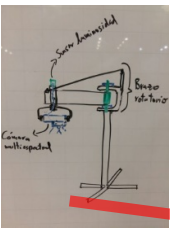
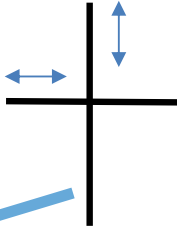
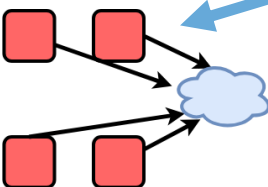
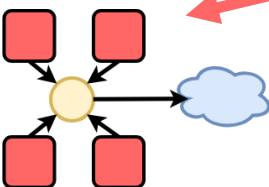
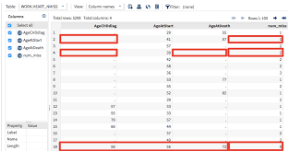
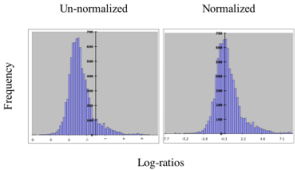
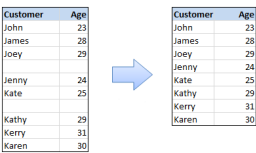
C MORPHOLOGICAL MATRIX

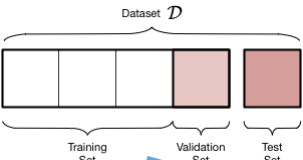
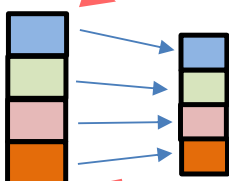
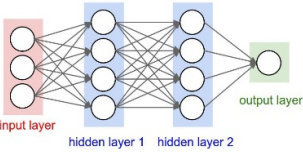
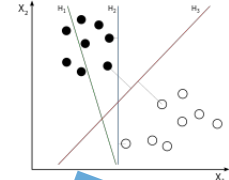
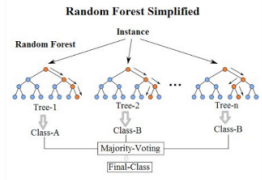
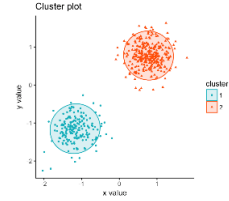
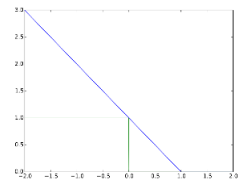
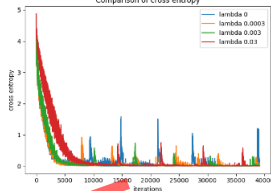
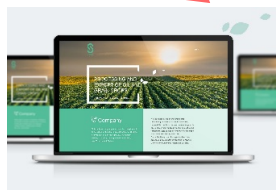
FUNCTION CARRIER				
Function	PROPOSAL 1	PROPOSAL 2	PROPOSAL 3	PROPOSAL 4
1. Group coffee plants	"Almacigo" group 	Coffee group with diverse ages 	Coffee group in production stage 	Coffee plants between "almacigo" and production stage 
2. Store organic matter	Plant pots 	Plant bags 	Crop beds 	Regular crop disposal 
3. Store fertilizer and fungicide	Bottles 	Tanks 	Spray 	
4. Transport H ₂ O	Pumping 	Gravity from tank 	Normal pressure from aqueduct system 	
5. Transform electrical energy	DC motor 	Stepper motor 	Servomotor 	
6. Combine coffee plants with organic matter	Regular crop 	Single plant care 	Raised garden beds 	

7. Distribute fertilizer and fungicide	Spray 	Pressurized knapsack 		
8. Simulate rain with H ₂ O	Holes in tubes 	Sprinkler bottom-up 	Sprinkler top-down 	
9. Divide coffee plants with organic matter	Wood, class IV, penetration index = 9 	Acrylic 	Bricks and cement 	Plastic, polyethylene and polipropylene. 
10. Inoculate Coffee Leaf Rust	In-vitro 	Natural (wind) 	Rub with Coffee Leaf Rust 	Brush 
11. Assemble inoculated and non-inoculated coffee plants	One single piece divided 	Independent lots 	Pairs of lots 	
12. Incorporate fungicide and fertilizer with coffee plants	Sprinkler 	Spray 	Pressurized knapsack 	Holes in tubes 

13. Perceive wind velocity	Ultrasonic wind sensor 	Dual wind sensor 	GPRS weather station 	
14. Perceive volumetric flow rate	Flow meter 	Pluviometer 	Electromagnetic flow meter 	
15. Position multispectral camera	Stepper motor 	Servomotor 	Manual 	
16. Position RGB camera	Movement by chains 	Curved rails 	Endless screw 	Small car positioning 
17. Isolate each lot of coffee plants	Plastic curtains 	Fabric curtains 	Blinds 	
18. Combine simulated rain with isolated lots of coffee plants	Holes in tubes raised from the lots 	Sprinkler above the lots 		

19. Integrate expert info with isolated lots of coffee plants	Remote monitoring 	Visual inspection 		
20. Perceive pH in isolated lots of coffee plants	Type drainage 	Type soil 	Type general 	
21. Perceive lux in isolated lots of coffee plants	Lux sensor 	Digital luminosity sensor 	LDR 	
22. Perceive humidity in isolated lots of coffee plants	Soil moisture capacitive sensor 	DHT-33 (x2) 	SHT-10 humidity sensor 	5TE / DHT-33 
23. Perceive temperature in isolated lots of coffee plants	Type soil 	Thermometer 	DHT-33 (x2) 	5TE / DHT-33 
24. Perceive electrical conductivity in isolated lots of coffee plants	Conductivity electrode 	5TE 		

25. Capture images from multispectral camera	Single rail 	Rotary arm 	Manual 	
26. Capture images from RGB camera	Zig-zag rail 	Straight rail 	2-axis rail 	Elliptical rail 
27. Cluster sensors' readings in a package	One package for each lot 	One package for the entire crop with lot ID 		
28. Store collected data locally	USB 	Micro-SD 		
29. Send data package to remote server	WiFi 	XBee 	Bluetooth 	
30. Pre-process data package	Missing values to zeros 	Normalization 	Remove records having missing values 	

31. Store pre-processed data package	Remote database 	Local database on the remote server 		
32. Divide data into training, validation and test sets	Random division 	Balance across classes + Shuffle 		
33. Process training dataset	Deep Neural Network 	Support Vector Machines 	Random forest 	Clustering 
34. Evaluate the candidate model with the validation dataset	Hinge-loss 	Cross-entropy 	Mean-Squared-Error $MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$	
35. Visualize the results through a simple interface using some test data	Develop a web interface 	Develop a software to automatically send e-mails 	Develop an app 	Use Thingworx PaaS 

References

- 1 – Proposal 1:** “Cultivemos café / Almacigo,” Quienes Somos, 15-Dec-2015. [Online]. Available: https://www.cenicafe.org/es/index.php/cultivemos_cafe/almacigo/Almacigo. [Accessed: 03-Feb-2019].
- 1 – Proposal 2:** Cultivo de platano - Anacafé. [Online]. Available: https://www.anacafe.org/glifos/index.php/Varietades_de_cafe. [Accessed: 03-Feb-2019].
- 1 – Proposal 3:** “Coffee Plants of the World,” Specialty Coffee Association. [Online]. Available: <https://sca.coffee/research/coffee-plants-of-the-world/>. [Accessed: 03-Feb-2019].
- 1 – Proposal 4:** Ray, “Vacation Planning,” Costa Rica Guide. [Online]. Available: <https://costa-rica-guide.com/photos/coffee-growing-picking/>. [Accessed: 03-Feb-2019].
- 2 – Proposal 1:** Homepage. [Online]. Available: <https://www.gsdistribution.ca/hc-canada>. [Accessed: 03-Feb-2019].
- 2 – Proposal 2:** J. Poindexter, “Growing Potatoes: A Beginner's Guide to Planting Big & Healthy Potatoes,” MorningChores, 02-Nov-2018. [Online]. Available: <https://morningchores.com/growing-potatoes/>. [Accessed: 03-Feb-2019].
- 2 – Proposal 3:** “Raised Garden Beds,” USA Garden Company. [Online]. Available: <http://www.usagardencompany.com/raised-garden-beds.html>. [Accessed: 03-Feb-2019].
- 2 – Proposal 4:** “Pics for you every day,” King Yu of Xia Dynasty also known as Xia Yu Wikipedia t. [Online]. Available: <http://naturtreenspicerx.pw/Healthy-soil-healthy-plants-healthy-people-articles-Soil-carbon.html>. [Accessed: 03-Feb-2019].
- 3 – Proposal 1:** “Signum Fungicide 2.5kg,” ProGreen. [Online]. Available: <https://www.progreen.co.uk/signum-fungicide-2-5kg>. [Accessed: 03-Feb-2019].
- 3 – Proposal 2:** Prabha Polycolor Pvt Ltd. [Online]. Available: <http://kkmasterbatches.com/>. [Accessed: 03-Feb-2019].
- 3 – Proposal 3:** 2019. [Online]. Available: <https://www.homedepot.ca/en/home/p.defender-1l-ready-to-use-garden-fungicide.1000107206.html>. [Accessed: 04-Feb-2019].
- 4 – Proposal 1:** “Evolution Series,” Dream Pond. [Online]. Available: <https://www.dreampond.com/koi-pond-water-garden/water-pumps/evolution-series.html>. [Accessed: 04-Feb-2019].
- 4 – Proposal 2:** Web-Know-How, “Water Delivery selection guide,” ROVER & ROVER Battery Single Point Watering System. [Online]. Available: http://www.aquapro.net/tank_info.html. [Accessed: 04-Feb-2019].
- 4 – Proposal 3:** “Canilla PVC para manguera manija mariposa Ø ½ DUKE- www.deplano.com.ar,” DePlano. [Online]. Available: <https://www.deplano.com.ar/canillas-pvc/2270-canilla-esf-pvc-mariposa-o-1-2-duke-60.html>. [Accessed: 04-Feb-2019].
- 5 – Proposal 1:** Buyers Products | Brands. [Online]. Available: <https://www.surpluscenter.com/Electrical/AC-Motors/>. [Accessed: 04-Feb-2019].
- 5 – Proposal 2:** “NEMA 23 Stepper Motor | 3A , 30kg-cm,” DIYElectronics. [Online]. Available: <https://www.diyelectronics.co.za/store/stepper-motors/478-nema-23-57bygh115-003.html>. [Accessed: 04-Feb-2019].
- 5 – Proposal 3:** “Hitec HS-645MG Servo Motor”, RobotShop, 2019. [Online]. Available: <https://www.robotshop.com/uk/hitec-hs-645mg-servo-motor.html>. [Accessed: 04-Feb-2019].
- 6 – Proposal 1:** “Coffee Plants of the World,” Specialty Coffee Association. [Online]. Available: <https://sca.coffee/research/coffee-plants-of-the-world/>. [Accessed: 04-Feb-2019].
- 6 – Proposal 2:** “100 litre Woven Planter Bags,” Nursery and Garden Supplies Australia. [Online]. Available: <https://nurseryandgardensupplies.com.au/product/woven-planter-bags-100-ltr/>. [Accessed: 04-Feb-2019].
- 6 – Proposal 3:** “How to Grow Plants in Raised Beds,” Thompson & Morgan. [Online]. Available: <https://www.thompson-morgan.com/how-to-grow-plants-in-raised-beds>. [Accessed: 04-Feb-2019].
- 7 – Proposal 1:** 2019. [Online]. Available: <https://www.homedepot.ca/en/home/p.defender-1l-ready-to-use-garden-fungicide.1000107206.html>. [Accessed: 04-Feb-2019].
- 7 – Proposal 2:** “Kingfisher Garden Pressure Sprayer 5L,” JTF. [Online]. Available: <https://www.jtf.com/sprayer-pressure-5-l.html>. [Accessed: 04-Feb-2019].
- 8 – Proposal 1:** Hydroponic Gardening & More with Brent, “PVC Watering System,” YouTube, 31-Dec-2015. [Online]. Available: <https://www.youtube.com/watch?v=AVWP7gjka0>. [Accessed: 04-Feb-2019].
- 8 – Proposal 2:** M. Caffeine, “How To: Build A DIY Rain Machine For Under \$20!,” YouTube, 18-Apr-2014. [Online]. Available: https://www.youtube.com/watch?v=XnHRB_P4F70. [Accessed: 04-Feb-2019].
- 8 – Proposal 3:** Dl.sciencesocieties.org, 2019. [Online]. Available: <https://dl.sciencesocieties.org/images/publications/jeq/35/6/2229fig2.jpeg>. [Accessed: 04-Feb-2019].
- 9 – Proposal 1:** “Travesses de fusta,” Fusta de serra. [Online]. Available: <https://www.fitorforestal.com/ca/travesses-de-fusta.html>. [Accessed: 04-Feb-2019].
- 9 – Proposal 2:** “MRC MERCANTILE,” MRC Mercantile Acrylic Plastics Cebu. [Online]. Available: <http://acrylicsheetcebu.com.ph/>. [Accessed: 04-Feb-2019].
- 9 – Proposal 3:** “He Made This Ugly Cinderblock Wall Beautiful With MORE Cinderblocks! - Page 2 of 2,” Pinterest. [Online]. Available: <https://co.pinterest.com/pin/163818505173203549/>. [Accessed: 04-Feb-2019].

- 9 – Proposal 4:** "Chopping Boards & Racks", Busy Chef, 2019. [Online]. Available: https://www.busychef.co.uk/collections/chopping-boards-racks/products/collection_tablecraft-cutting-board. [Accessed: 04- Feb- 2019].
- 10 – Proposal 1:** "Field Inoculation and Classification of Maize Ear Rot Caused by *Fusarium verticillioides* ", Bradford Protein Assay - BIO-PROTOCOL. [Online]. Available: <https://bio-protocol.org/e3099>. [Accessed: 04-Feb-2019].
- 10 – Proposal 2:** Inspiring.vector@gmail.com, "diente de león blanco," Depositphotos. [Online]. Available: <https://mx.depositphotos.com/111413150/stock-illustration-dandelion-on-white.html>. [Accessed: 04-Feb-2019].
- 10 – Proposal 3:** "Cómo limpiar las plantas," MiaRevista.es. [Online]. Available: <https://www.miarevista.es/hogar/articulo/como-limpiar-las-plantas-481444999448>. [Accessed: 04-Feb-2019].
- 10 – Proposal 4:** "5 tuti tipp, hogy tisztán tartsd a szobanövényeidet," Gardenista, 03-Mar-2018. [Online]. Available: <https://gardenista.hu/2018/03/04/5-tut-tipp-hogy-tisztan-tartsd-a-szobanovenyeidet/>. [Accessed: 04-Feb-2019].
- 12 – Proposal 1:** "What is the difference between sprinkler irrigation and furrow irrigation?", Quora, 2019. [Online]. Available: <https://www.quora.com/What-is-the-difference-between-sprinkler-irrigation-and-furrow-irrigation>. [Accessed: 04- Feb- 2019].
- 12 – Proposal 2:** 2019. [Online]. Available: <https://www.homedepot.ca/en/home/p.defender-1l-ready-to-use-garden-fungicide.1000107206.html>. [Accessed: 04- Feb- 2019].
- 12 – Proposal 3:** "Kingfisher Garden Pressure Sprayer 5L," JTF. [Online]. Available: <https://www.jtf.com/sprayer-pressure-5-l.html>. [Accessed: 04-Feb-2019].
- 12 – Proposal 4:** Daisy Creek Farms with Jag Singh, "Drill Holes in 1/2' PVC Pipe and This Happens," YouTube, 08-May-2018. [Online]. Available: <https://www.youtube.com/watch?v=hazWfD45wNM>. [Accessed: 04-Feb-2019].
- 13 – Proposal 1:** "Wind Speed and Direction Sensor | Dust Sentry & AQM 65," Aeroqual. [Online]. Available: <https://www.aeroqual.com/product/gill-wind-sensor>. [Accessed: 04-Feb-2019].
- 13 – Proposal 2:** "Wind Speed/Direction, Modbus, 10 m Cable - IC-WSD-1-002," Instrument Choice. [Online]. Available: <https://www.instrumentchoice.com.au/instrument-choice/weather-stations/proffesional-weather-stations/dyacon-weather-station-accessories-and-sensors/wind-speed-direction-modbus-10-m-cable-ic-wsd-1-002>. [Accessed: 04-Feb-2019].
- 13 – Proposal 3:** T. Twmffat and Instructables, "Arduino GPRS Weather Station - Part 2: Upgraded Wind Sensors and Improved Energy Efficiency," Instructables.com, 20-Jan-2019. [Online]. Available: <https://www.instructables.com/id/Arduino-GPRS-Weather-Station-Part-2-Upgraded-Wind-/>. [Accessed: 04-Feb-2019].
- 14 – Proposal 1:** "water flow meter liquid flowmeter rotameter meter flow 94mm-in Flow Meters from Tools on Aliexpress.com | Alibaba Group," aliexpress.com. [Online]. Available: <https://www.aliexpress.com/item/water-flow-meter-liquid-flowmeter-rotameter-meter-flow-94mm/32373639290.html>. [Accessed: 04-Feb-2019].
- 14 – Proposal 2:** "Pluviometer," Weer en klimaat. [Online]. Available: <https://weerkenklimaat.weebly.com/pluviometer.html>. [Accessed: 04-Feb-2019].
- 14 – Proposal 3:** "Electro Magnetic Water Meter," IndiaMART.com. [Online]. Available: <https://www.indiamart.com/proddetail/electro-magnetic-water-meter-17008786755.html>. [Accessed: 04-Feb-2019].
- 15 – Proposal 1:** "NEMA 23 Stepper Motor | 3A , 30kg-cm," DIYElectronics. [Online]. Available: <https://www.diyelectronics.co.za/store/stepper-motors/478-nema-23-57bygh115-003.html>. [Accessed: 04-Feb-2019].
- 15 – Proposal 2:** "Hitec HS-645MG Servo Motor", RobotShop, 2019. [Online]. Available: <https://www.robotshop.com/uk/hitec-hs-645mg-servo-motor.html>. [Accessed: 04- Feb- 2019].
- 15 – Proposal 3:** Micrologia, "con una sola mano tirando la cuerda sobre fondo blanco," Depositphotos. [Online]. Available: <https://mx.depositphotos.com/27740207/stock-photo-single-hand-pulling-rope-on.html>. [Accessed: 04-Feb-2019].
- 16 – Proposal 1:** Ltd, "Dişli Zincir Dişli Zincir / Urunlerimiz / Zincir Kızakları," Dişli Zincir / Türkiye. [Online]. Available: <http://www.dislizincir.com.tr/Urunlerimiz/Zincir-Kizaklari>. [Accessed: 04-Feb-2019].
- 16 – Proposal 2:** D. Carpenter, "Curved Linear Rail," Ballscrews | 16mm & Below For Industrial Applications. [Online]. Available: <https://www.tpa-us.com/curved-linear-rail.html>. [Accessed: 04-Feb-2019].
- 16 – Proposal 3:** "Konmison Sfu1605 Linear Guide Rail 400mm Linear Motion Rail Cnc Linear Rail Shaft Guide with Stepper Motor Nema 23 Support," Dorling Kindersley Science Encyclopedia: Amazon.co.uk: Dorling Kindersley Publishing: 9780751356410: Books. [Online]. Available: <https://www.amazon.co.uk/Konmison-Sfu1605-Linear-Stepper-Support/dp/B016MC2W3C>. [Accessed: 04-Feb-2019].
- 16 – Proposal 4:** J. Petersen and L. Myers, "Wooden Train Freight Cars -- The Play Trains! Ultimate Wooden Train Guide," Play Trains!, 09-Dec-2017. [Online]. Available: <https://play-trains.com/wooden-train-freight-cars-with-removable-freight/>. [Accessed: 04-Feb-2019].
- 17 – Proposal 1:** "Online Home Store for Furniture, Decor, Outdoors & More," Wayfair. [Online]. Available: <https://www.wayfair.com/bed-bath/pdx/ben-and-jonah-clean-home-peva-single-shower-curtain-liner-banj1011.html>. [Accessed: 04-Feb-2019].
- 17 – Proposal 2:** Pam, Kristin, Toni, Hannah, Claire, C. Drake, and Denise, "DIY No-Sew Drop Cloth Curtains a CHEAP DIY Curtain Rod," The Crazy Craft Lady, 11-Jan-2019. [Online]. Available: <https://thecrazycraftlady.com/no-sew-drop-cloth-curtains/>. [Accessed: 04-Feb-2019].
- 17 – Proposal 3:** "project," Variegateinteriors FALSE CEILING AND WINDOW BLINDS. [Online]. Available: <http://www.variegateinteriors.com/false-ceiling-and-window-blinds>. [Accessed: 04-Feb-2019].

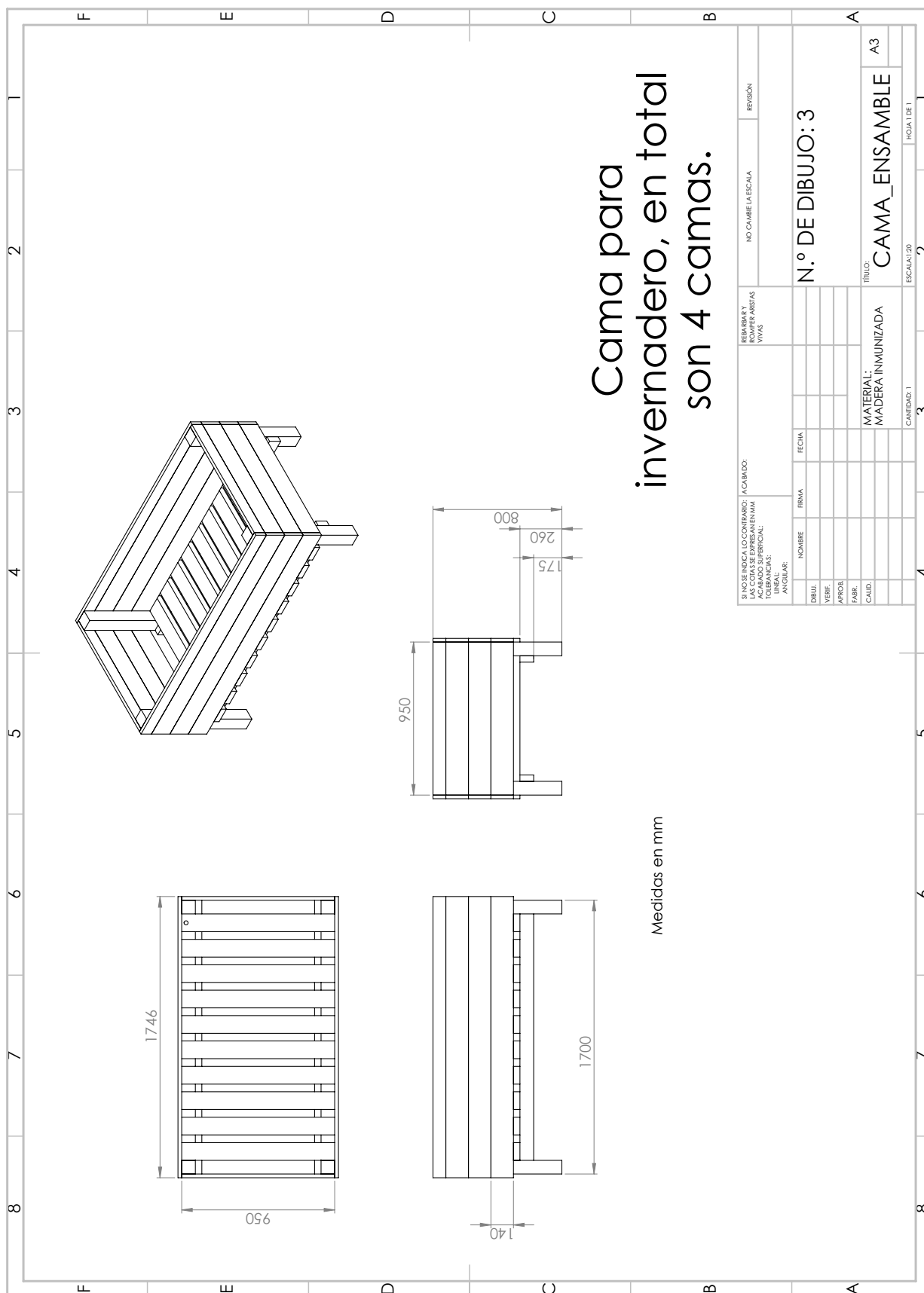
- 19 – Proposal 1:** D. Anthony, “Top 5: Best Drones for Real Estate Photography - MyDroneChoice,” My Drone Choice, 11-Jun-2017. [Online]. Available: <https://www.mydronechoice.com/best-drones-real-estate-photography/>. [Accessed: 04-Feb-2019].
- 19 – Proposal 2:** J. Erickson, “Parasitic fungi and the battle against coffee rust disease,” University of Michigan News, 13-Nov-2015. [Online]. Available: <https://news.umich.edu/parasitic-fungi-and-the-battle-against-coffee-rust-disease/>. [Accessed: 04-Feb-2019].
- 20 – Proposal 1:** “How To Install pH Sensors,” www.wateronline.com. [Online]. Available: <https://www.wateronline.com/doc/how-to-install-ph-sensors-0001>. [Accessed: 04-Feb-2019].
- 20 – Proposal 2:** Turf-Tec Soil Compaction Tester / Dial Penetrometer. [Online]. Available: <http://www.turf-tec.com/PHDS1lit.html>. [Accessed: 04-Feb-2019].
- 20 – Proposal 3:** “Measure Your Gut Health With a Stool pH Test,” Wellness Expert Specializing in Digestion & Joint Health. [Online]. Available: <https://www.drdauidwilliams.com/measure-gut-health-with-stool-ph-test>. [Accessed: 04-Feb-2019].
- 21 – Proposal 1:** “Terms & Conditions,” Skye Instruments. [Online]. Available: <http://www.skyeinstruments.com/news-events/lux-sensors/>. [Accessed: 04-Feb-2019].
- 21 – Proposal 2:** Zen Cart™ Team, “TSL2561 digital luminosity / lux / light sensor,” SK Pang Electronics Ltd. [Online]. Available: <http://skpang.co.uk/catalog/tsl2561-digital-luminosity-lux-light-sensor-p-1131.html>. [Accessed: 04-Feb-2019].
- 21 – Proposal 3:** “Diode House +918030007370”, Diode House +918030007370, 2019. [Online]. Available: <http://diodehouse.in/ldr-light-dependent-resistor-sensor-ldr/p104>. [Accessed: 04- Feb- 2019].
- 22 – Proposal 1:** “DFROBOT Gravity: Analog Capacitive Soil Moisture Sensor- Corrosion Resistant,” Amazon. [Online]. Available: <https://www.amazon.com/DFROBOT-Gravity-Capacitive-Corrosion-Resistant/dp/B01GHY0N4K>. [Accessed: 04-Feb-2019].
- 22 – Proposal 2:** R. Team, “DHT11 Temperature - Humidity Sensor Module : rhydoLABZ INDIA,” rhydoLABZ.com. [Online]. Available: https://www.rhydolabz.com/sensors-weather-sensors-c-137_147/dht11-temperature-humidity-sensor-module-p-2044.html. [Accessed: 04-Feb-2019].
- 22 – Proposal 3:** “NEW PRODUCT – Soil Temperature/Moisture Sensor – SHT10,” Adafruit Industries - Makers, hackers, artists, designers and engineers!, 14-Mar-2013. [Online]. Available: <https://blog.adafruit.com/2013/03/14/new-product-soil-temperaturemoisture-sensor-sht10/>. [Accessed: 04-Feb-2019].
- 22 – Proposal 4:** Soil Sensors. [Online]. Available: <http://www.eagle-tek.com/en/category/products/f>. [Accessed: 04-Feb-2019].
- R. Team, “DHT11 Temperature - Humidity Sensor Module : rhydoLABZ INDIA,” rhydoLABZ.com. [Online]. Available: https://www.rhydolabz.com/sensors-weather-sensors-c-137_147/dht11-temperature-humidity-sensor-module-p-2044.html. [Accessed: 04-Feb-2019].
- 23 – Proposal 1:** J. Shama, J. Shama, Forum News Service, and Forum News Service, “Tracking soil temperature an important step before planting,...”, Agweek, 21-Nov-2017. [Online]. Available: <http://www.agweek.com/news/4254715-tracking-soil-temperature-important-step-planting-experts-say>. [Accessed: 04-Feb-2019].
- 23 – Proposal 2:** violetsupply.com. [Online]. Available: <https://www.violetsupply.com/product/thermometer-humidity-gauge/>. [Accessed: 04-Feb-2019].
- 23 – Proposal 3:** R. Team, “DHT11 Temperature - Humidity Sensor Module : rhydoLABZ INDIA,” rhydoLABZ.com. [Online]. Available: https://www.rhydolabz.com/sensors-weather-sensors-c-137_147/dht11-temperature-humidity-sensor-module-p-2044.html. [Accessed: 04-Feb-2019].
- 23 – Proposal 4:** Soil Sensors. [Online]. Available: <http://www.eagle-tek.com/en/category/products/f>. [Accessed: 04-Feb-2019].
- R. Team, “DHT11 Temperature - Humidity Sensor Module : rhydoLABZ INDIA,” rhydoLABZ.com. [Online]. Available: https://www.rhydolabz.com/sensors-weather-sensors-c-137_147/dht11-temperature-humidity-sensor-module-p-2044.html. [Accessed: 04-Feb-2019].
- 24 – Proposal 1:** “Gravity: Analog Electrical Conductivity Sensor / Meter(K=10),” DFRobot. [Online]. Available: <https://www.dfrobot.com/product-1797.html>. [Accessed: 04-Feb-2019].
- 24 – Proposal 2:** Soil Sensors. [Online]. Available: <http://www.eagle-tek.com/en/category/products/f>. [Accessed: 04-Feb-2019].
- 25 – Proposal 3:** T. Online, “The new 4GEE Capture Cam means you won't miss a thing,” T3, 01-Mar-2016. [Online]. Available: <https://www.t3.com/news/the-new-4gee-capture-cam-means-you-wont-miss-a-thing>. [Accessed: 04-Feb-2019].
- 26 – Proposal 4:** “Steel metal ellipse frame border vector image on VectorStock,” VectorStock. [Online]. Available: <https://www.vectorstock.com/royalty-free-vector/steel-metal-ellipse-frame-border-vector-3509355>. [Accessed: 04-Feb-2019].
- 28 – Proposal 1:** Compugold. [Online]. Available: http://www.plaza-tecnologia.com/mxProductDetails.php?id_product=105688. [Accessed: 04-Feb-2019].
- 28 – Proposal 2:** “Kingston Digital 16GB microSDHC Class 10 UHS-I 45MB/s Read Card with SD Adapter (SDC10G2/16GB),” Amazon. [Online]. Available: <https://www.amazon.com/Kingston-Digital-16GB-microSDHC-SDC10G2/dp/B0162YQEIE>. [Accessed: 04-Feb-2019].
- 29 – Proposal 1:** “Wireless Access Point PNG icons,” Free PNG and Icons Downloads. [Online]. Available: <https://www.iconspng.com/image/9945/wireless-access-point>. [Accessed: 04-Feb-2019].
- 29 – Proposal 2:** ADXL345 - ADXL345 triple axis, digital accelerometer libr... | Mbed. [Online]. Available: <https://os.mbed.com/cookbook/Xbee-Pro>. [Accessed: 04-Feb-2019].
- 29 – Proposal 3:** “Bluetooth Logo,” 1000 Logos The Famous Brands and Company Logos in the World. [Online]. Available: <http://1000logos.net/bluetooth-logo/>. [Accessed: 04-Feb-2019].

- 30 – Proposal 1:** “How to Deal With Missing Values in SAS,” SASCRUNCH TRAINING. [Online]. Available: <https://www.sascrunch.com/dealing-with-missing-values.html>. [Accessed: 04-Feb-2019].
- 30 – Proposal 2:** “PUMAdb : Normalization Help,” Princeton University. [Online]. Available: https://puma.princeton.edu/help/results_normalization.shtml. [Accessed: 04-Feb-2019].
- 30 – Proposal 3:** “Delete Blank Rows in Excel, Remove Blank Cells in Excel,” Chandoo.org - Learn Excel, Power BI & Charting Online, 25-Jan-2010. [Online]. Available: <https://chandoo.org/wp/delete-blank-rows-excel/>. [Accessed: 04-Feb-2019].
- 31 – Proposal 1:** “Network data server 3d cloud computing concept realistic vector illustration,” City Of Forks City Limit Sign - FORKS / WASHINGTON. APRIL 13, 2017 Stock Footage Video 28374238 | Shutterstock, 04-Dec-2014. [Online]. Available: <https://www.shutterstock.com/image-vector/network-data-server-3d-cloud-computing-235210087>. [Accessed: 04-Feb-2019].
- 31 – Proposal 2:** “Production line quality control management software - SCS Concept”, SCS Concept, 2019. [Online]. Available: <http://www.scsconcept.com/production-line-quality-control-management-software/>. [Accessed: 04-Feb-2019].
- 32 – Proposal 1:** 2019. [Online]. Available: <https://am207.github.io/2017/wiki/validation.html>. [Accessed: 04-Feb-2019].
- 33 – Proposal 1:** CS231n Convolutional Neural Networks for Visual Recognition. [Online]. Available: <http://cs231n.github.io/neural-networks-1/>. [Accessed: 04-Feb-2019].
- 33 – Proposal 2:** Mindmajix, “Learn How Support Vector Machine Algorithm Works With Examples,” Mindmajix, 11-Nov-2017. [Online]. Available: <https://mindmajix.com/support-vector-machine-algorithm>. [Accessed: 04-Feb-2019].
- 33 – Proposal 3:** “How Random Forests Can Keep You From Decision Tree,” Sefik Ilkin Serengil, 12-Mar-2018. [Online]. Available: <https://sefiks.com/2017/11/19/how-random-forests-can-keep-you-from-decision-tree/>. [Accessed: 04-Feb-2019].
- 33 – Proposal 4:** Kassambara, “CLARA - Clustering Large Applications,” Correlation matrix : A quick start guide to analyze, format and visualize a correlation matrix using R software - Easy Guides - Wiki - STHDA, 09-Apr-2017. [Online]. Available: <http://www.sthda.com/english/articles/tag/pam-clustering/>. [Accessed: 04-Feb-2019].
- 34 – Proposal 1:** “Hinge loss,” Wikipedia, 15-Jan-2019. [Online]. Available: https://en.wikipedia.org/wiki/Hinge_loss. [Accessed: 04-Feb-2019].
- 34 – Proposal 2:** S. Singla and S. Singla, “Experiments with a New Loss Term Added to the Standard Cross entropy,” medium.com, 09-Sep-2017. [Online]. Available: <https://medium.com/mlreview/experiments-with-a-new-loss-term-added-to-the-standard-cross-entropy-85b080c42446>. [Accessed: 04-Feb-2019].
- 34 – Proposal 3:** MosheWorld, “MosheWorld/Machine-Learning,” GitHub. [Online]. Available: <https://github.com/MosheWorld/Machine-Learning>. [Accessed: 04-Feb-2019].
- 35 – Proposal 1:** Sinteza Agro - WHY NOT?[Online]. Available: https://whynot.com.ua/en/portfolio/branding/sinteza_agro/. [Accessed: 04-Feb-2019].
- 35 – Proposal 2:** “How to Setup Email Notification,” Geotab Blog. [Online]. Available: <https://www.geotab.com/video/setup-email-notification/>. [Accessed: 04-Feb-2019].
- 35 – Proposal 3:** “How mobile apps are helping agriculture in achieving sustainable development?,” SourceTrace Systems, 26-Nov-2018. [Online]. Available: <http://www.sourcetrace.com/mobile-apps-for-agriculture/>. [Accessed: 04-Feb-2019].
- 35 – Proposal 4:** “Developer Portal | Developer Portal : ThingWorx,” Brand. [Online]. Available: <https://developer.thingworx.com/>. [Accessed: 04-Feb-2019].

**D ADVANTAGES, DISADVANTAGES AND
SOLUTIONS FOR BOTH CONCEPTS**

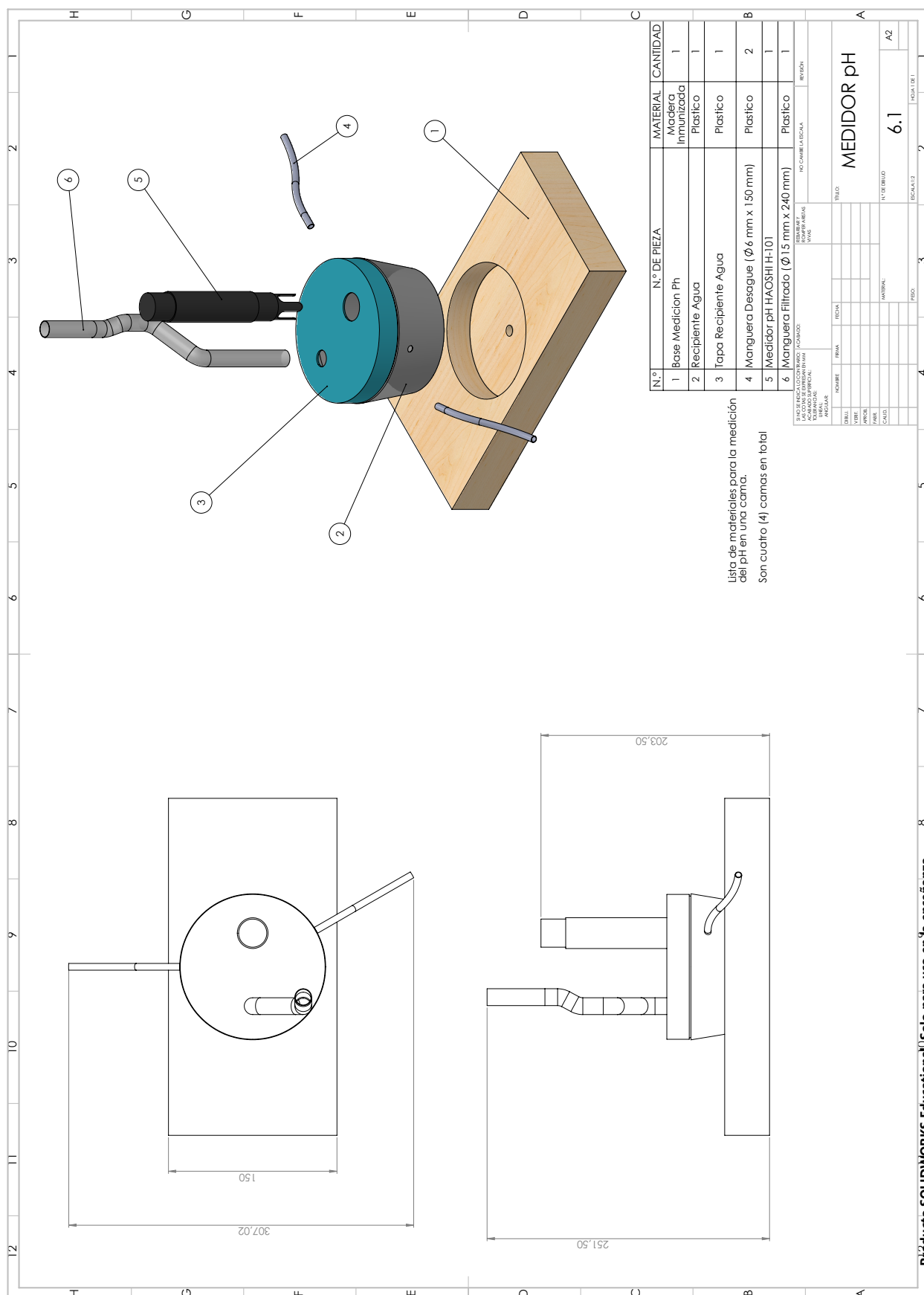
Concepts		
	Concept 1	Concept 2
Advantages	* Normal pressure from aqueduct system is easy to implement (4). * Raised beds facilitate water drainage (6). * Raised beds isolate lot soil from the greenhouse's (6). * Rain simulation system is easy to implement (8). * Independent lots are simple to assemble (11). * Electromagnetic flow meter measurements are reliable (14). * Stepper motors have inherent ability to control position (15). * Holes in tubes raised from the lots are easy to implement and are a good approximation of the rainfall (18). * pH measurement through drainage allows data acquisition automation (20). * The 5TE sensor measures the soil's humidity, temperature and electrical conductivity directly (22, 23). * The rotary arm provides a consistent position of the multispectral camera over the lots (25). * The elliptical rail is easy to construct (26). * One data package for the entire crop with lot ID requires only one connection to the Internet (27). * Pre-processing the data using normalization speeds up the creation of the model and reduce the biases due to high values (30). * A remote database outside the remote server offers relevant support for data management (31). * Balancing across classes and shuffling reduce bias of the diagnosis model (32). * Using a Deep Neural Network add non-linearity to the generated diagnosis model (33). * A web interface is controllable, extensible and freely customizable (35).	* The water pressure can be adjusted (4). * Raised beds do not deteriorate easily in time (6). * Raised beds isolate lot soil from the greenhouse's (6). * Rain simulation system properly irrigates the plants (8). * Servomotors allow to control the position, velocity and acceleration in a precise manner (15). * Sprinklers above the lots guarantee proper irrigation of the plants (18). * pH measurement through soil is optimal for crop conditions (20). * The 2-axis rail allows a precise positioning of the RGB camera (26). * If the Internet connection fails, some lot's data can however be persisted (27). * Pre-processing the data setting all missing values to zero is fast (30). * A local database on the remote server allows fast connections (31). * Random data division is fast and easy to implement (32). * The implementation of Support Vector Machines is simple and the diagnosis model's output can be satisfactory (33). * The visualization of results using Thingworx is easy and can be implemented fast (35).
Disadvantages	* The water pressure may not be enough (4). * Raised beds must withstand the lot's weight (6). * Rain simulation system may not produce a consistent rainfall (8). * Wood can deteriorate in time (9). * Holes in tubes raised from the lots might not be sufficient for proper irrigation of the plants (18). * pH values obtained in the water might differ from the soil's (20). * The rotary arm might collide with the crop's structure (25). * The elliptical rail might not allow a precise position for the RGB camera (26). * If the Internet connection fails, the whole data package cannot be persisted (27). * Balancing across classes and shuffling can take significant processing time (32). * A Deep Neural Network is processing-intensive and has high technological requirements (33). * Developing a web interface takes more time and needs more resources (35).	* Using a pumping device consumes more energy (4). * Raised beds must withstand the lot's weight (6). * Plastic raised beds are difficult to construct (9). * One single piece divided does not fit through the greenhouse's door (11). * Data acquisition using the pluviometer cannot be automated (14). * Sprinklers above the lots can negatively affect multispectral captures (18). * pH measurement through soil doesn't allow data acquisition automation (20). * The DHT-33 sensor cannot be buried, therefore can't directly measure the soil's temperature and humidity (22, 23). * With the single rail, the multispectral camera's position over the lots may be altered (25). * Sending one data package per lot requires multiple connections to the Internet (27). * Converting all missing values to zero can negatively affect the distribution of the data (30). * Random data division increases the bias of the diagnosis model, therefore negatively affects the predictions (32). * Support Vector Machines present difficulties with multiclass classification (33).
Solutions	* The water pressure can be increased using a pumping device (4). * Use a transversal element to strengthen the raised bed's structure (6). * Water repellent wood preservatives can be used (9). * Tube's height can be elevated (18). * The rotary arm can be elevated (25). * Make a protocol for reconnecting to the Internet in case of failures (27). * A web programming framework can speed up the implementation of the web interface (35).	* Use a transversal element to strengthen the raised bed's structure (6). * Use an additional sensor (e.g. ultrasonic sensor) to digitize pluviometer data (14). * Sprinklers above the lots can be displaced when capturing multispectral images (18). * Isolate the DHT-33 sensor in order to bury it in the soil (22, 23). * Add mechanical stoppers at the capturing positions (25). * A Recurrent Neural Network can be used to predict the missing values and fill them in the data (30).

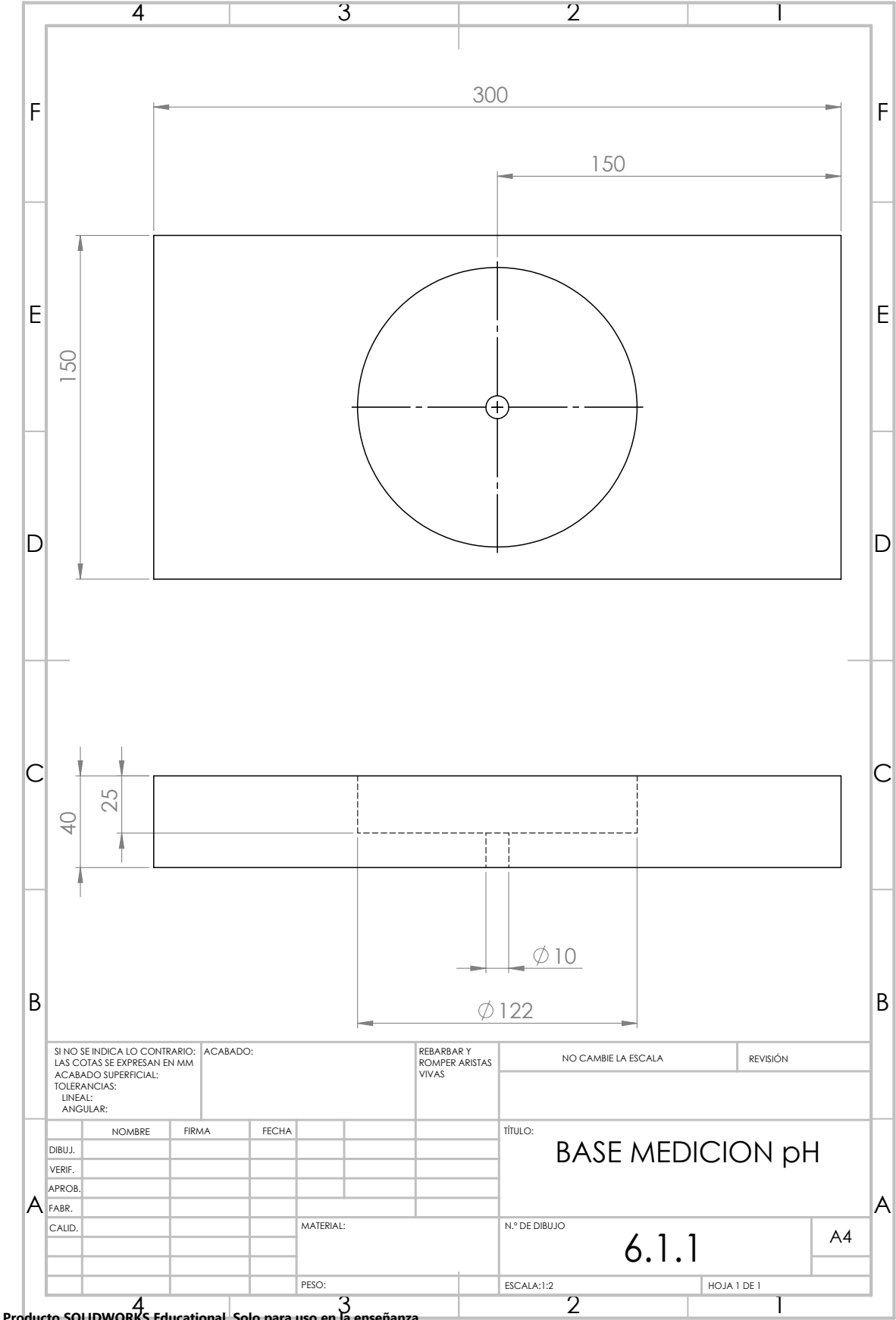
E COFFEE CROP LOTS

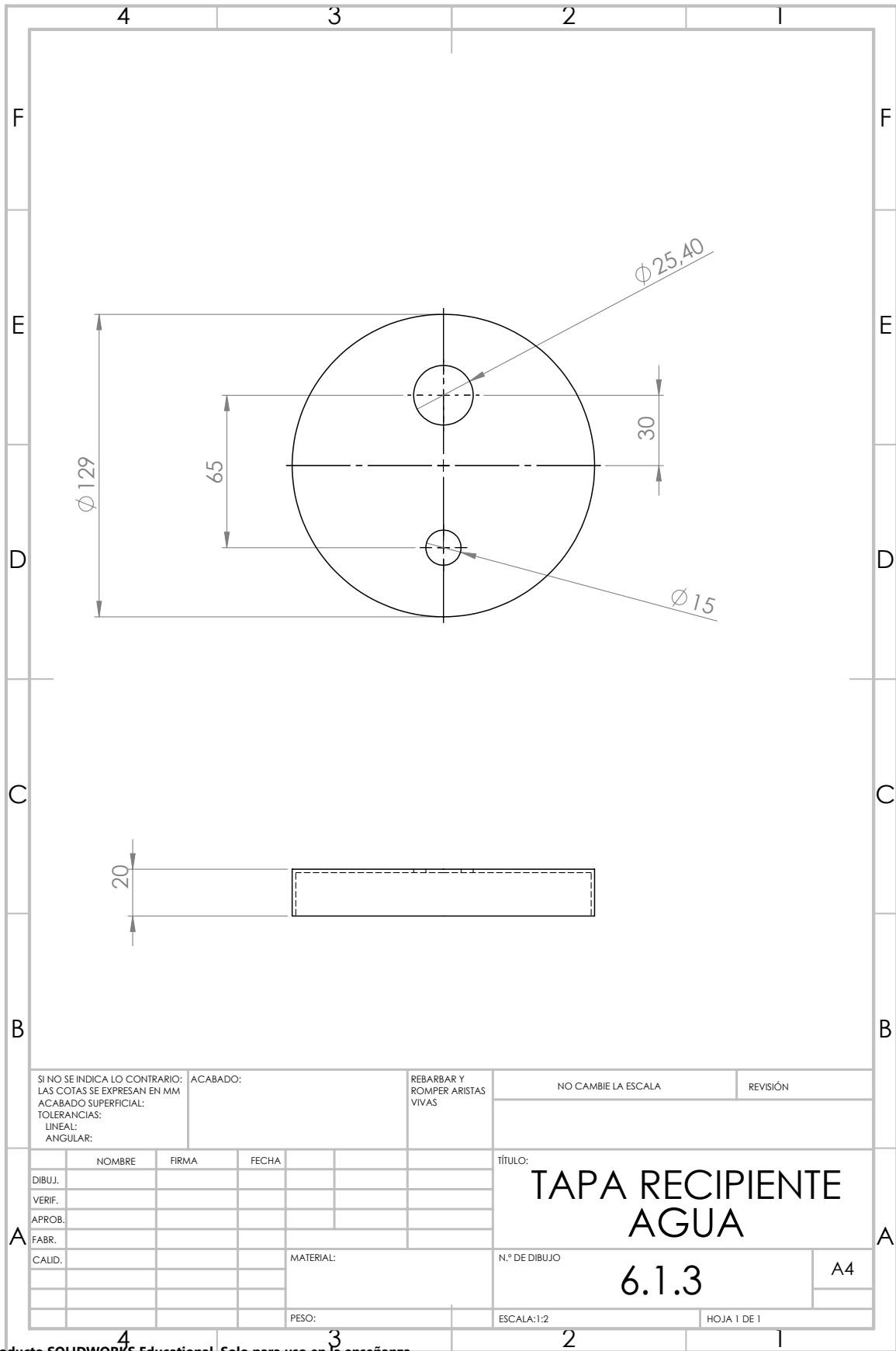


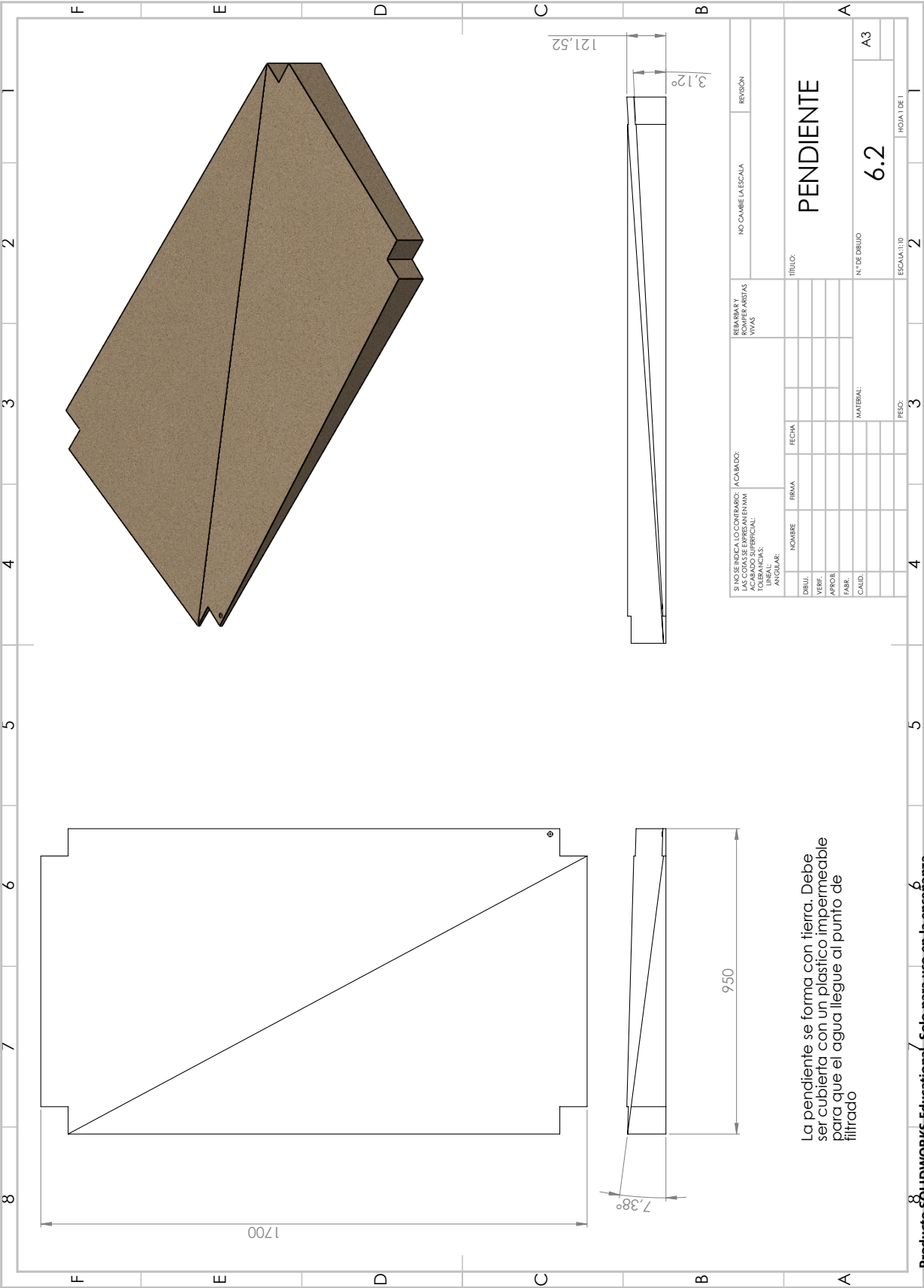
[illegible]

F PH MEASUREMENT SYSTEM

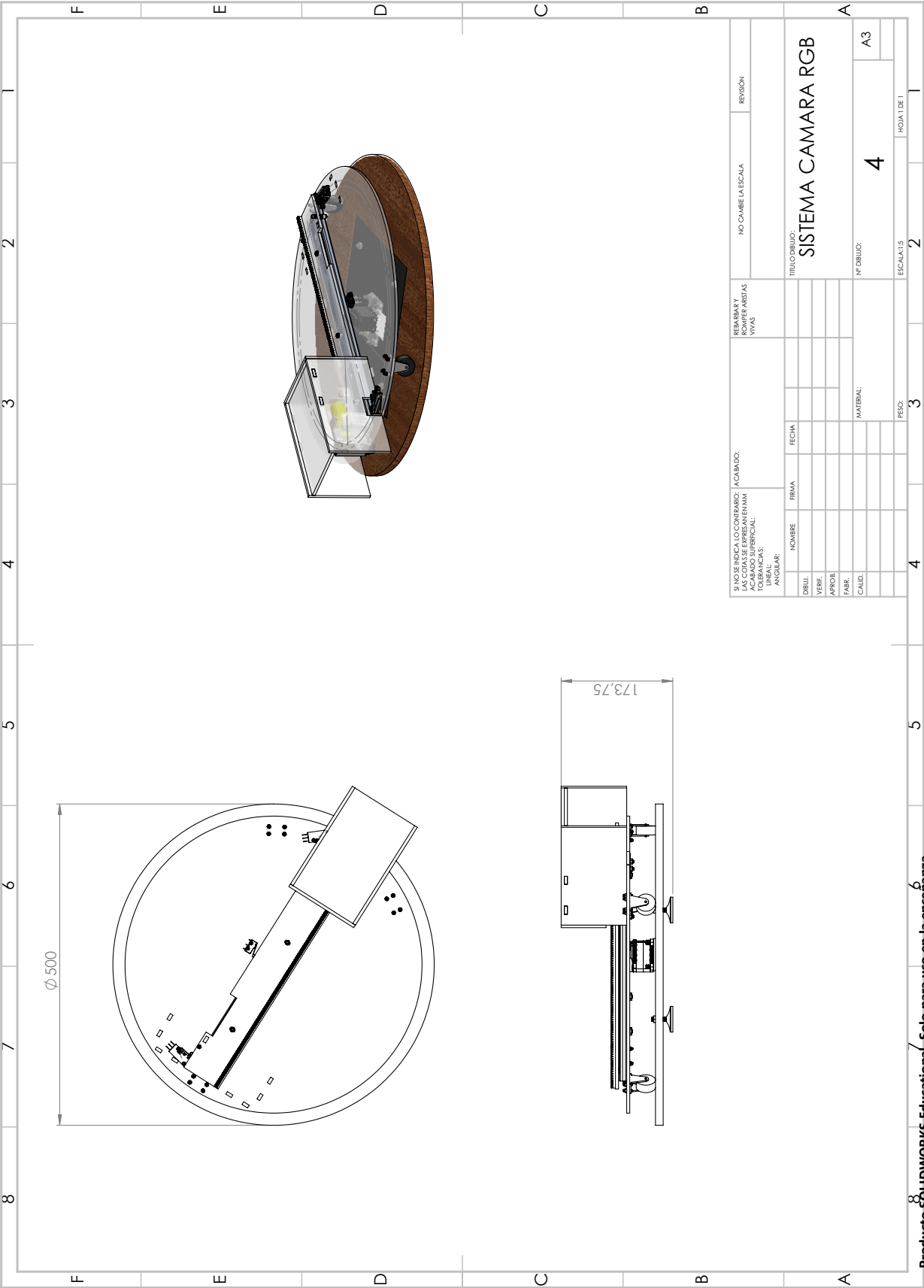




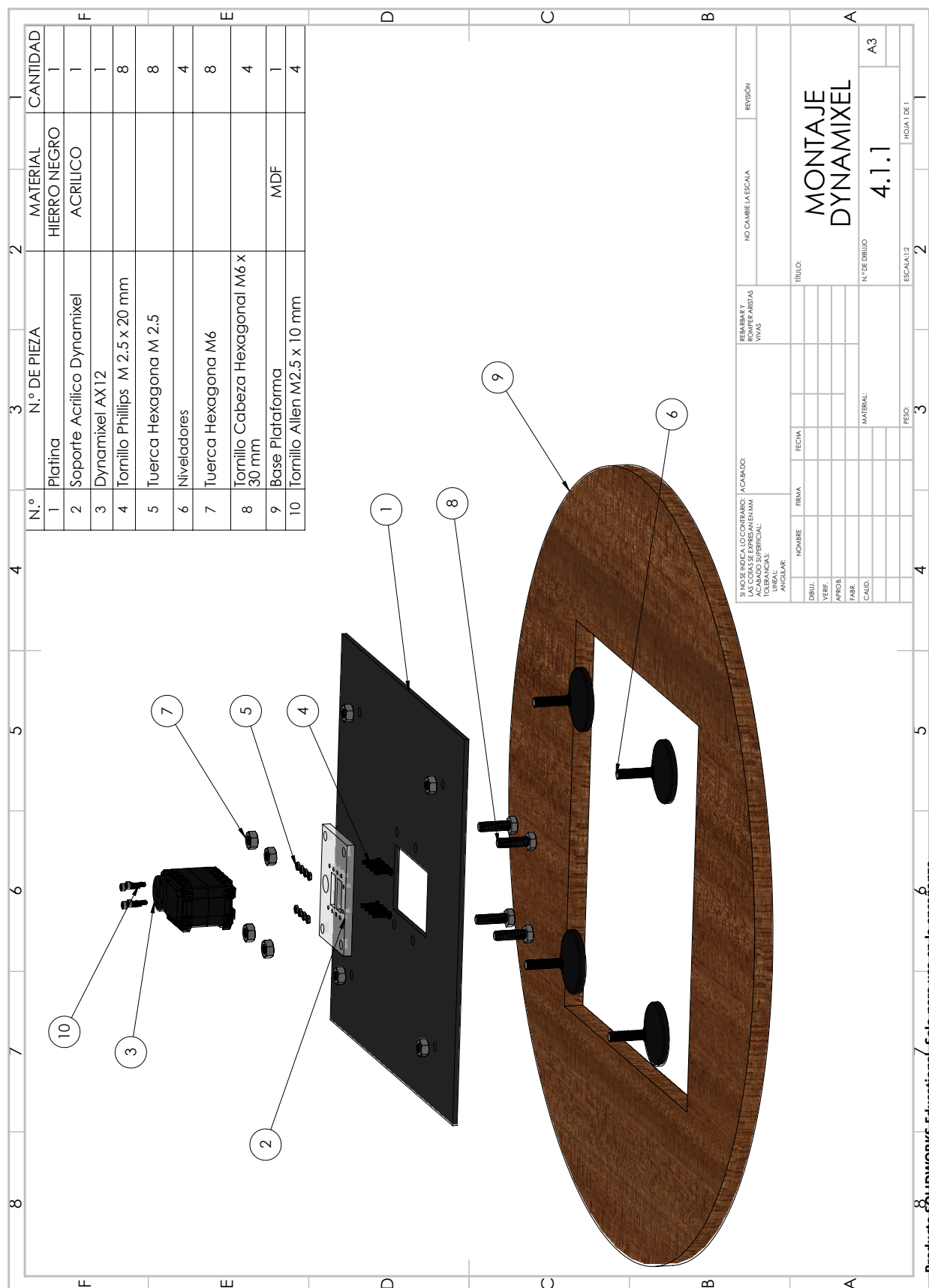


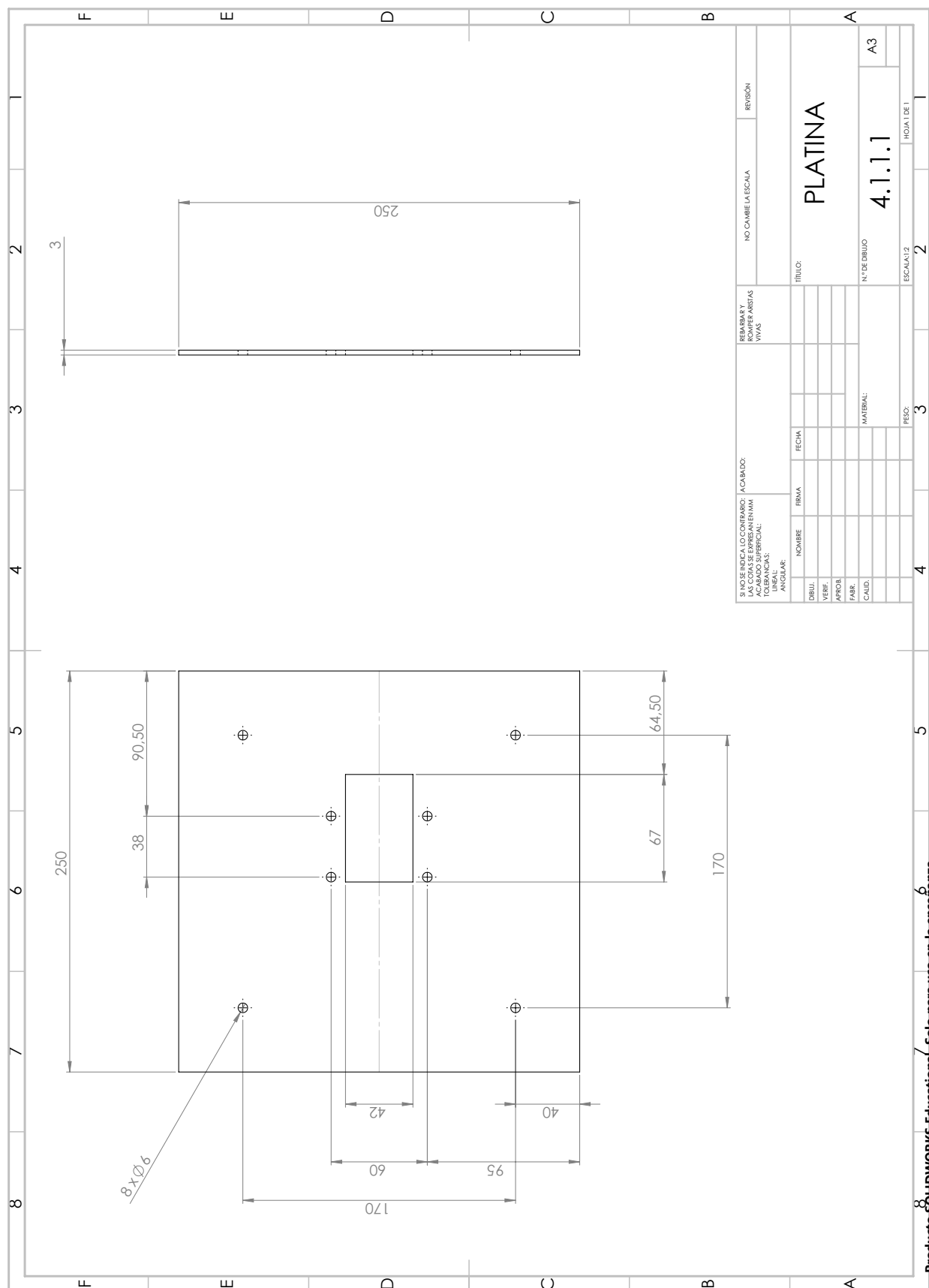


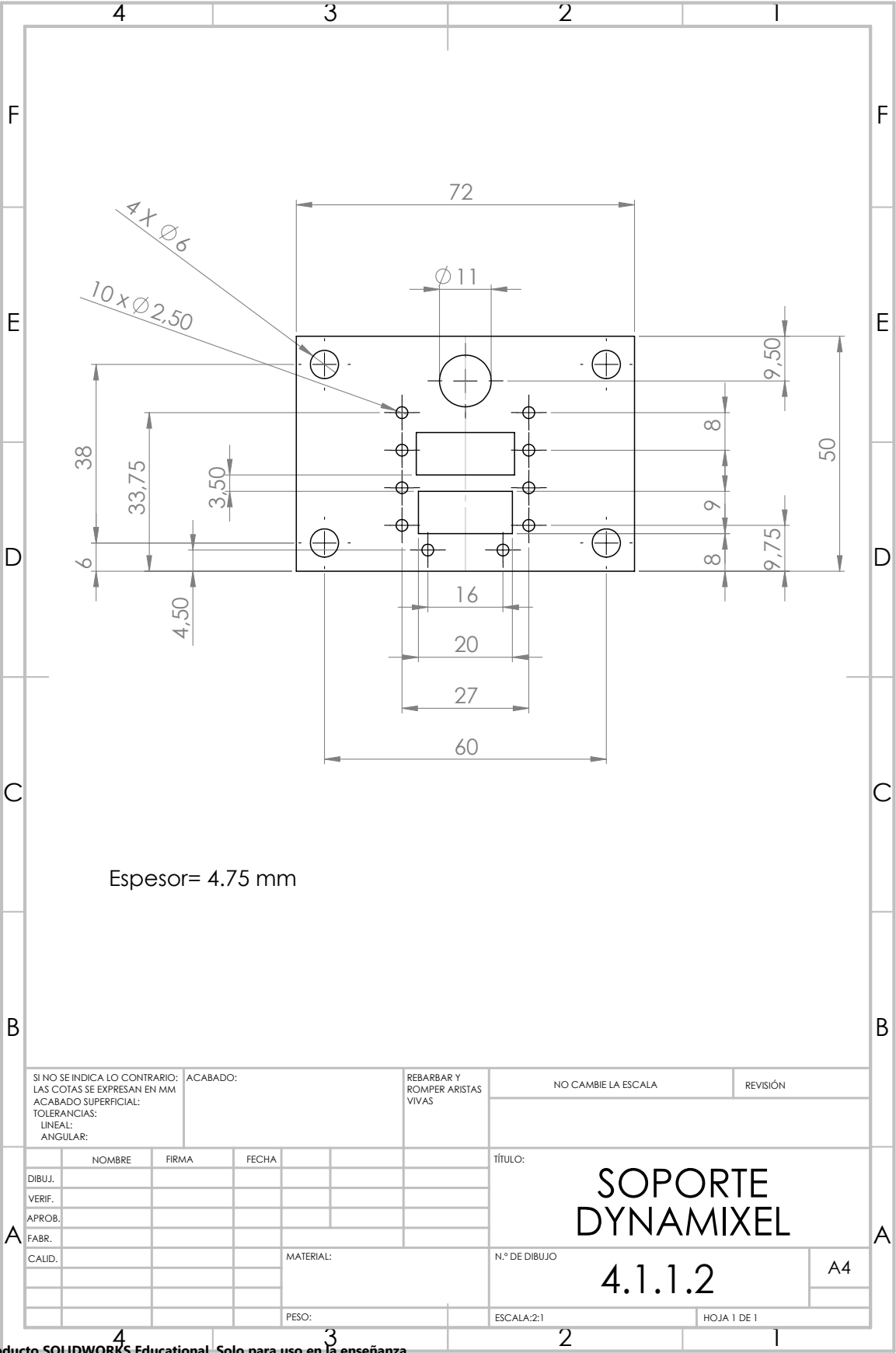
G ROTARY PLATFORM

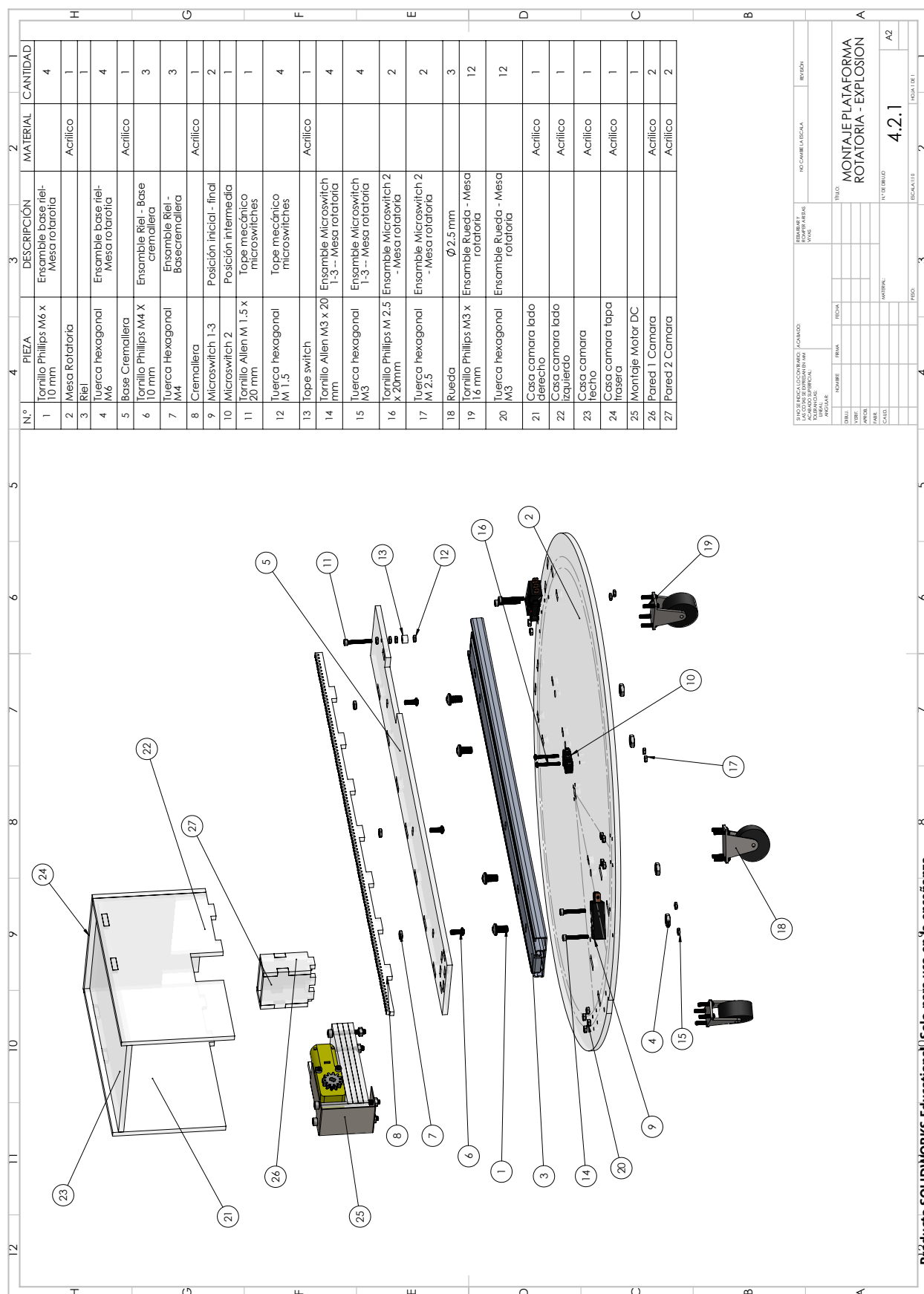


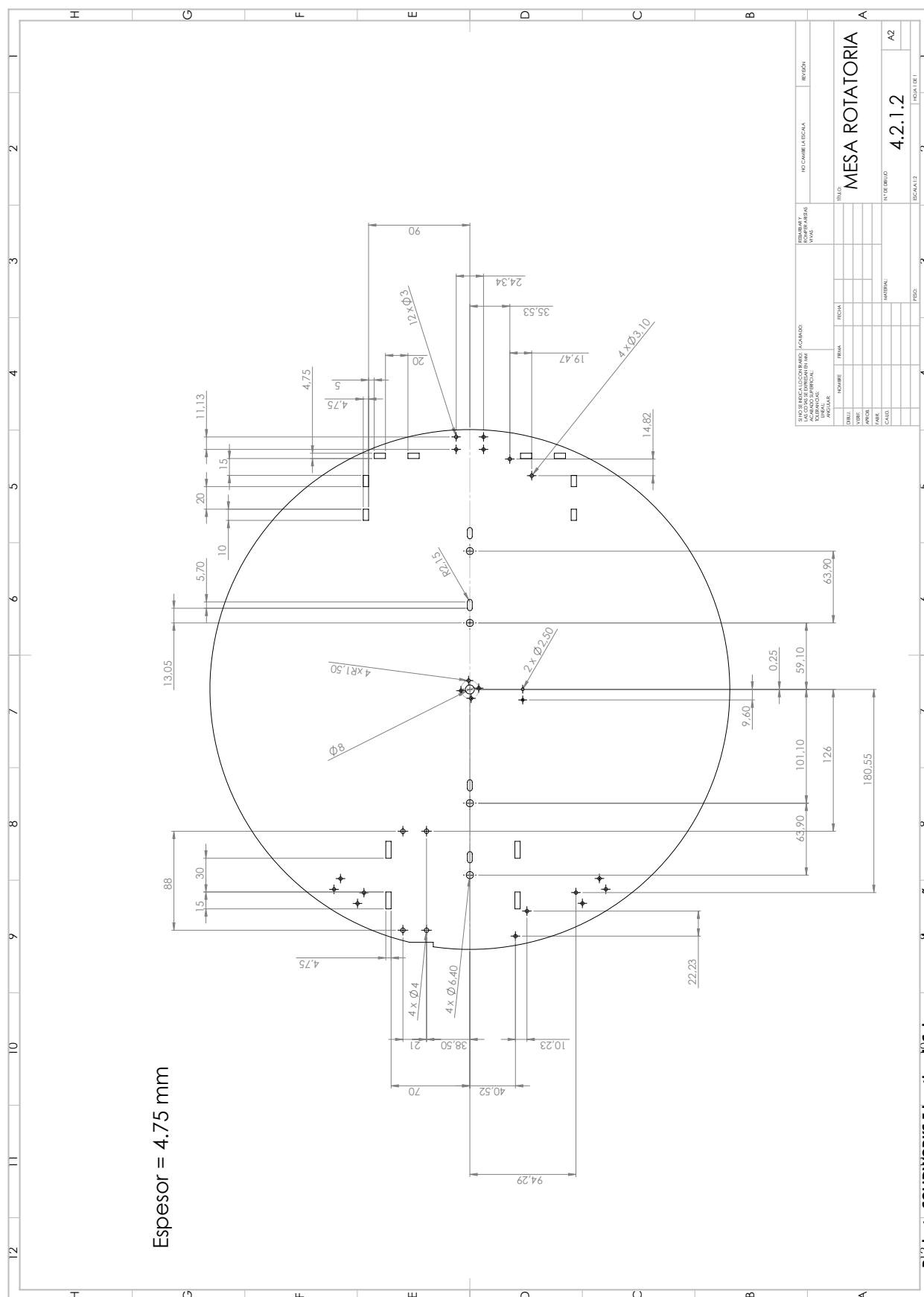


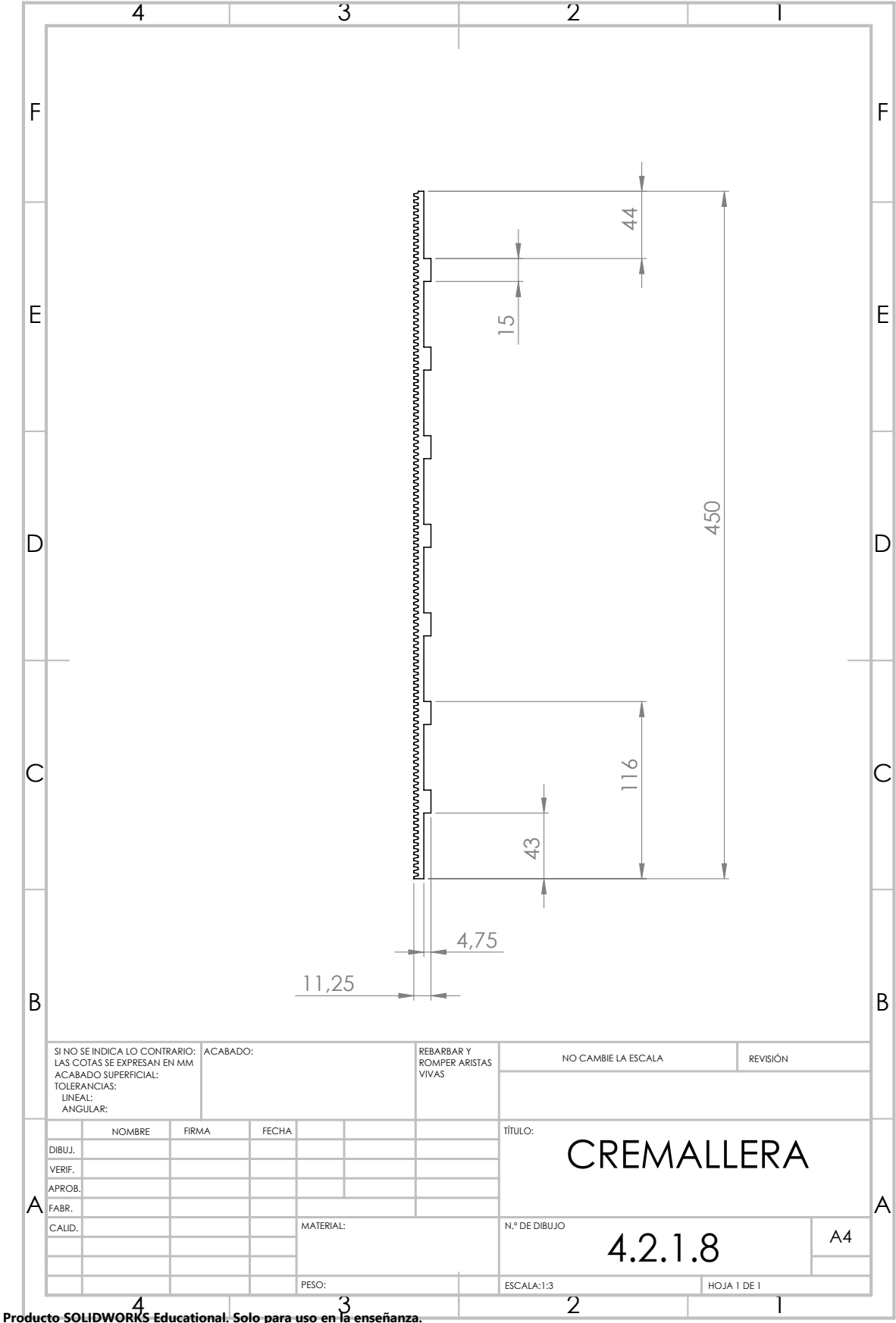


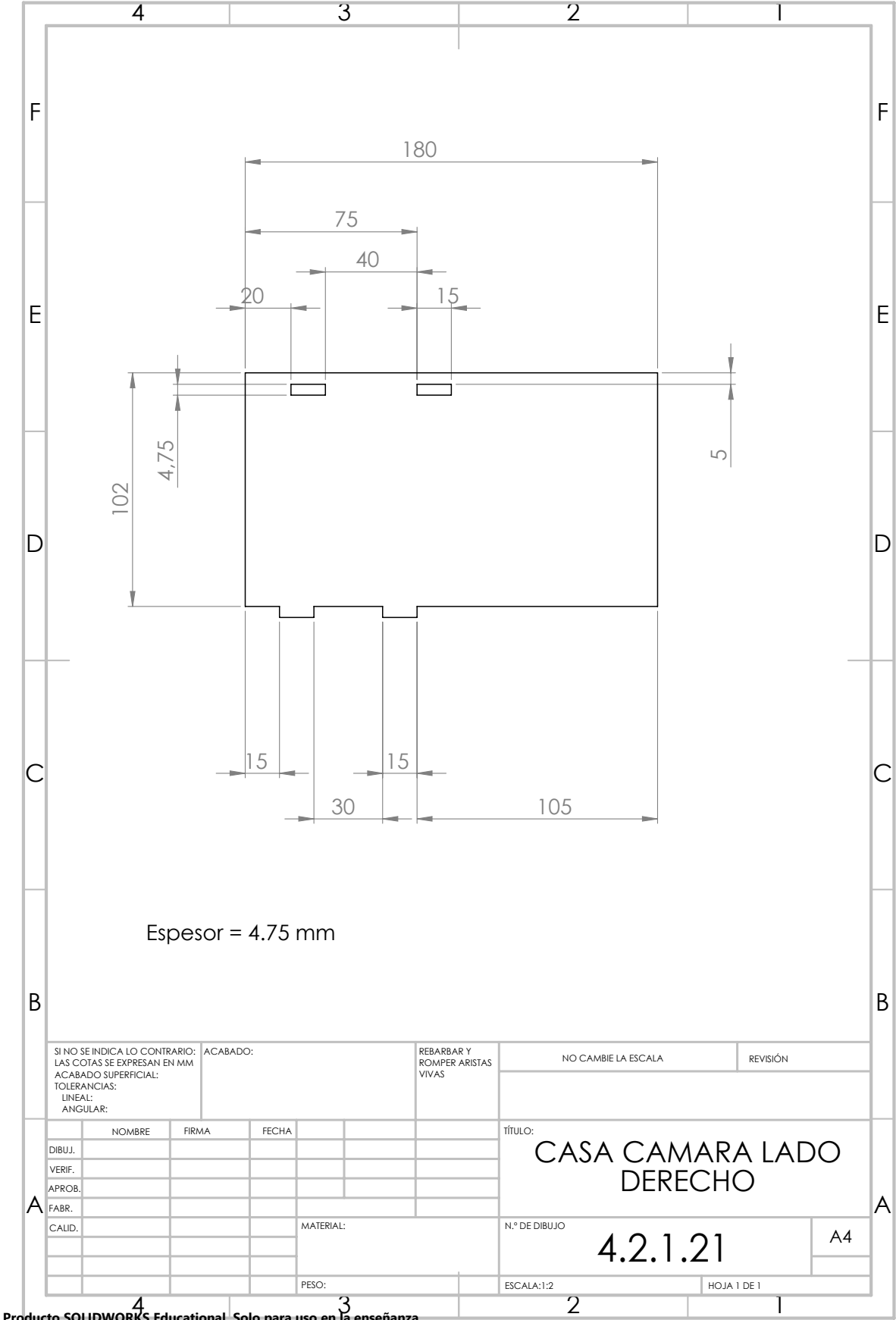


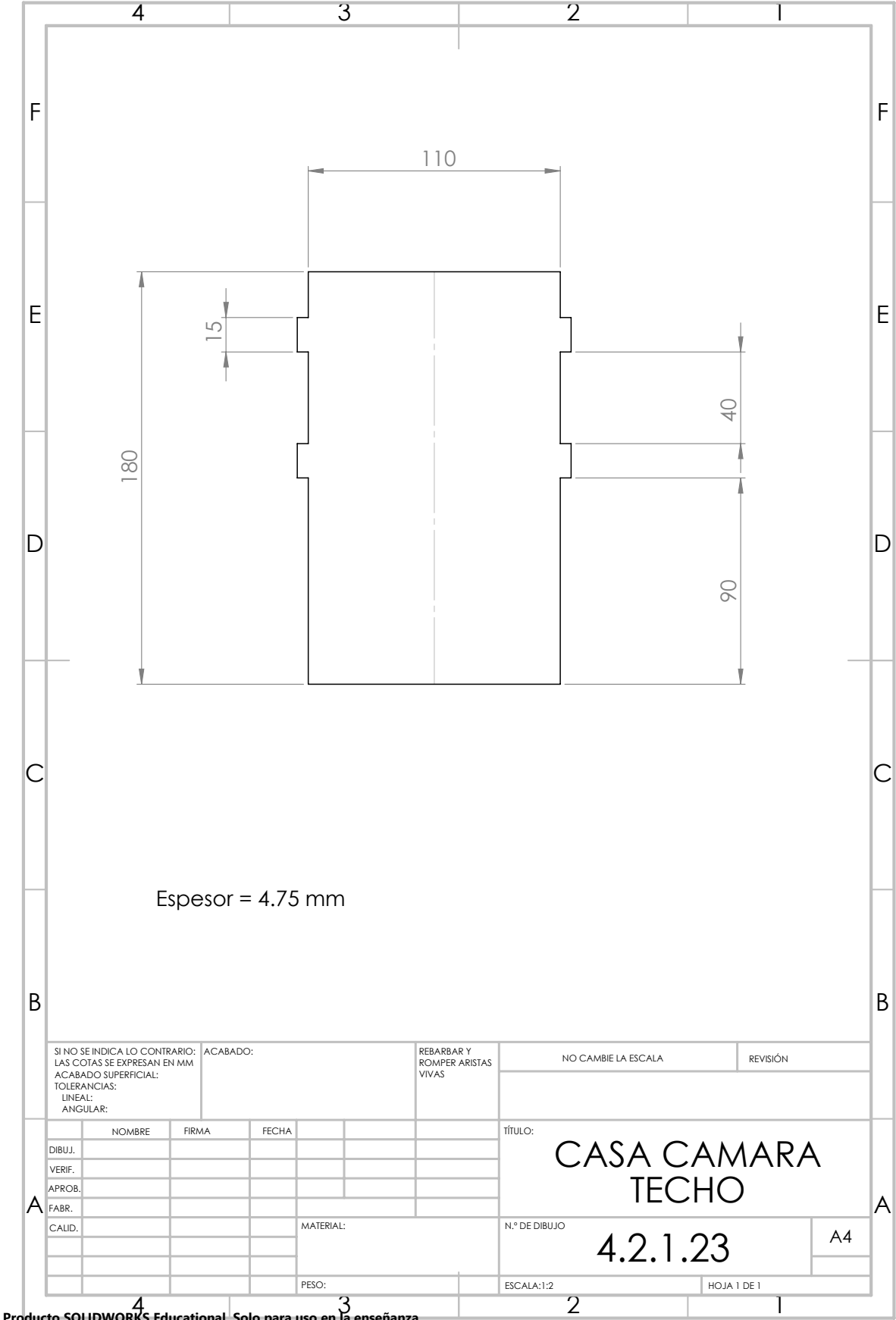


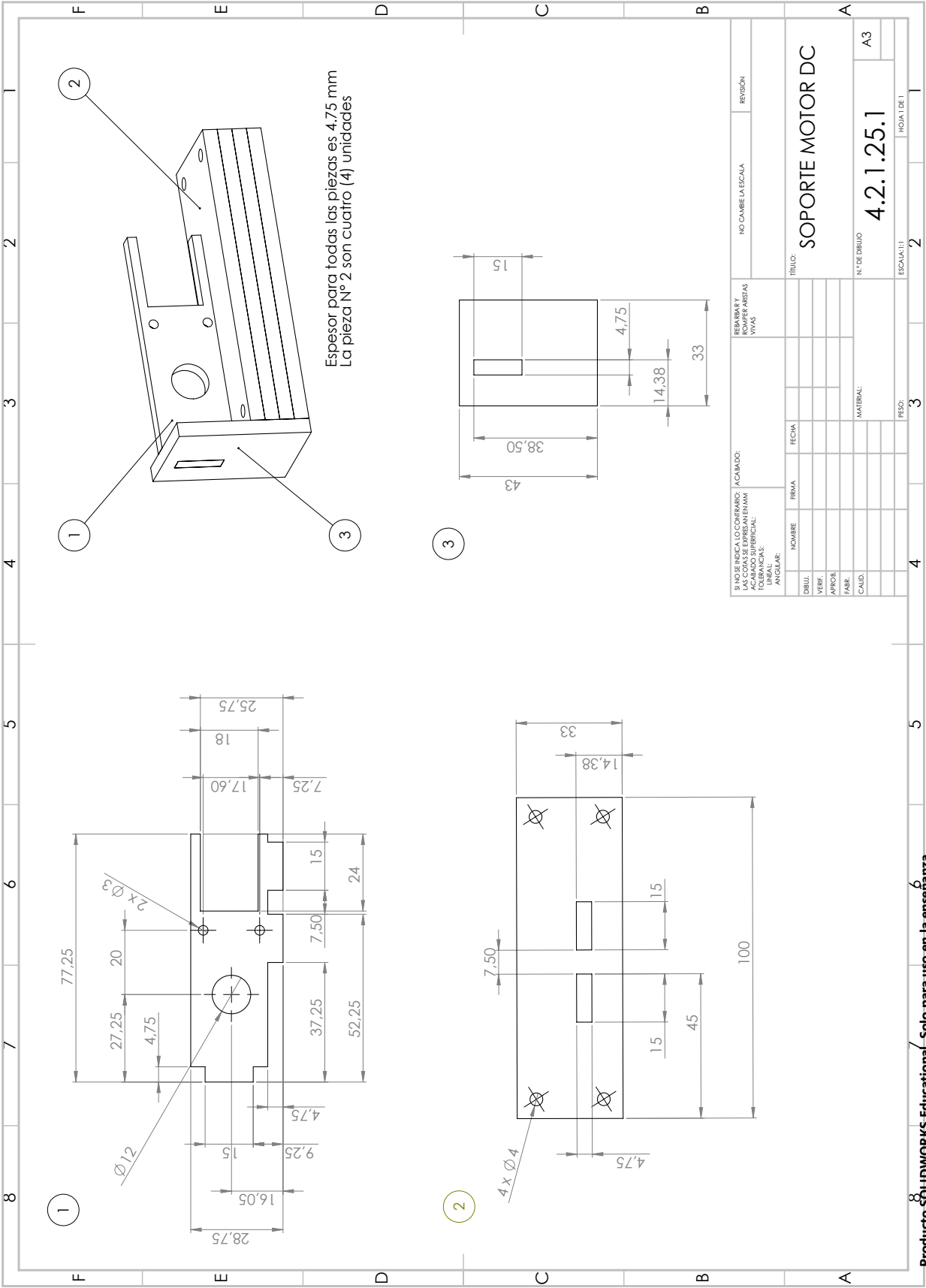


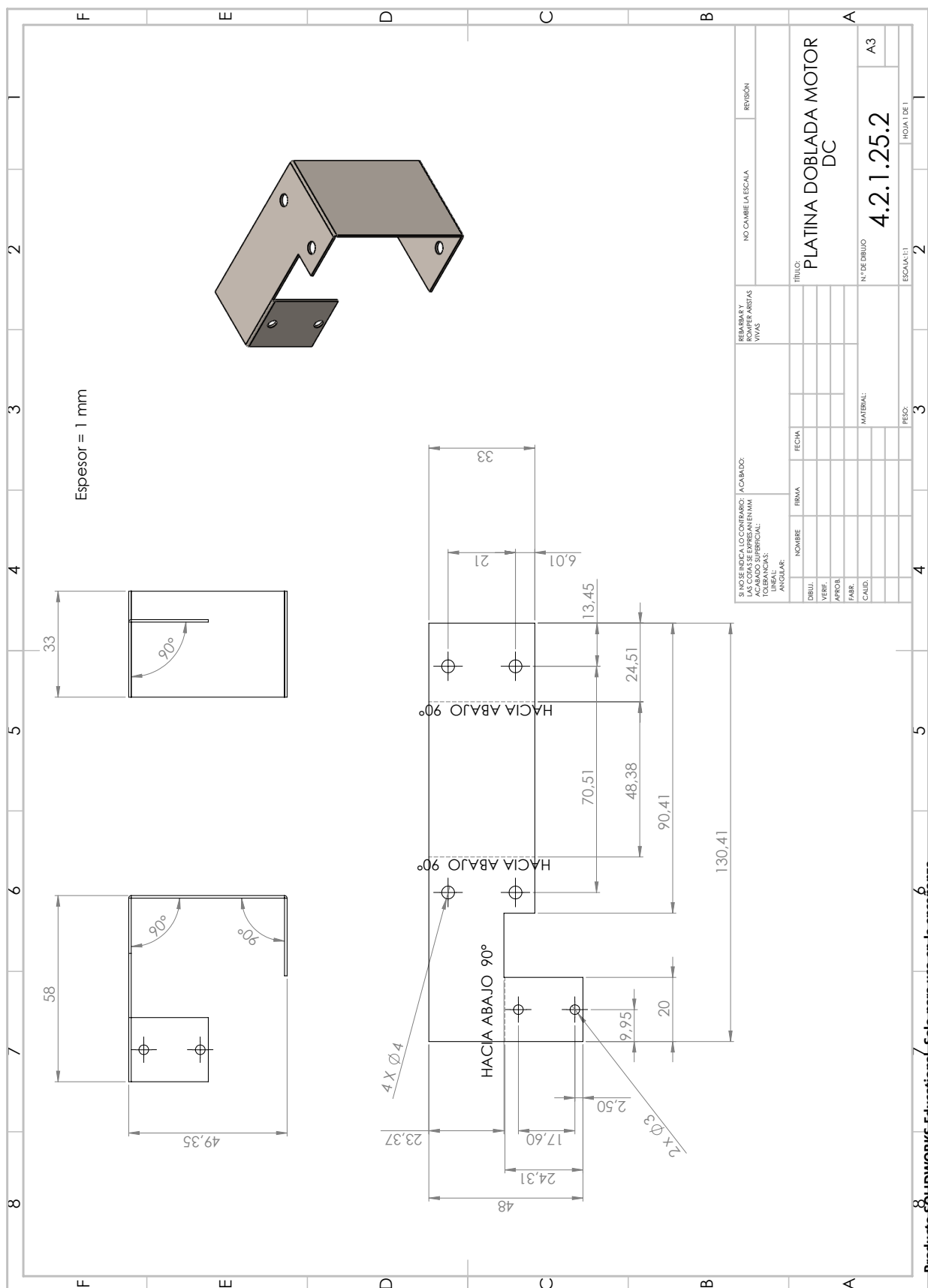


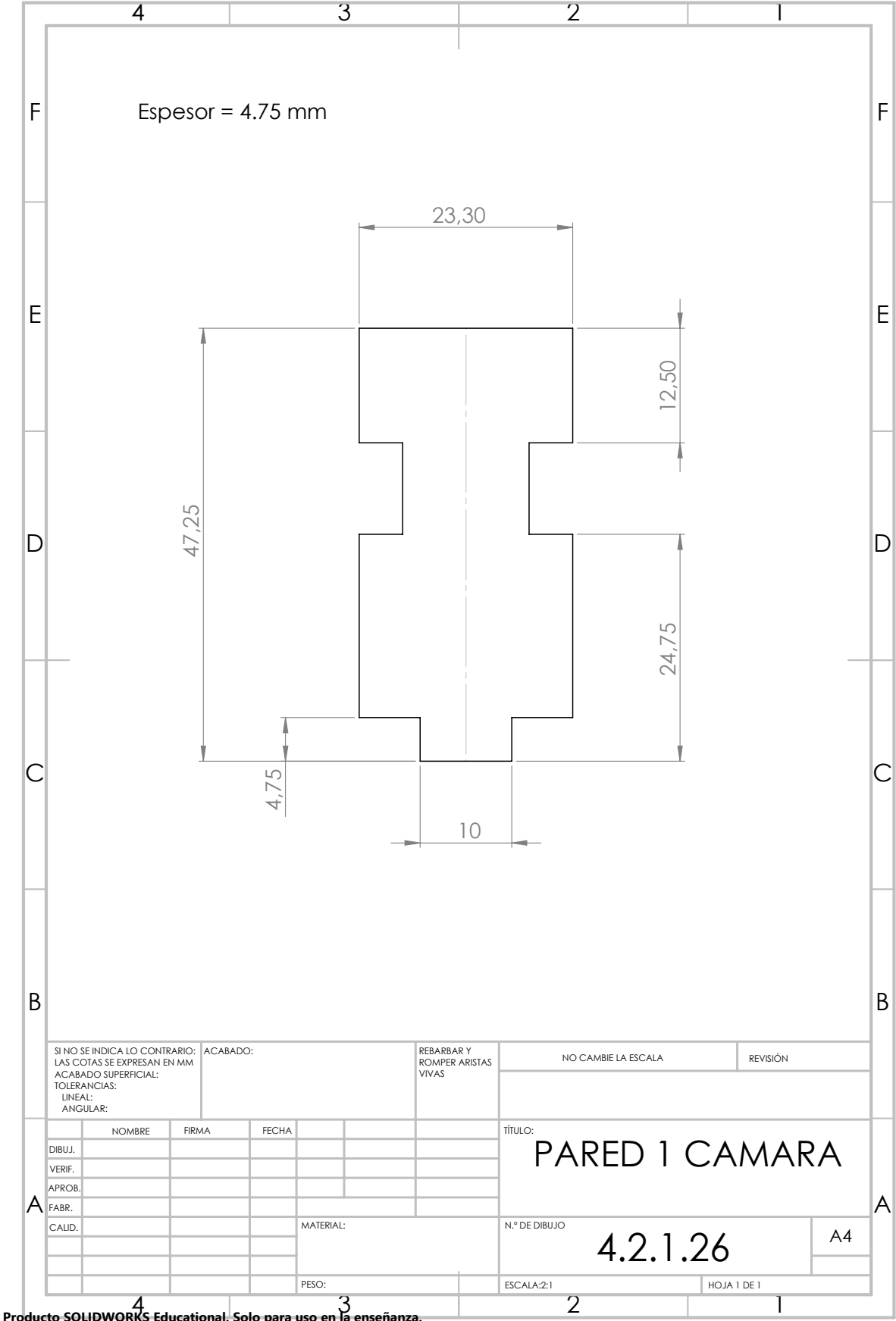


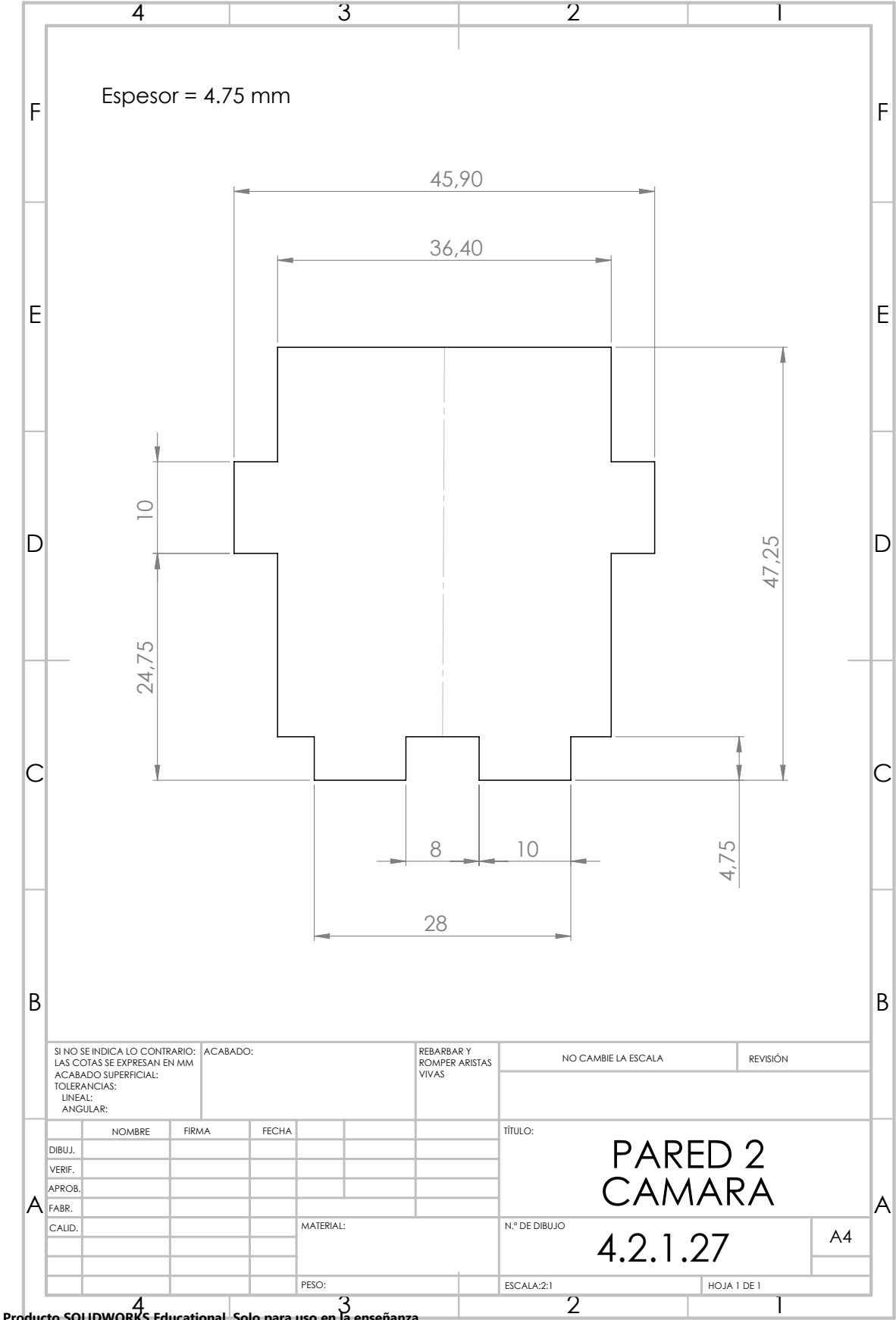




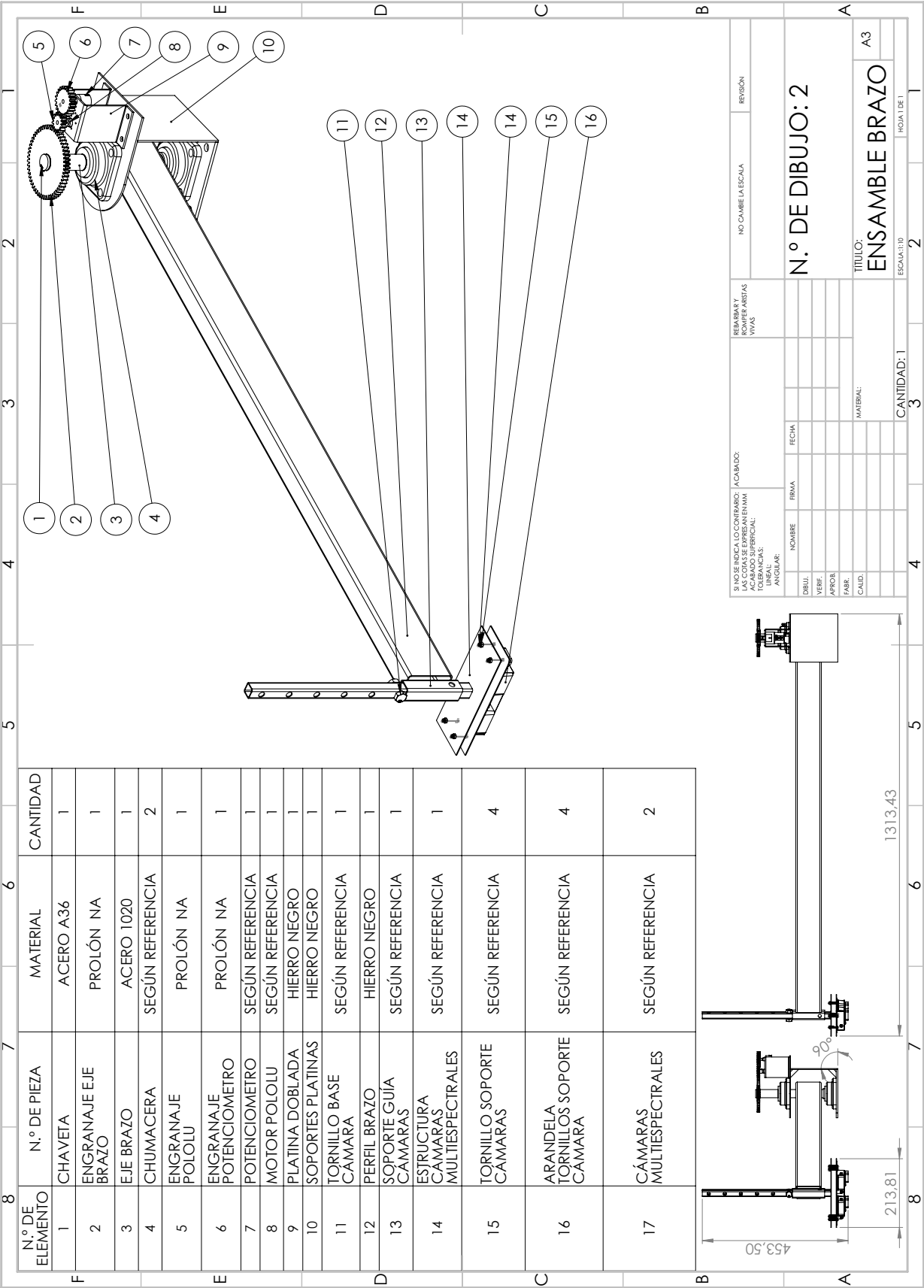


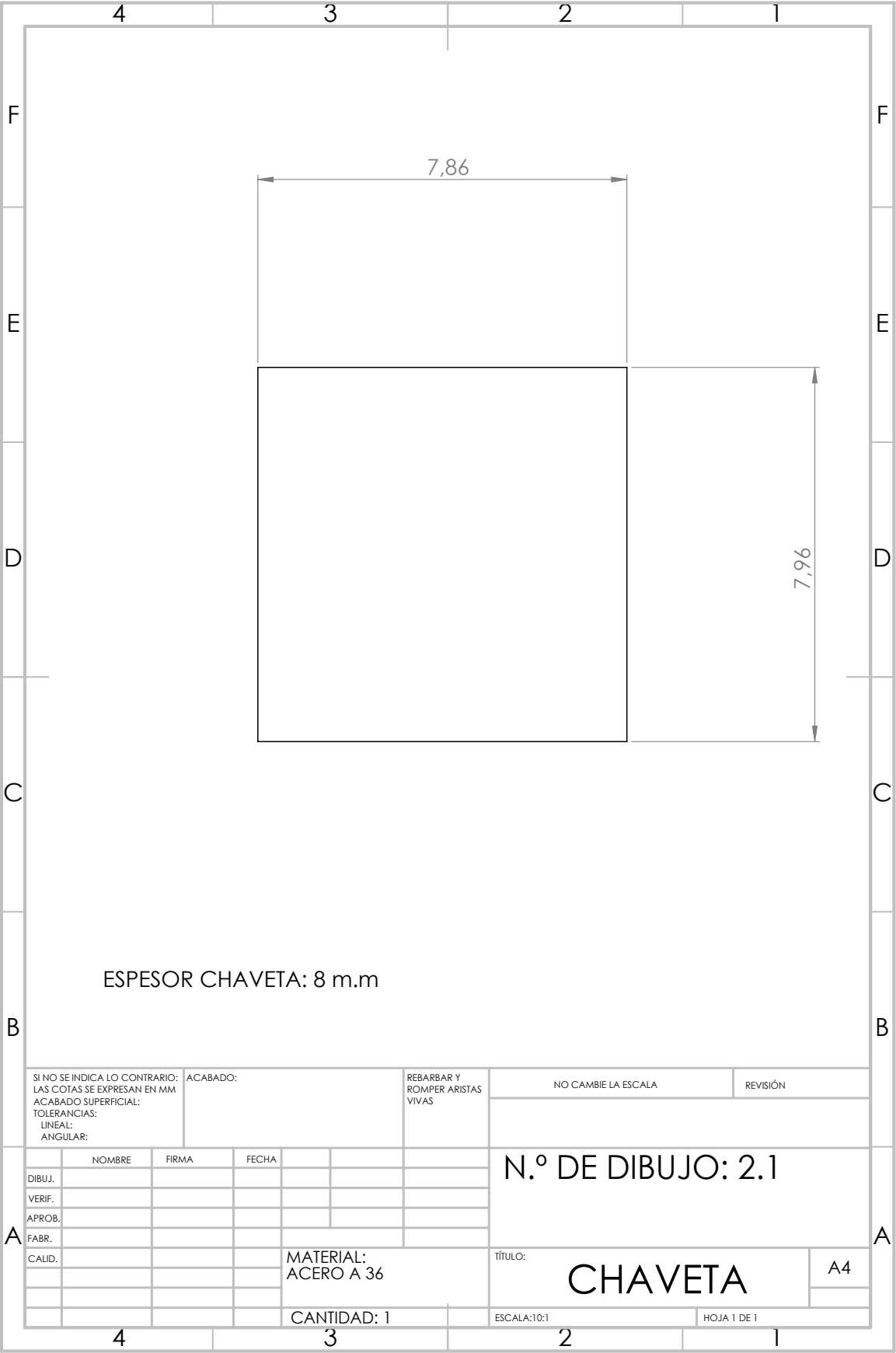


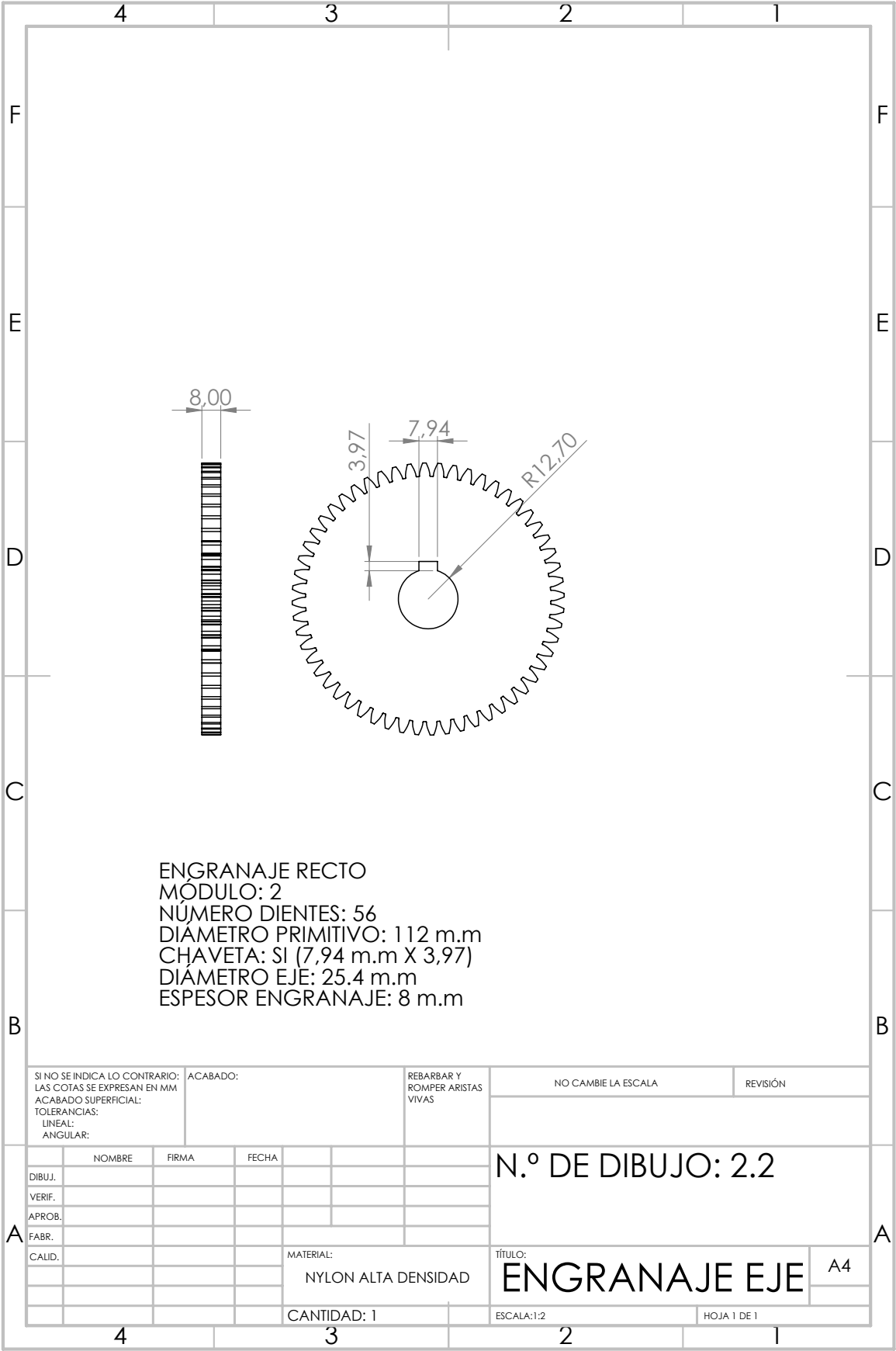


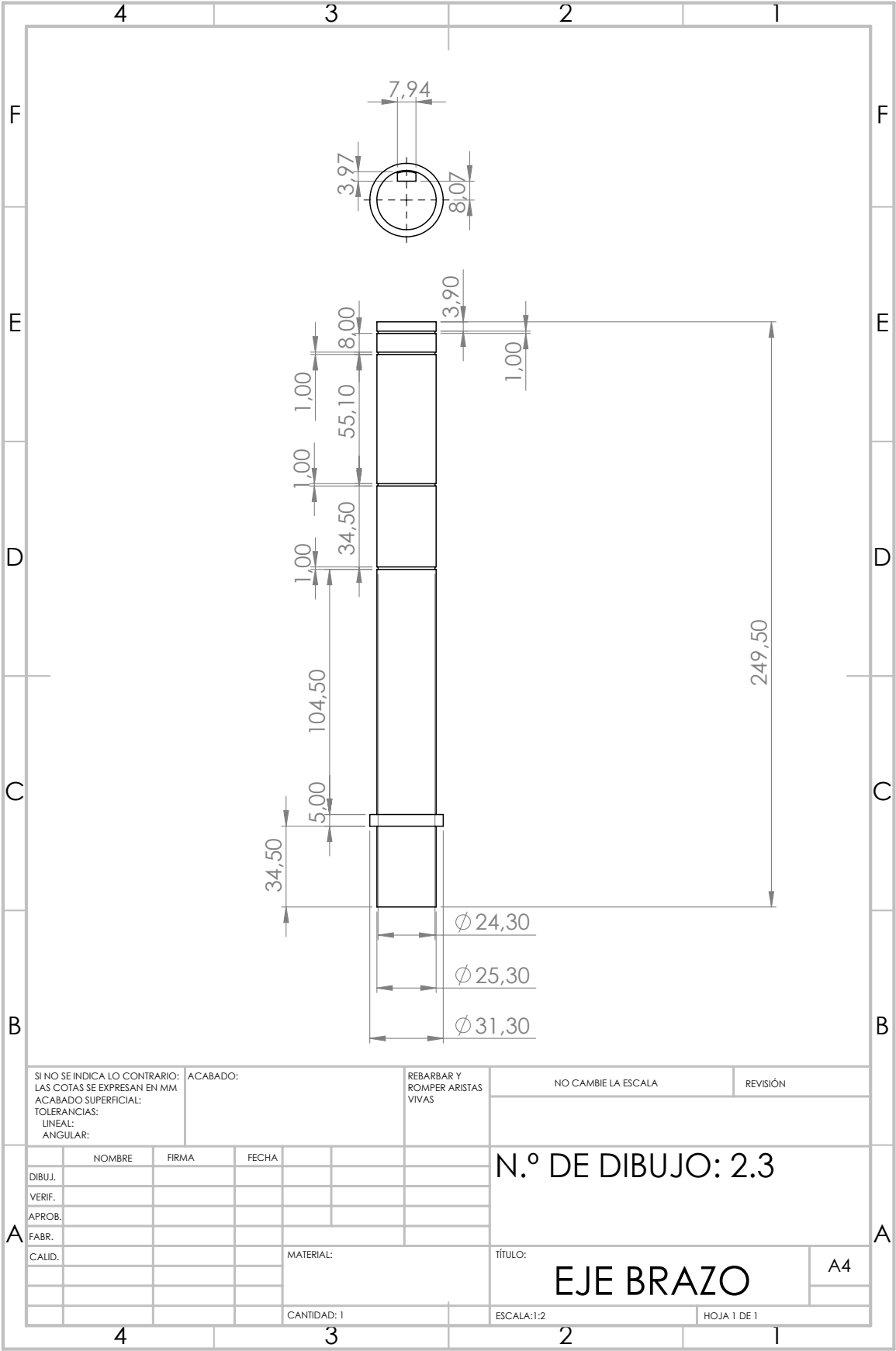


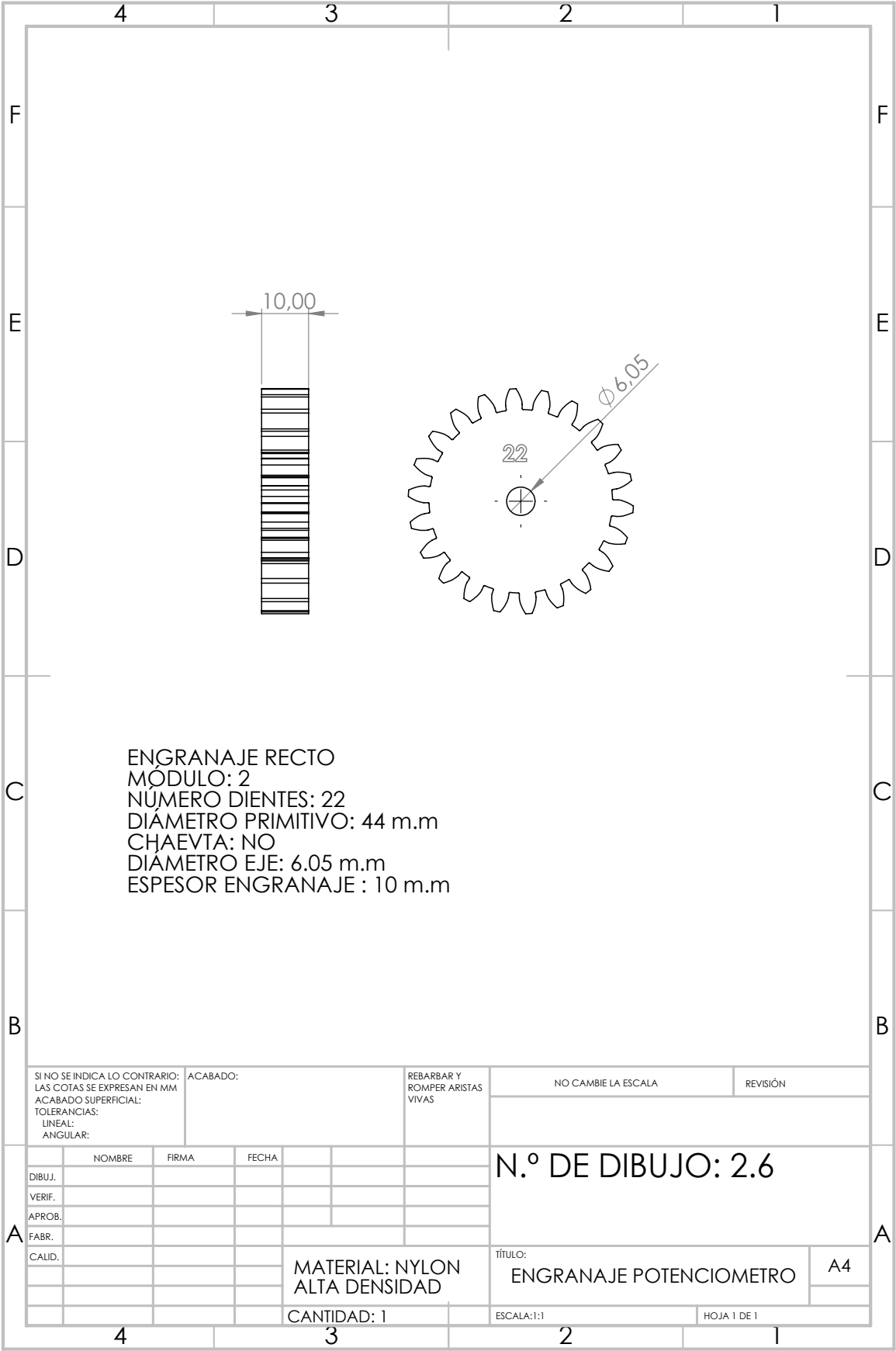
H ROTARY ARM

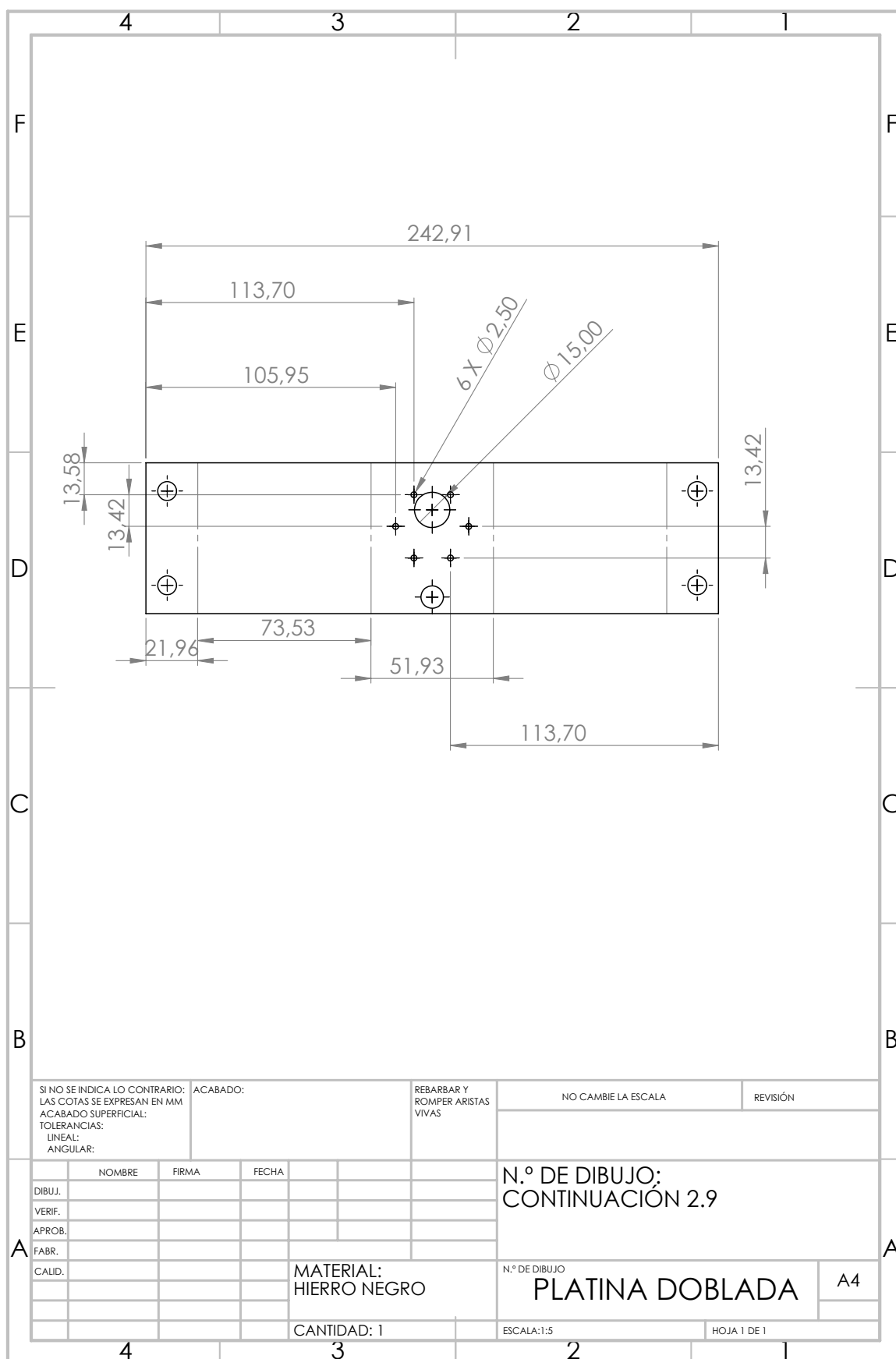


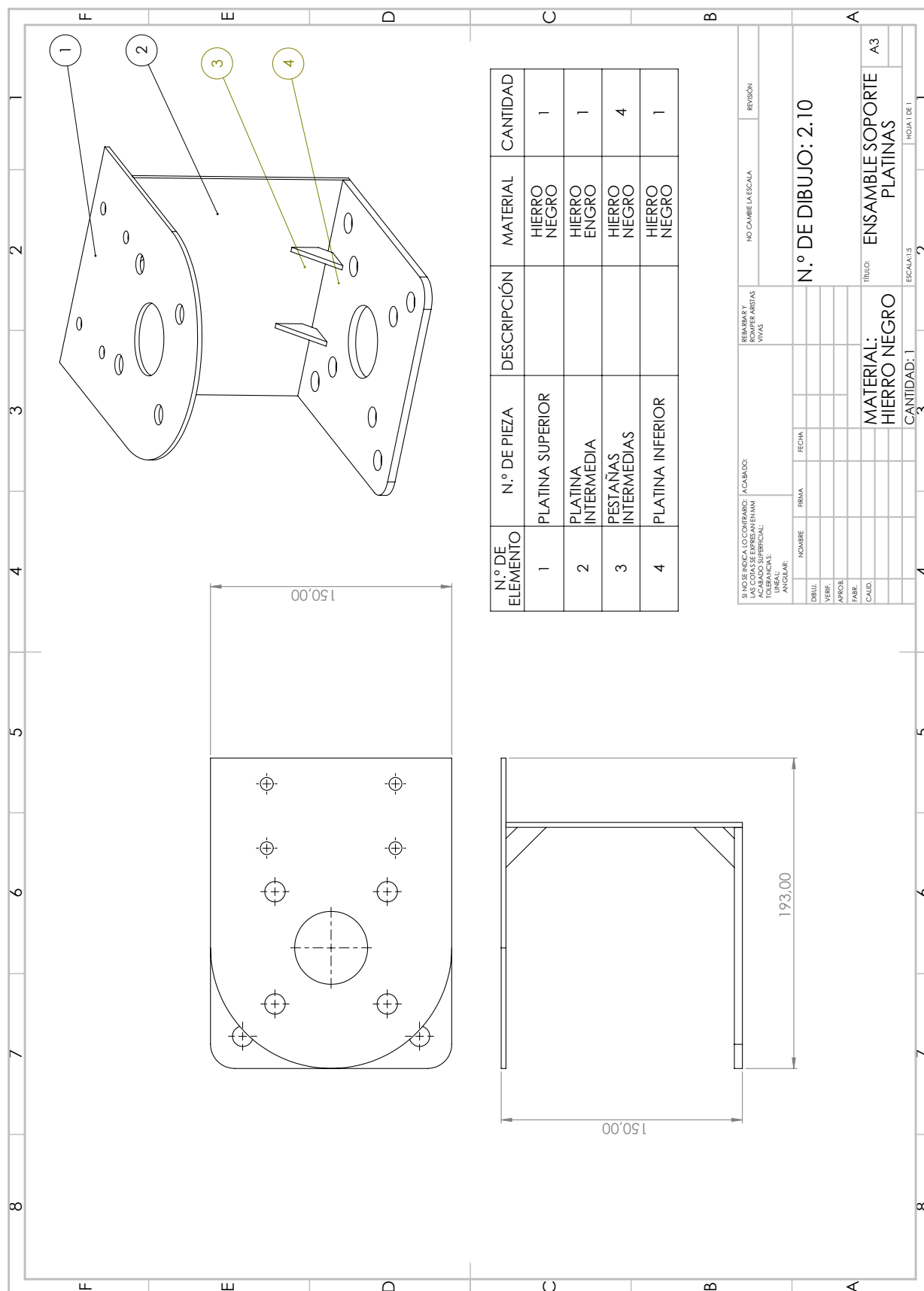


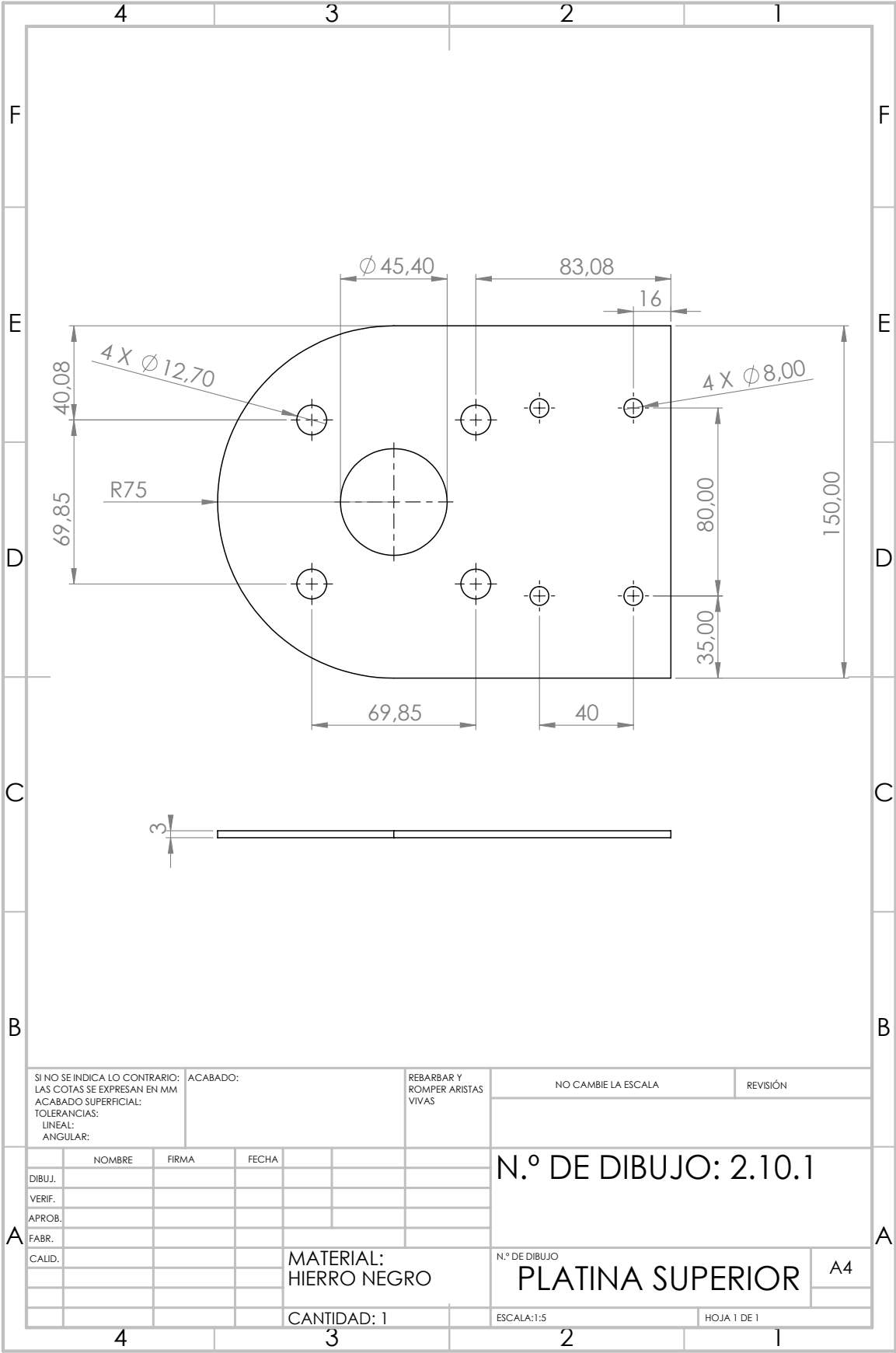




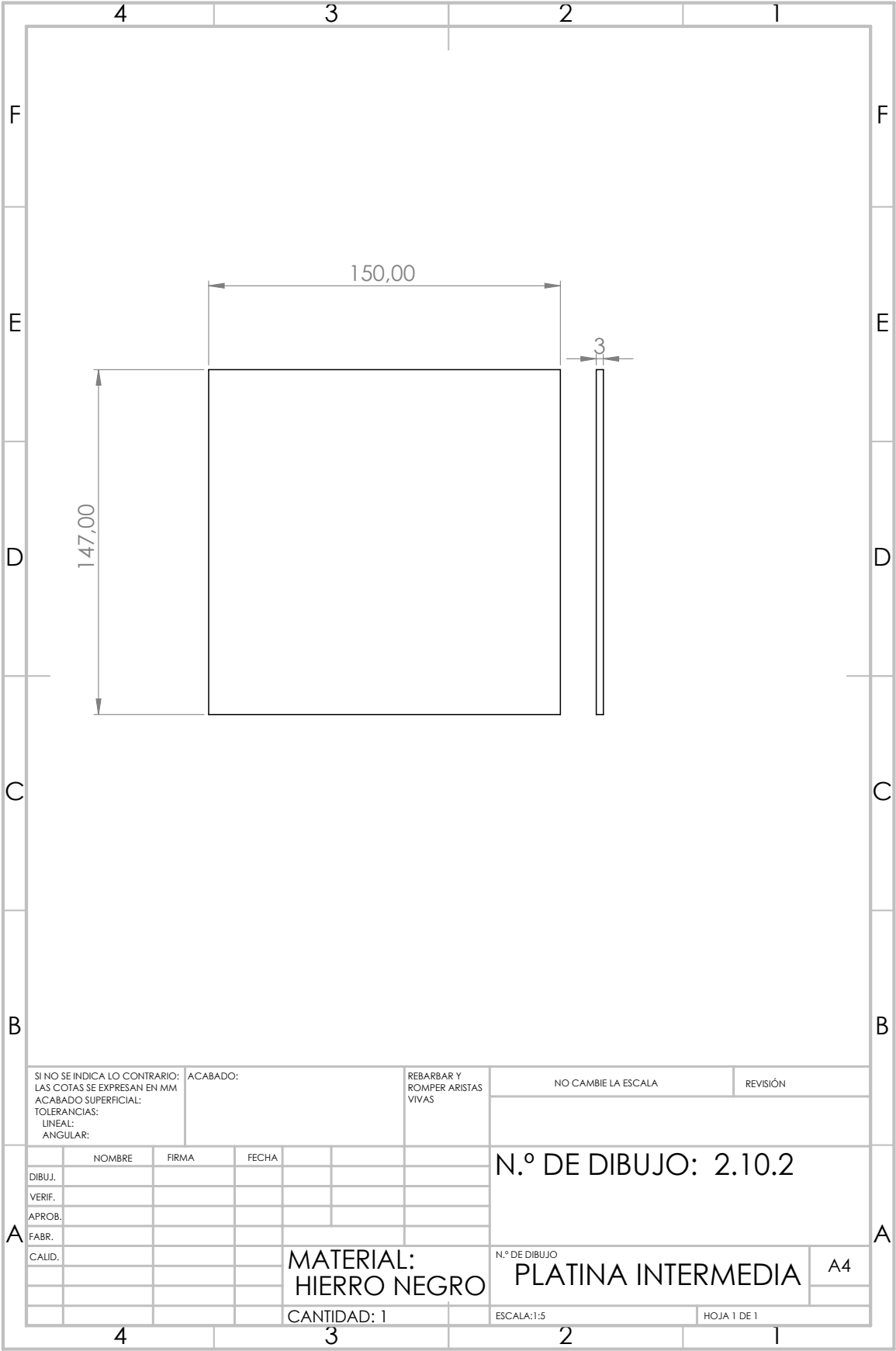


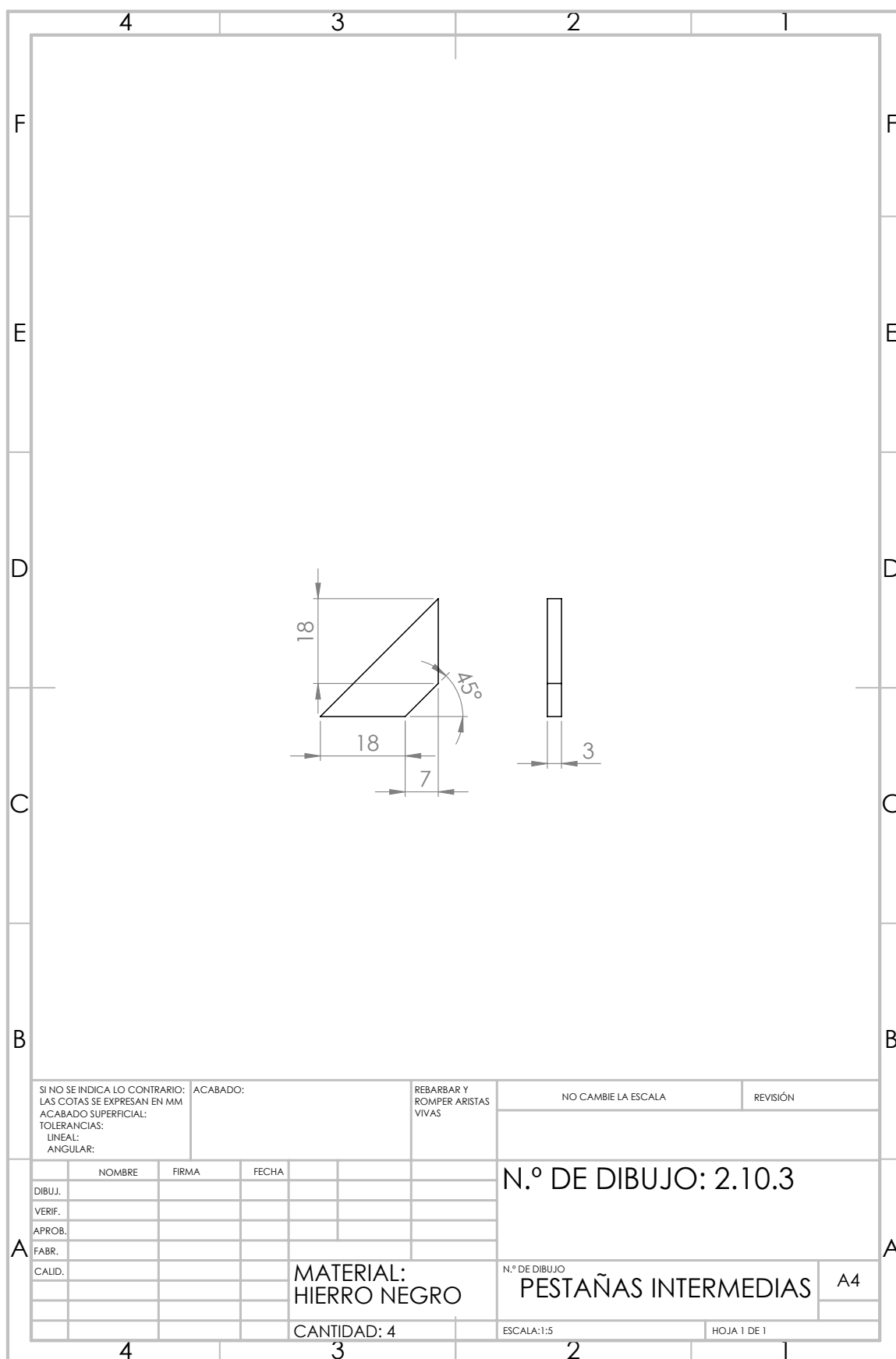


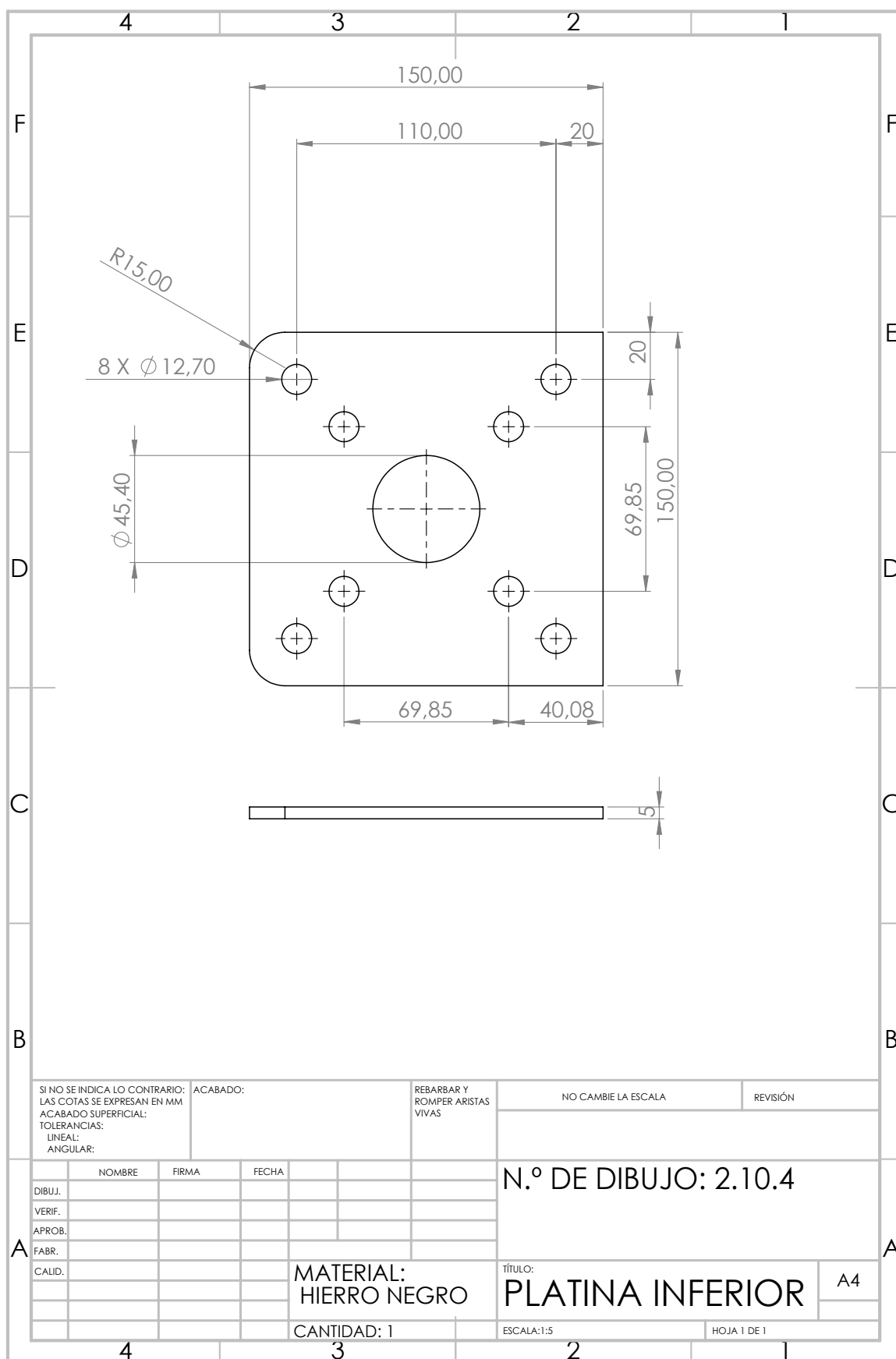


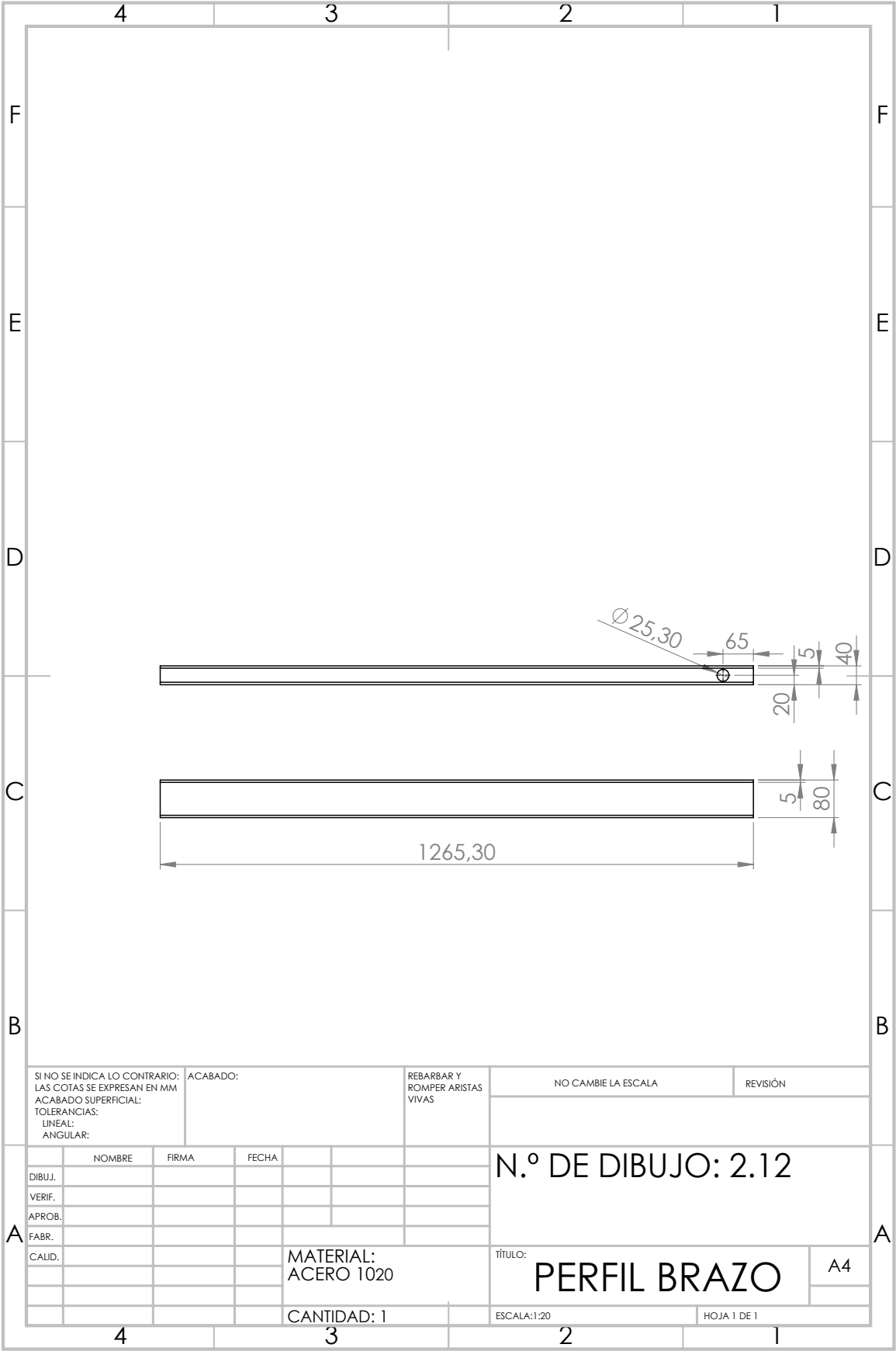


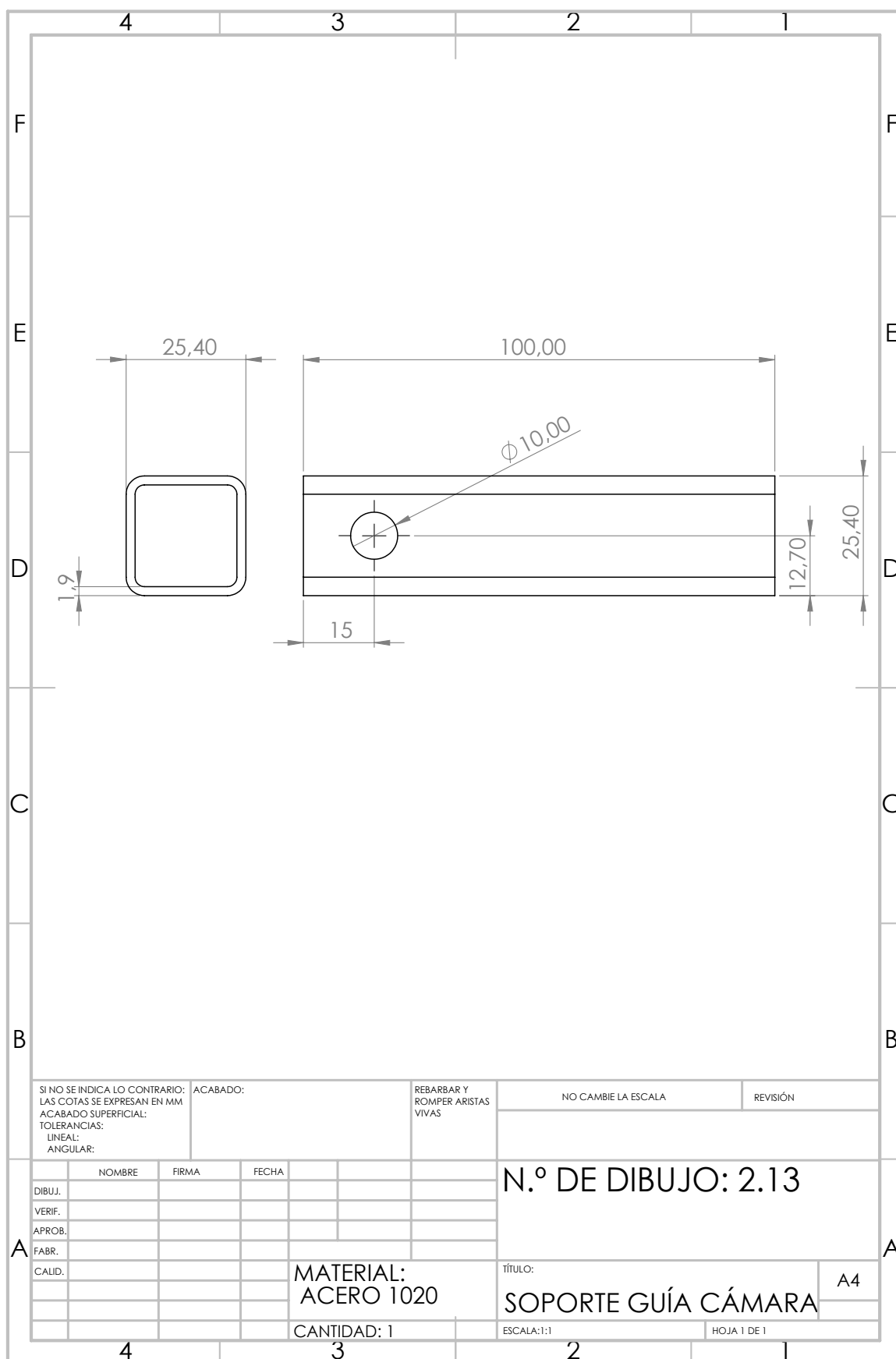
SI NO SE INDICA LO CONTRARIO: LAS COTAS SE EXPRESAN EN MM ACABADO SUPERFICIAL: TOLERANCIAS: LINEAL: ANGULAR:				ACABADO:		REBARBAR Y ROMPER ARISTAS VIVAS		NO CAMBIE LA ESCALA		REVISIÓN	
		NOMBRE		FIRMA		FECHA				N.º DE DIBUJO: 2.10.1	
DIBUJ.											
VERIF.											
APROB.											
FABR.										N.º DE DIBUJO PLATINA SUPERIOR	
CALID.											
										A4	
						MATERIAL: HIERRO NEGRO					
						CANTIDAD: 1				ESCALA:1:5	
										HOJA 1 DE 1	

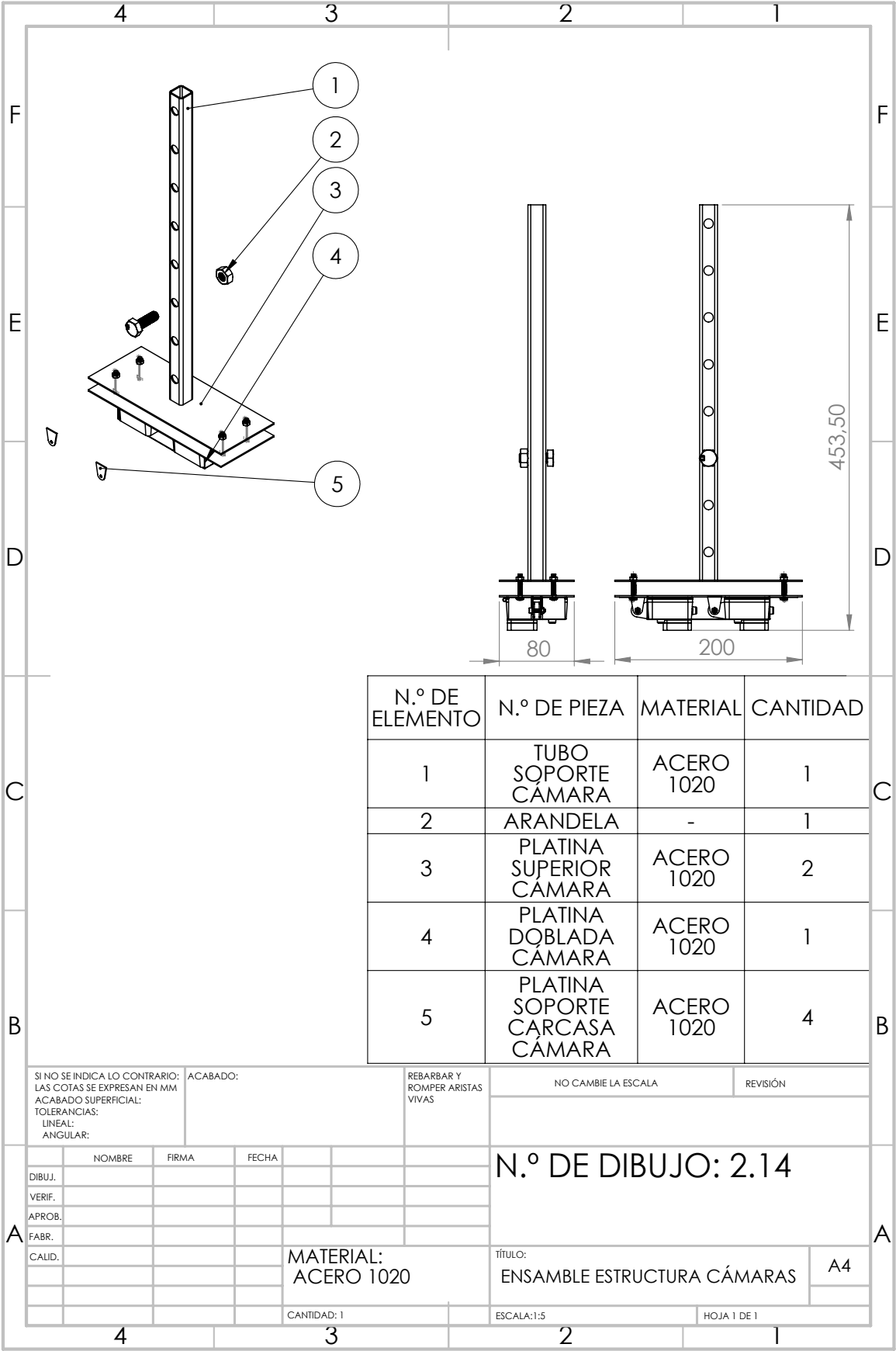


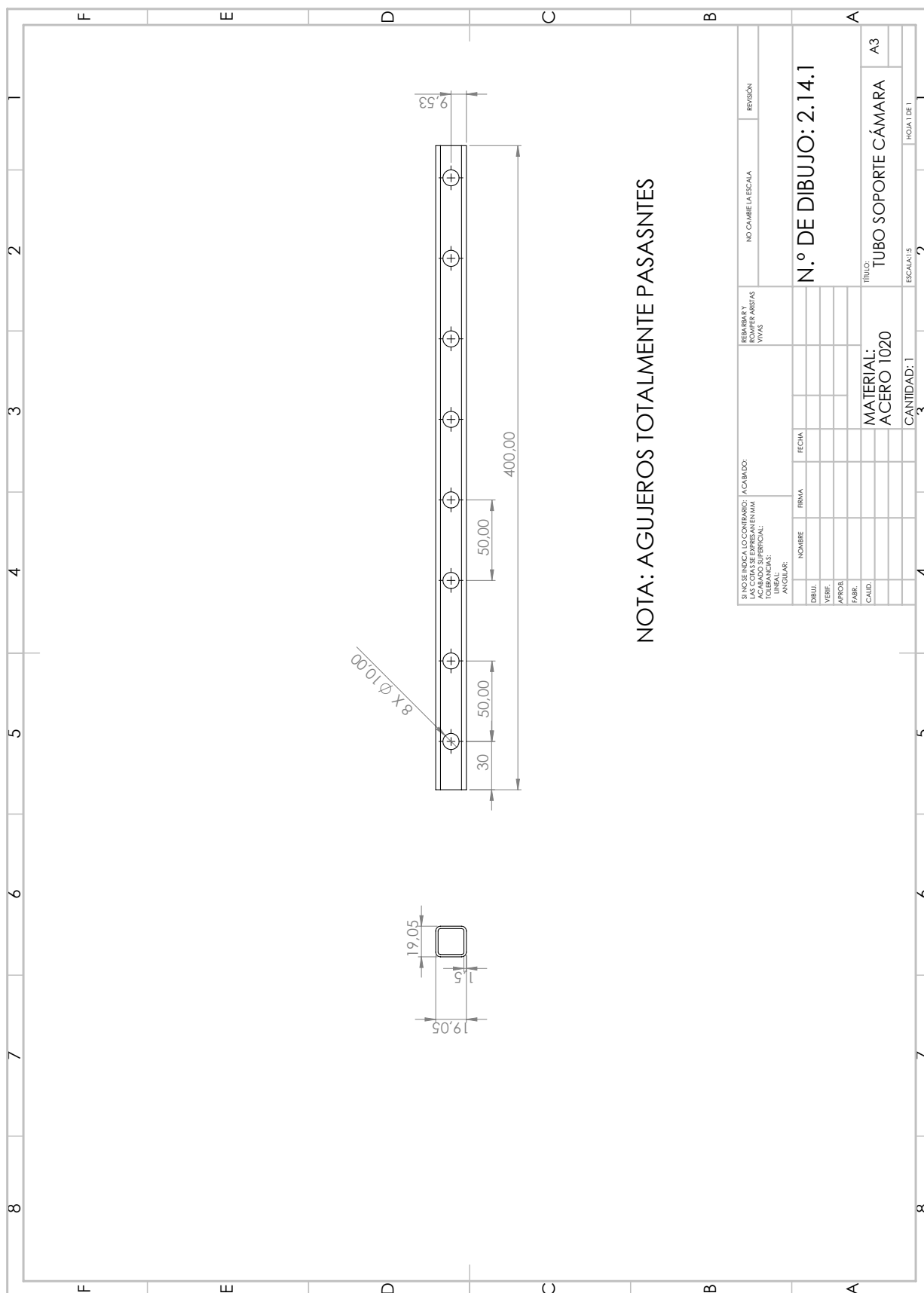


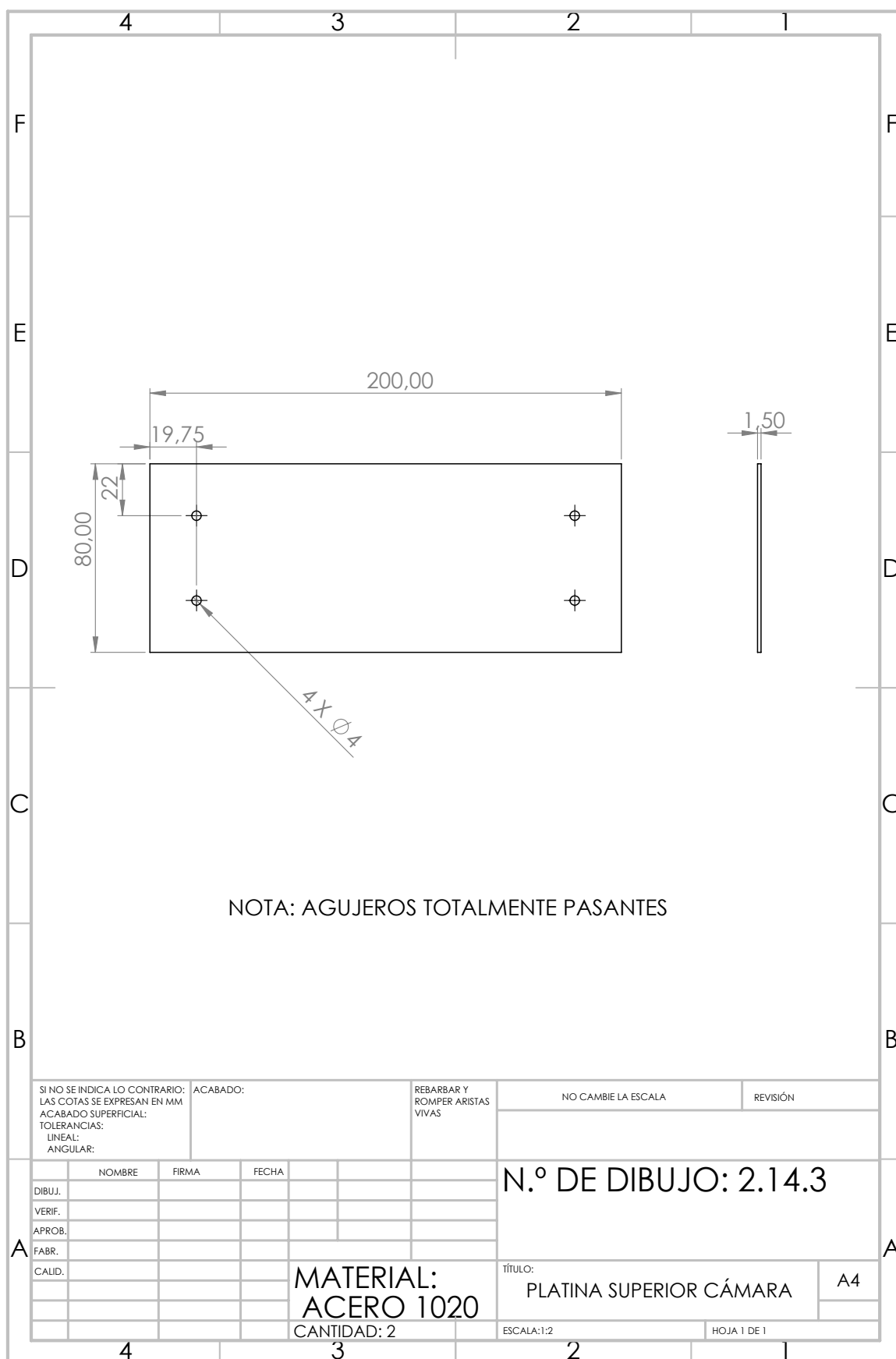


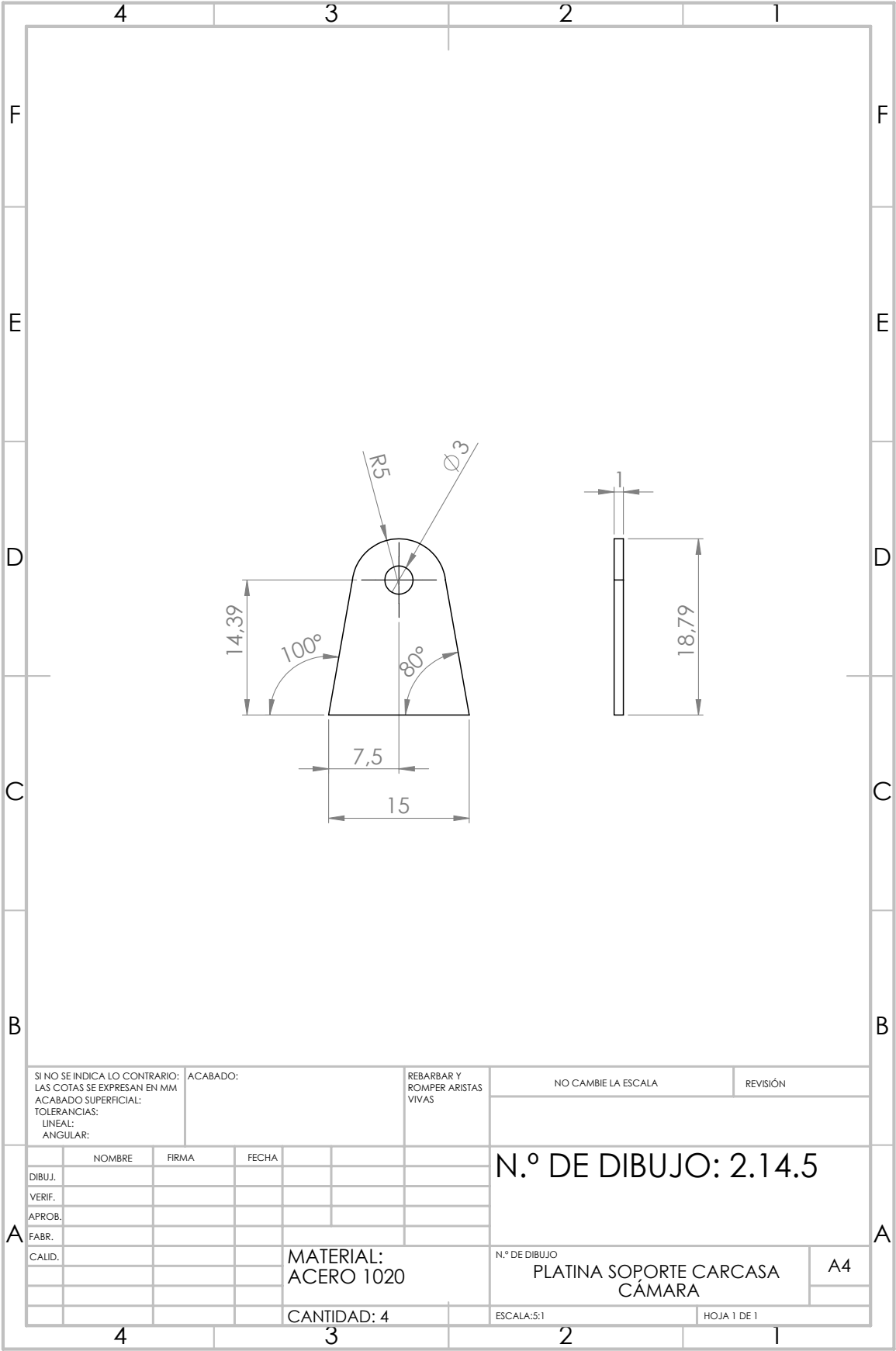




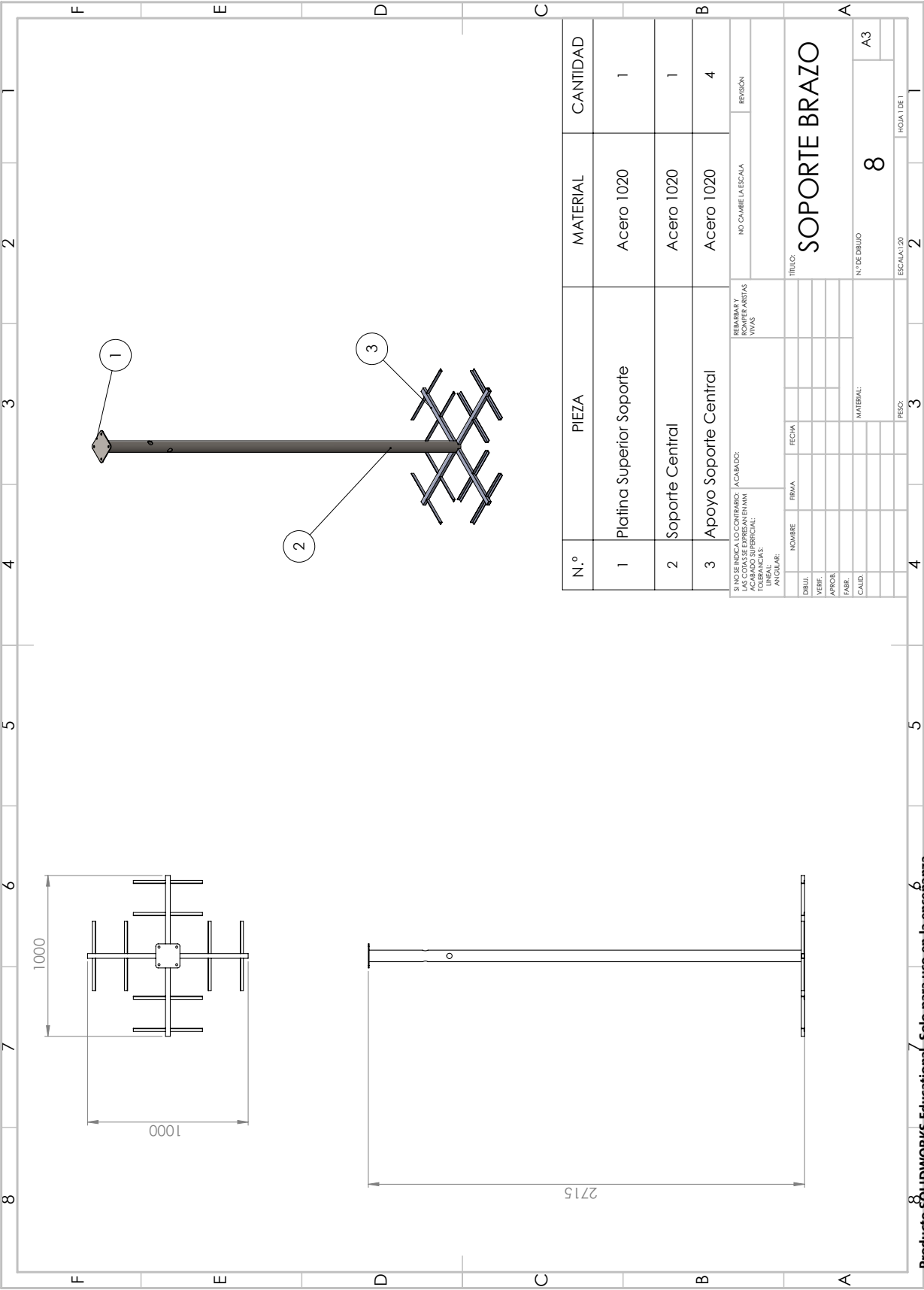


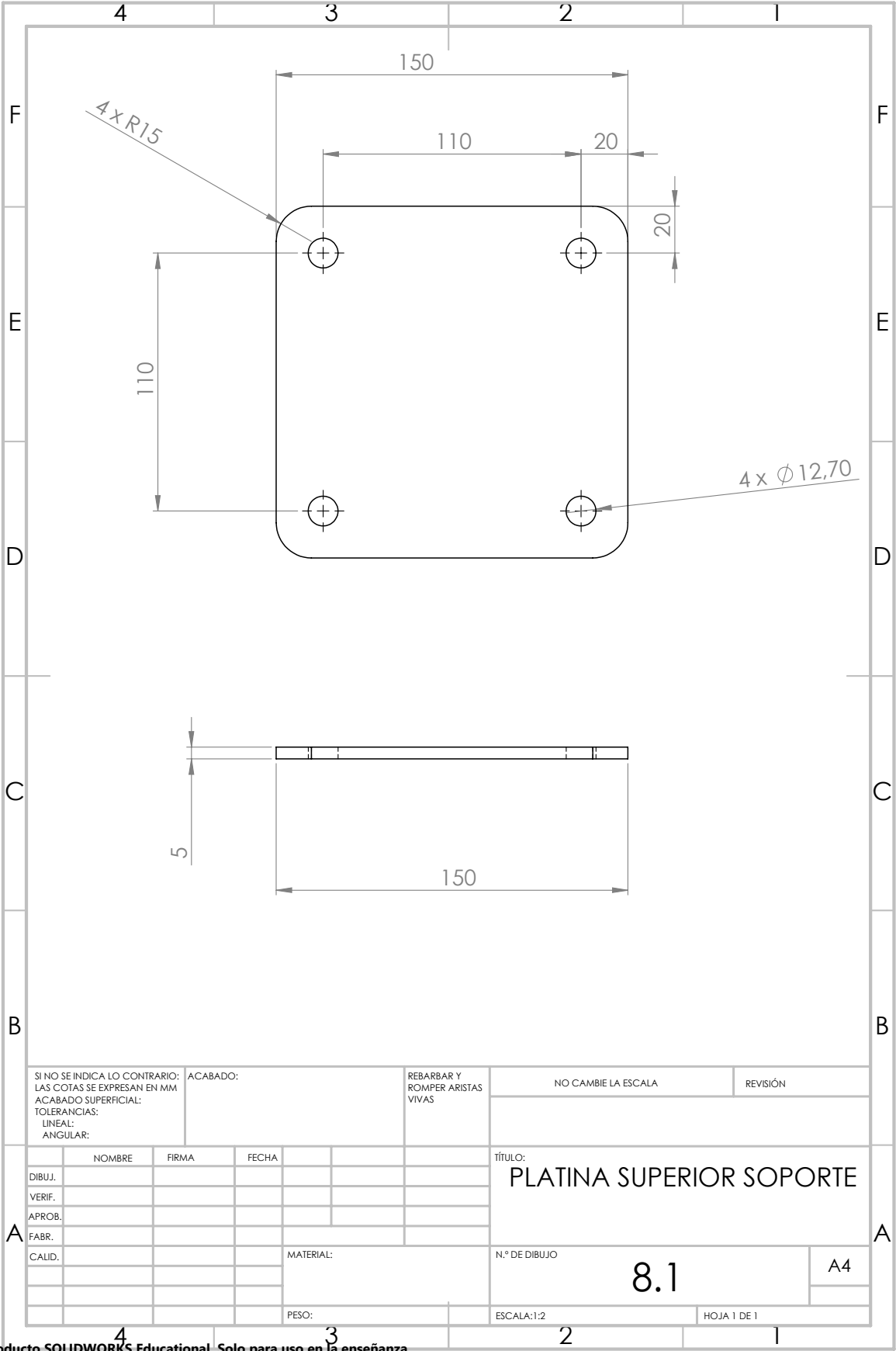


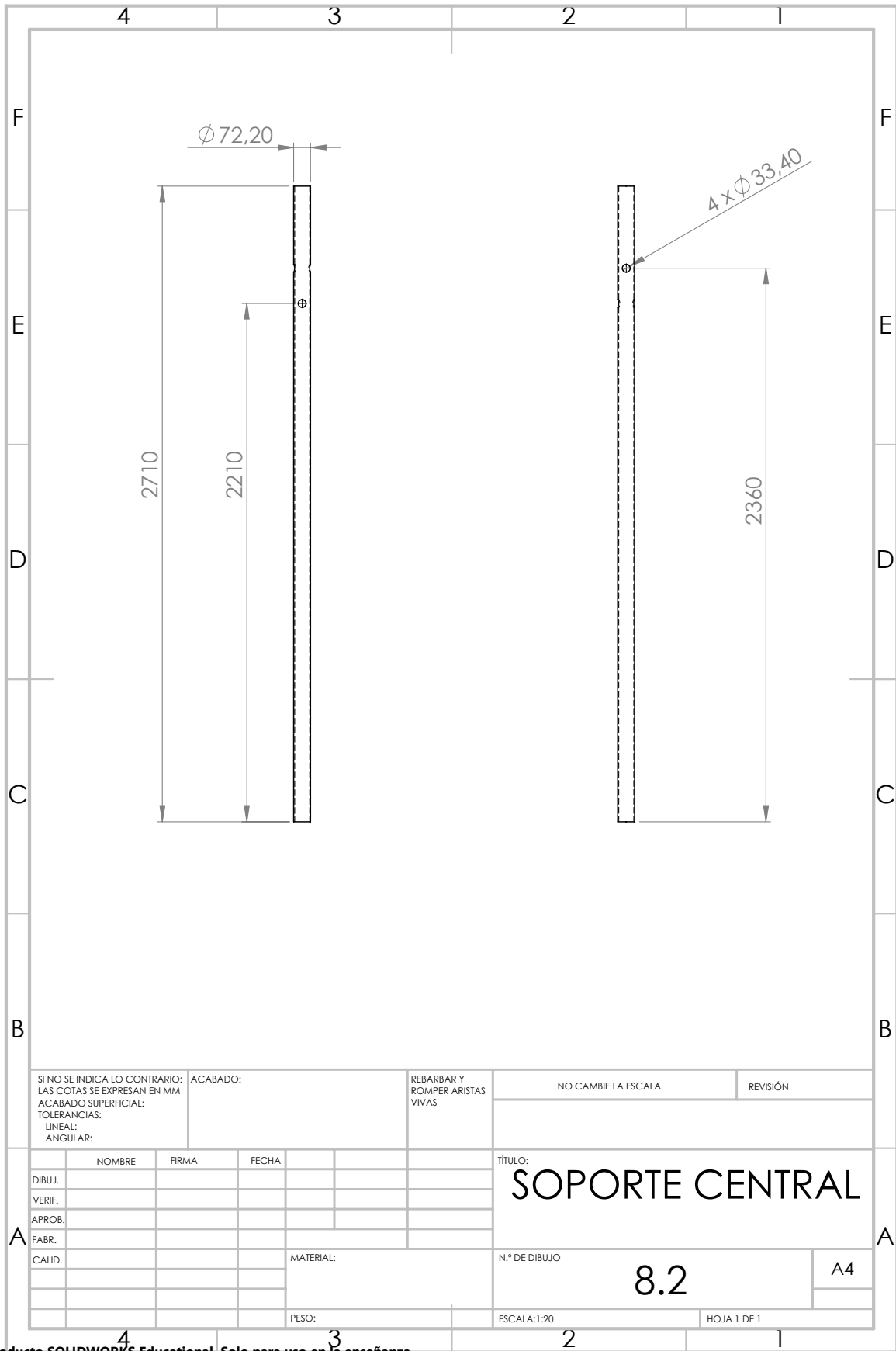




I ARM SUPPORT STRUCTURE







4	3	2	1
F			F
E			E
D			D
C			C
B			B
A			A

N.º	PIEZA	MATERIAL	CANTIDAD
1	Base Central Apoyo	Acero 1020	1
2	Lateral Apoyo	Acero 1020	4

SI NO SE INDICA LO CONTRARIO: LAS COTAS SE EXPRESAN EN MM. ACABADO SUPERFICIAL: TOLERANCIAS: LINEAL: ANGULAR:	ACABADO:	REBARBAR Y ROMPER ARISTAS VIVAS	NO CAMBIE LA ESCALA	REVISIÓN

	NOMBRE	FIRMA	FECHA			TÍTULO: APOYO SOPORTE CENTRAL
DIBUJ.						
VERIF.						
APROB.						
FABR.						
CALID.						N.º DE DIBUJO 8.3
						A4

4	3	2	1

Producto SOLIDWORKS Educational. Solo para uso en la enseñanza.

4	3	2	1
F	E	D	C
B	A		

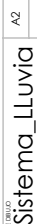
Technical drawing showing a lateral support (LATERAL APOYO) with dimensions: 200 mm height and 20 mm thickness.

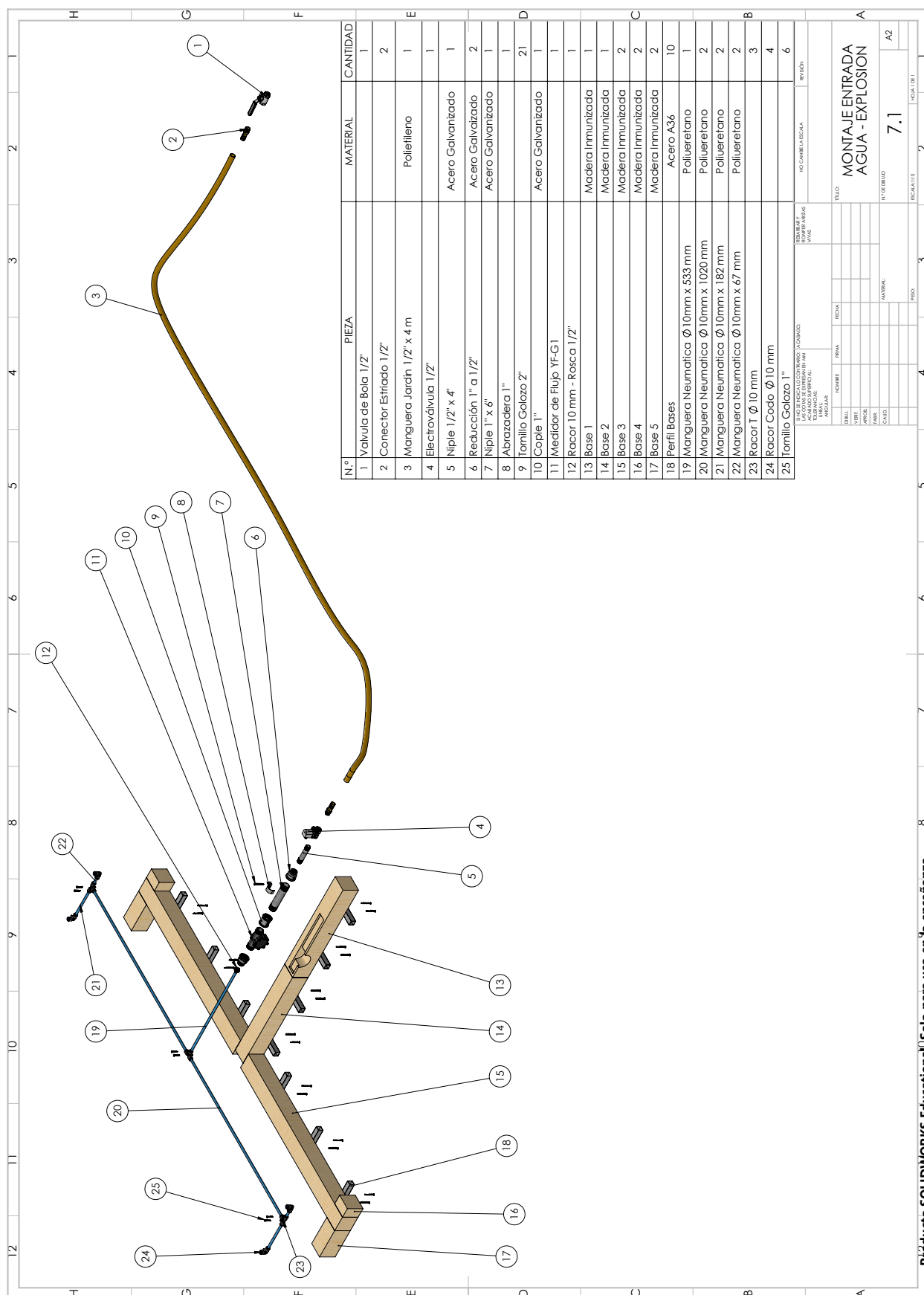
SI NO SE INDICA LO CONTRARIO: LAS COTAS SE EXPRESAN EN MM ACABADO SUPERFICIAL: TOLERANCIAS: LINEAL: ANGULAR:						ACABADO:		REBARBAR Y ROMPER ARISTAS VIVAS		NO CAMBIE LA ESCALA		REVISIÓN
NOMBRE	FIRMA	FECHA				LATERAL APOYO						
DIBUJ.												
VERIF.												
APROB.												
FABR.												
CALID.				MATERIAL:		8.3.2						
				PESO:		ESCALA: 1:2		HOJA 1 DE 1				

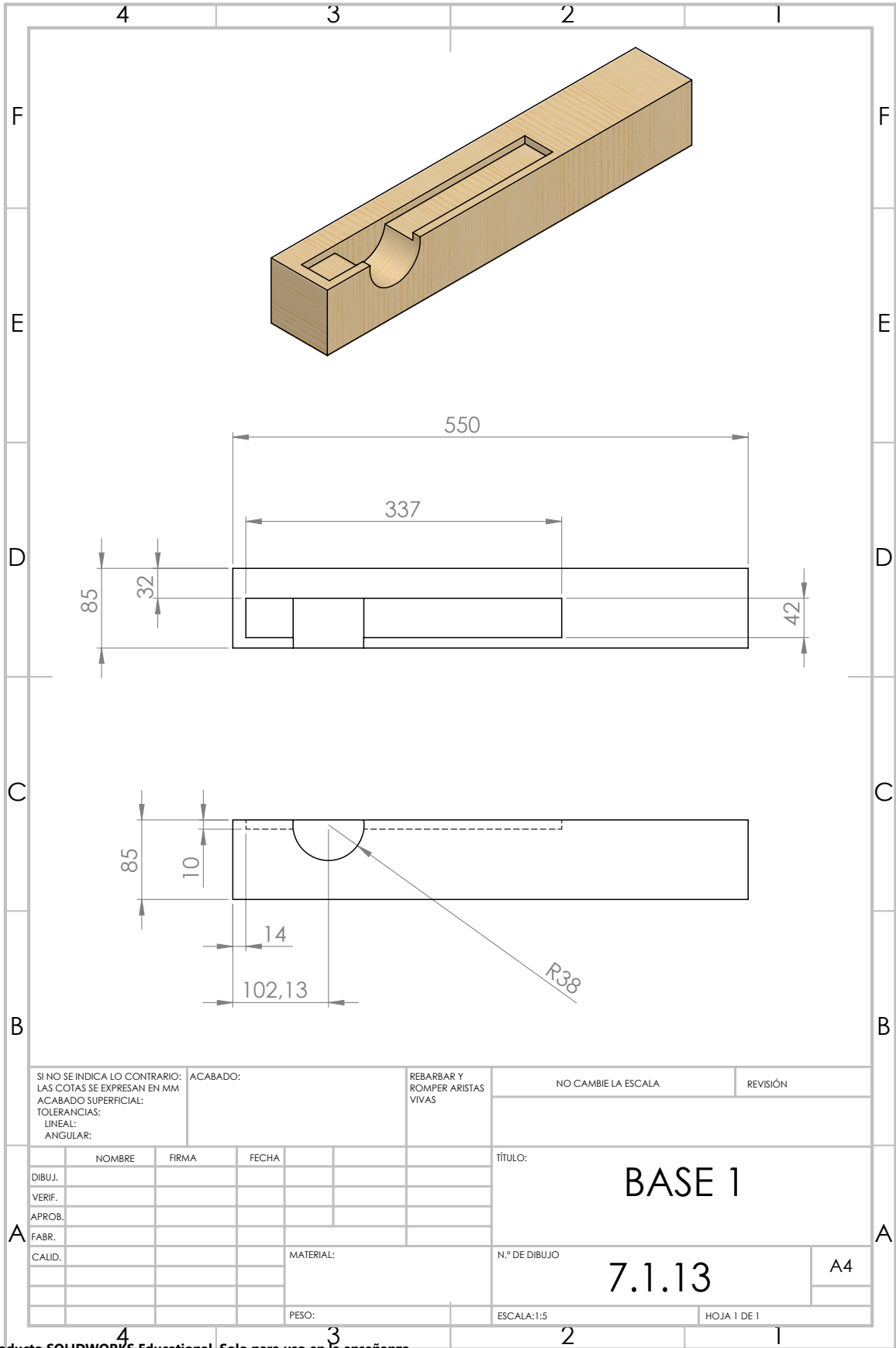
4 3 2 1

Producto SOLIDWORKS Educational. Solo para uso en la enseñanza.

J RAIN SYSTEM





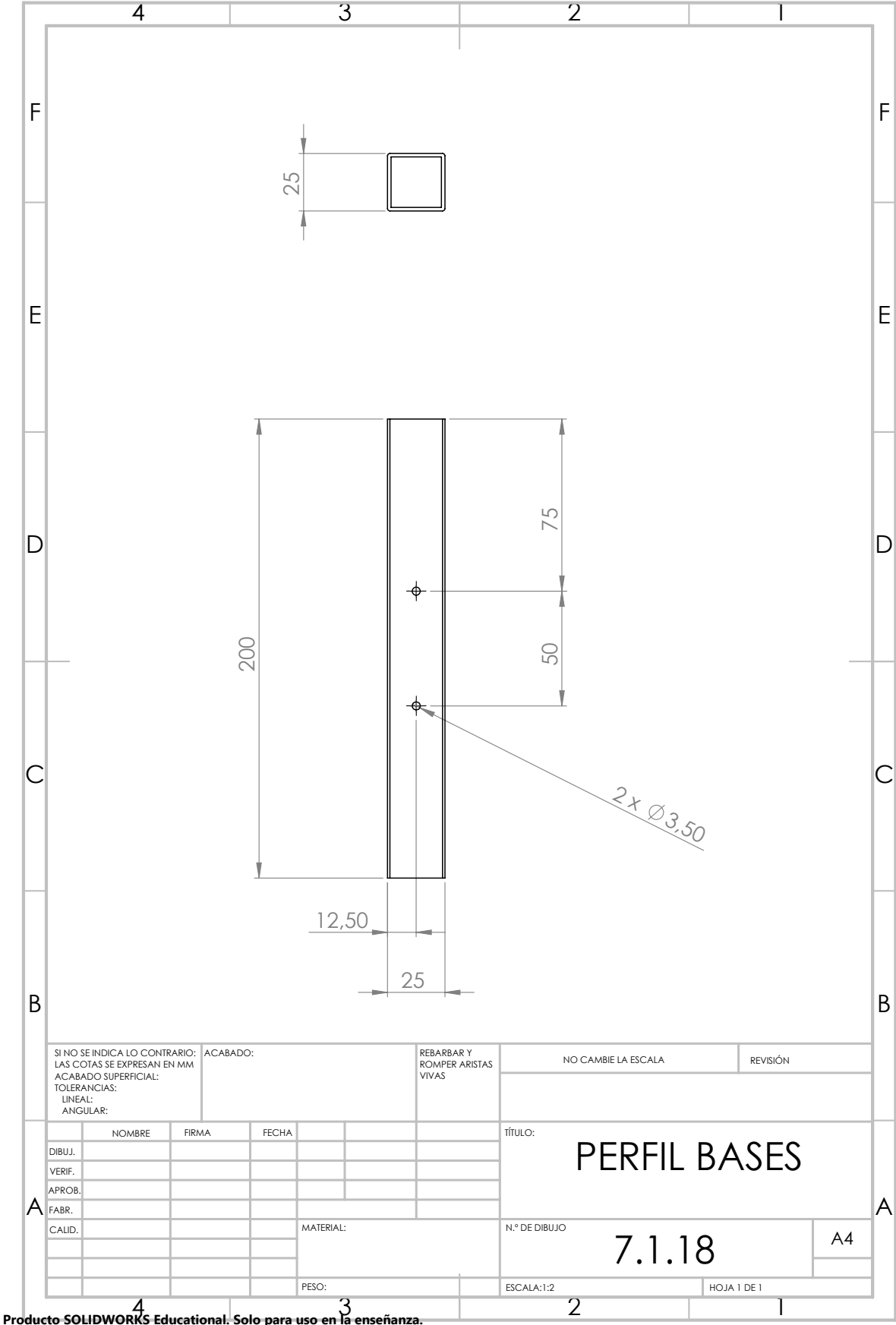


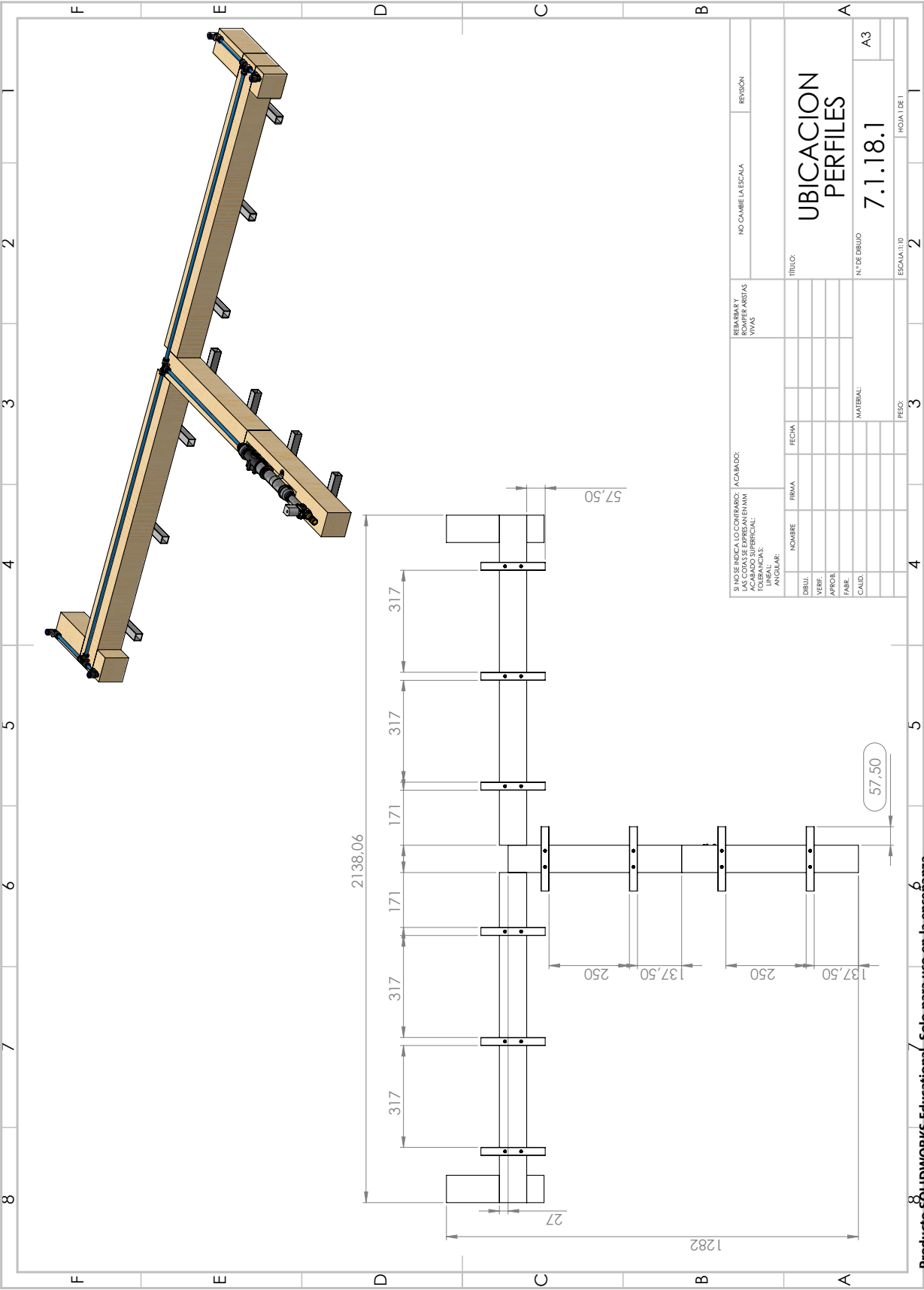
4	3	2	1
F			F
E			E
D			D
C			C
B			B
A			A

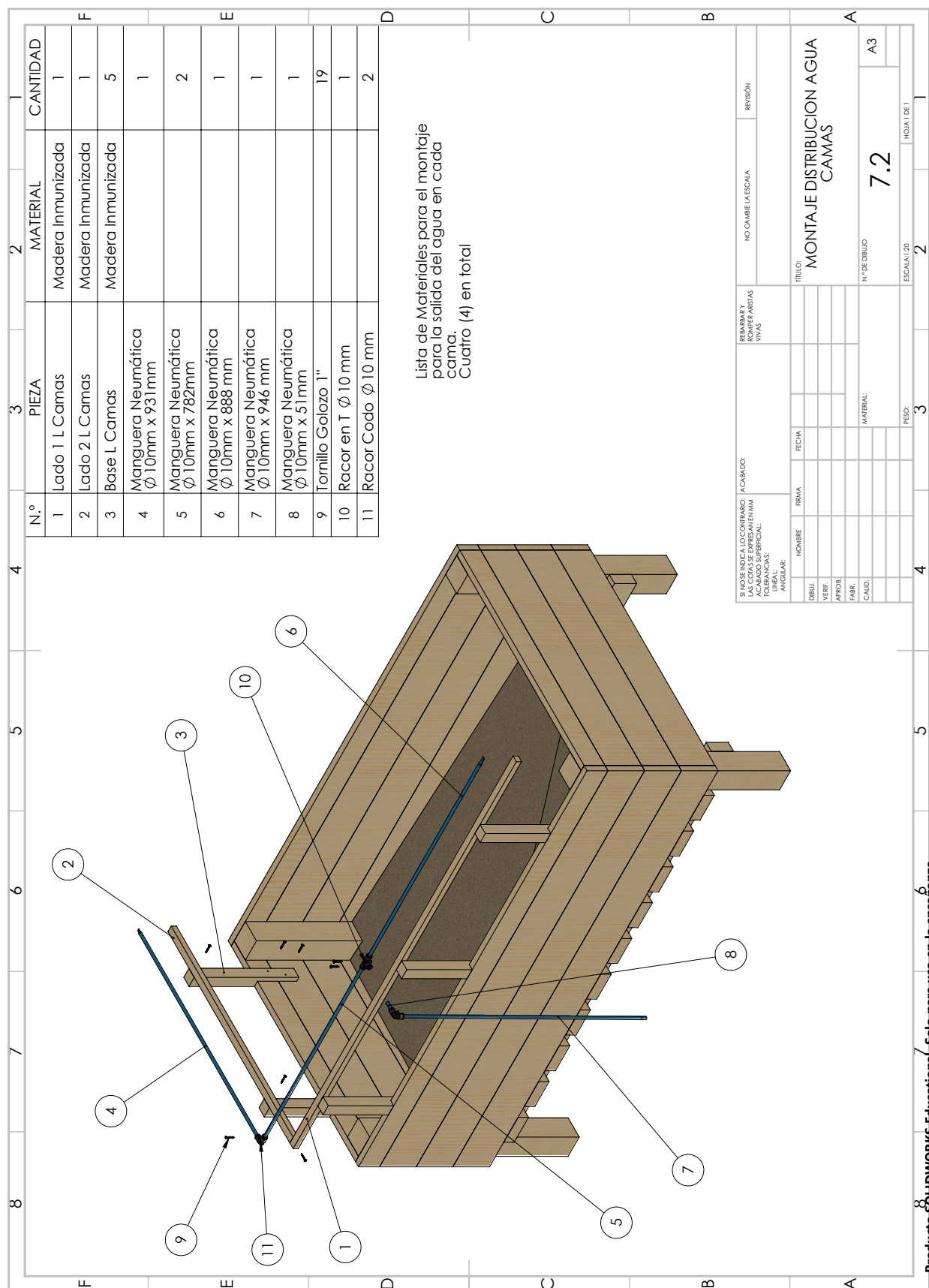
Technical drawing of a square base. The drawing shows a square with a side length of 85 mm. Dimension lines indicate the width and height, both labeled as 85.

PIEZA	LONGITUD (mm)
Base 2	540
Base 3	1027
Base 4	165
Base 5	54

SI NO SE INDICA LO CONTRARIO: LAS COTAS SE EXPRESAN EN MM ACABADO SUPERFICIAL: TOLERANCIAS: LINEAL: ANGULAR:		ACABADO:		REBARBAR Y ROMPER ARISTAS VIVAS		NO CAMBIE LA ESCALA		REVISIÓN
						TÍTULO: BASES SISTEMA LLUVIA		
						N.º DE DIBUJO 7.1.14-17		A4
						ESCALA: 1:1		HOJA 1 DE 1

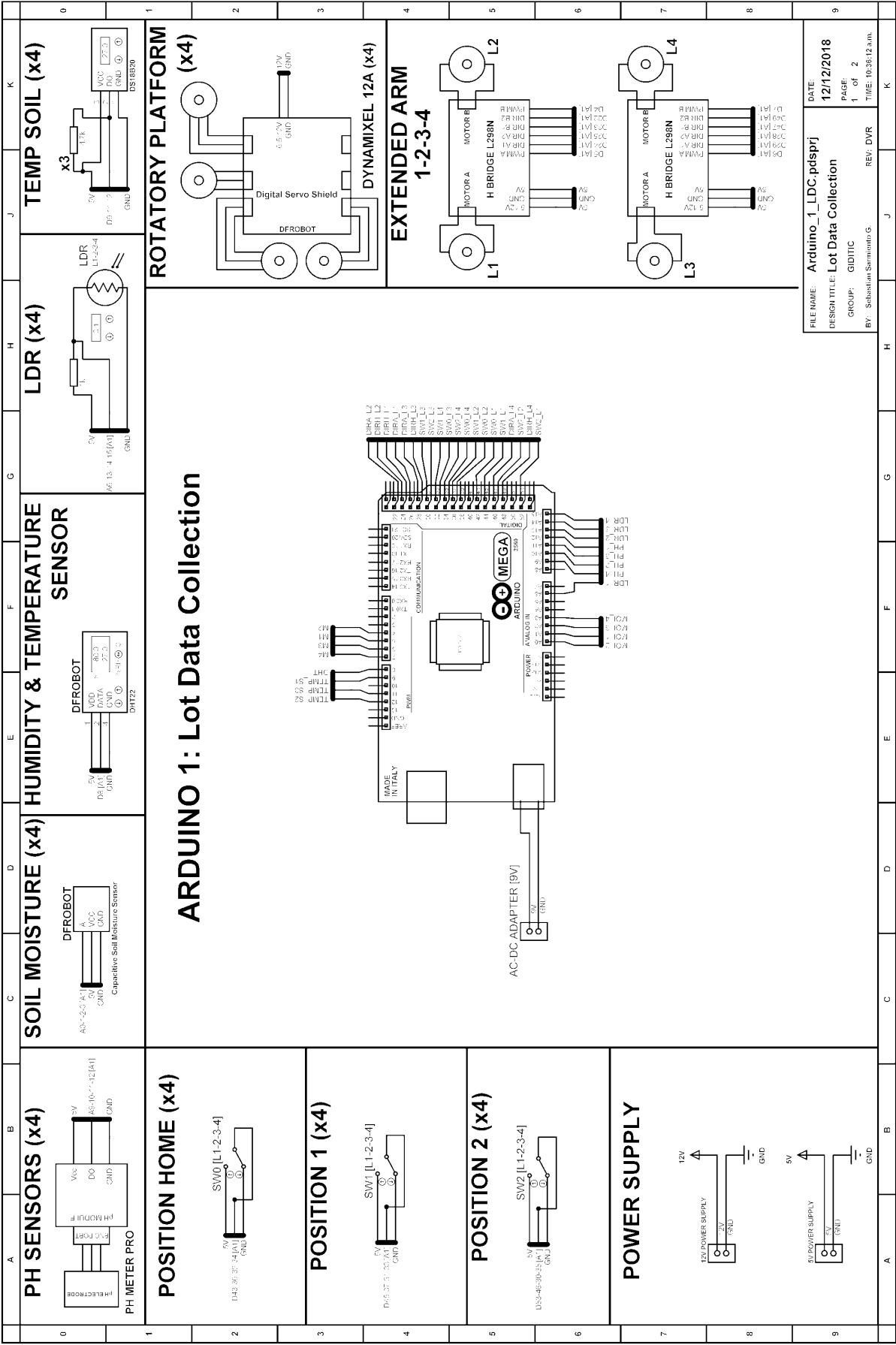




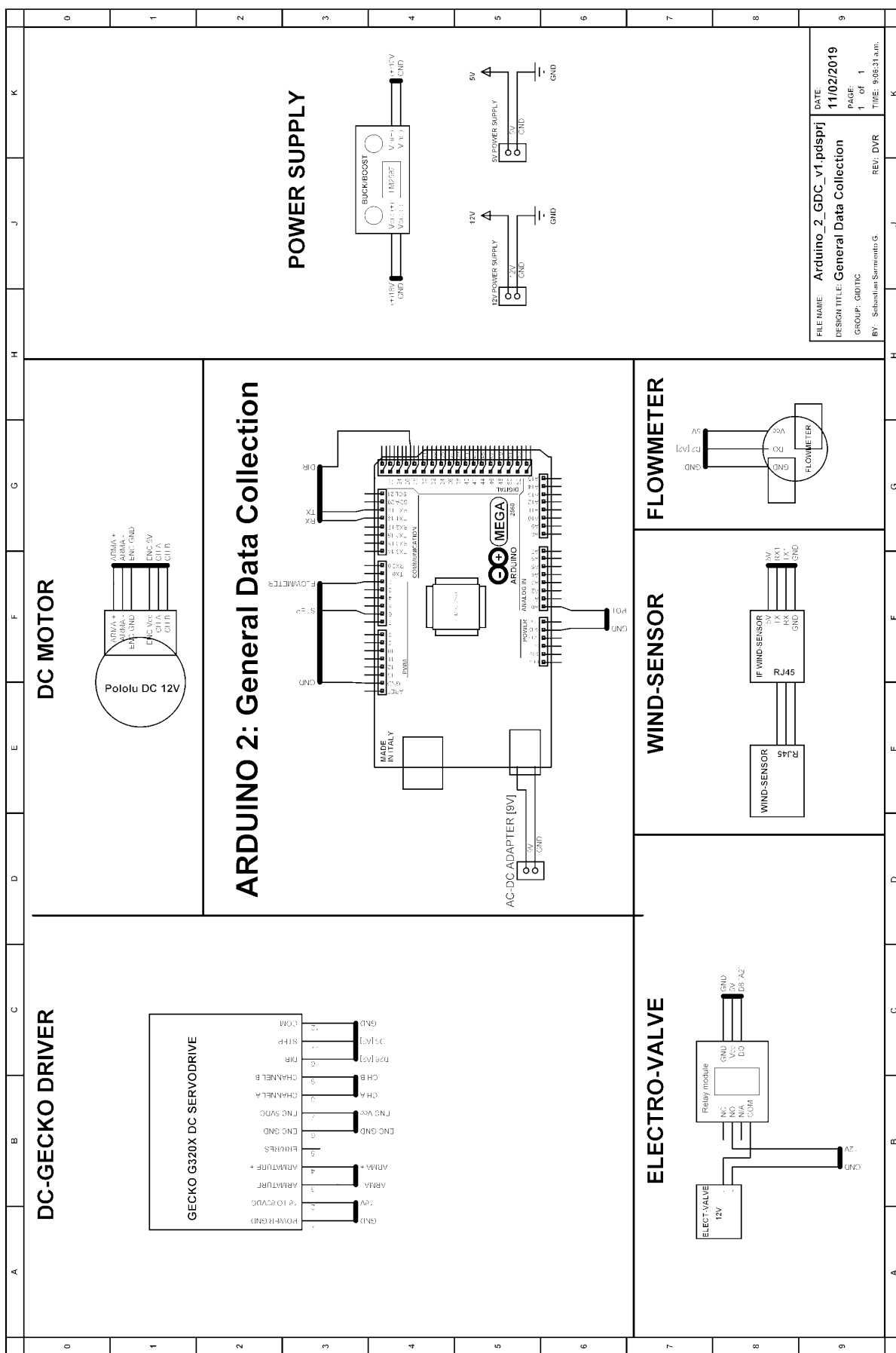


[illegible]

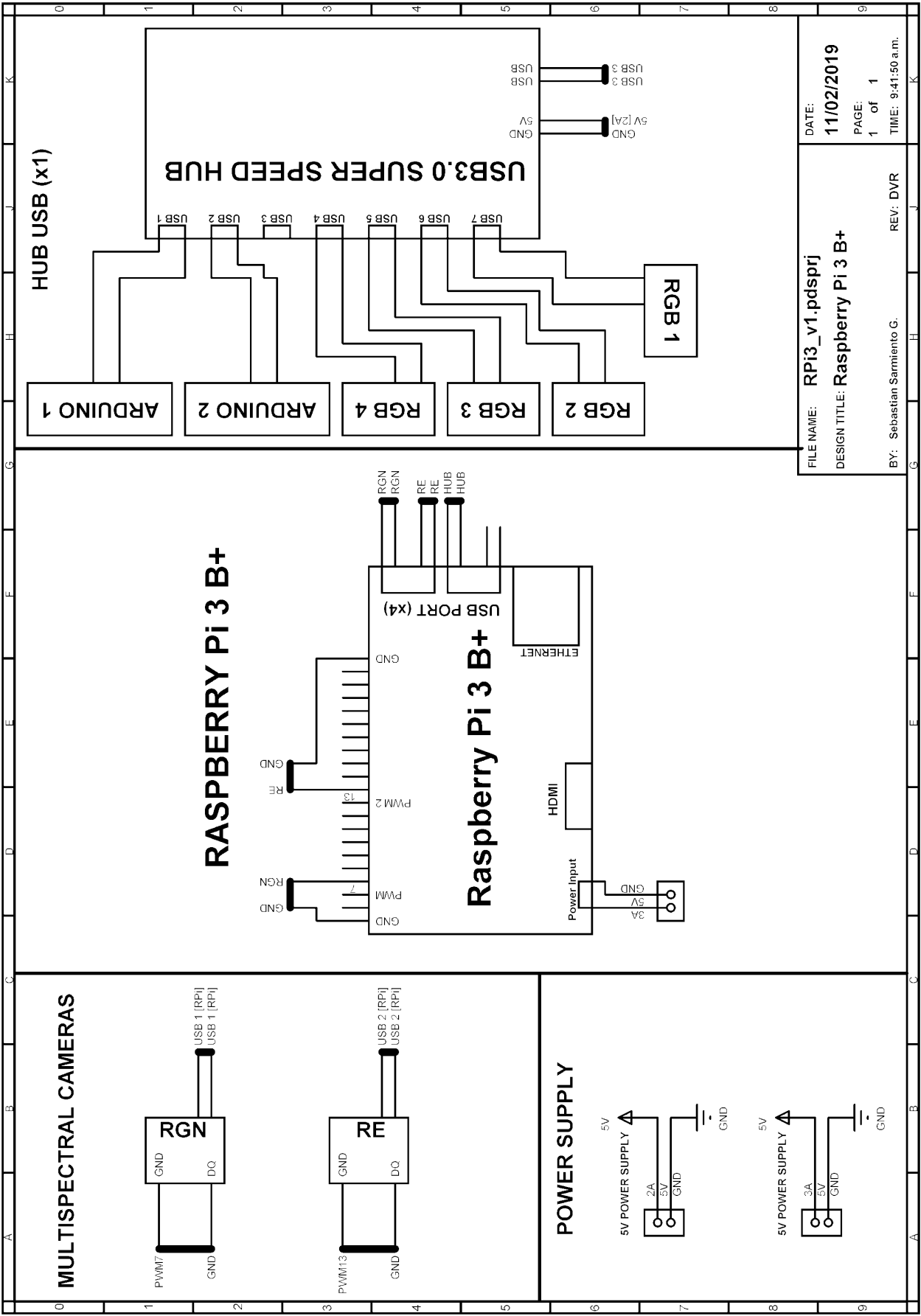
K ARDUINO 1: LOT DATA COLLECTION



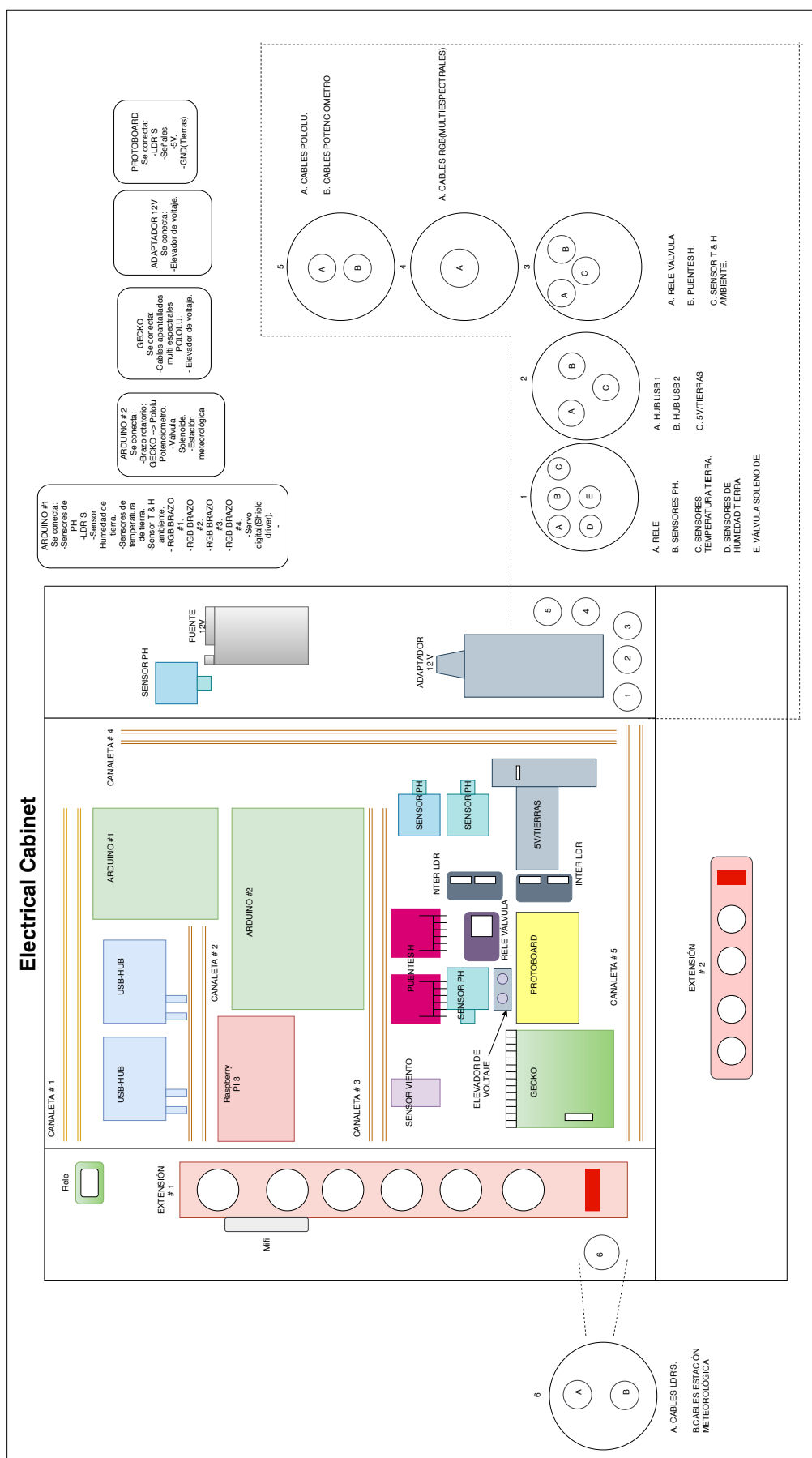
L ARDUINO 2: GENERAL DATA COLLECTION



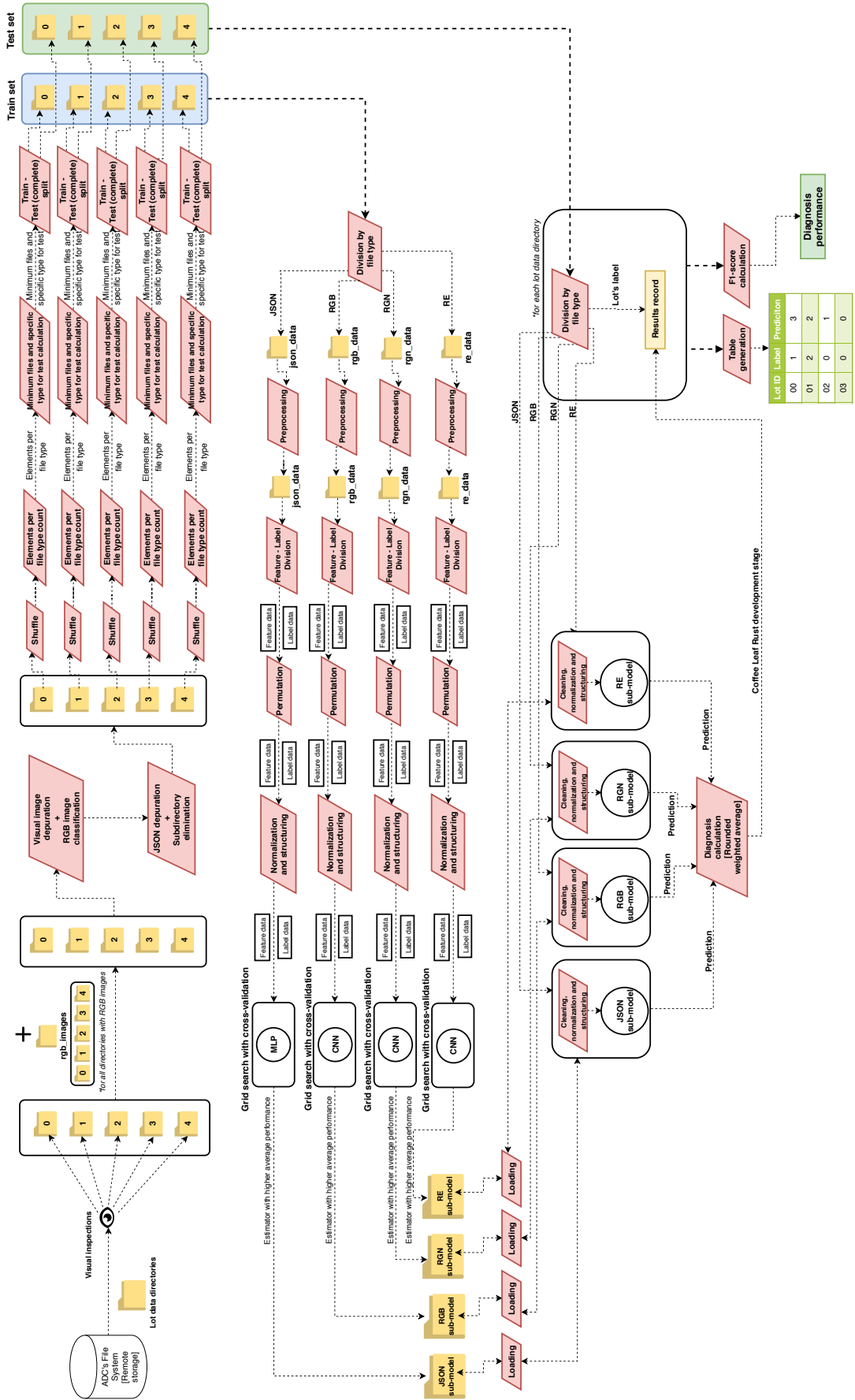
M RASPBERRY PI®



N ELECTRICAL CABINET



**O COLLECTED DATA MANAGEMENT,
DIAGNOSTIC MODEL GENERATION AND
PERFORMANCE EVALUATION**



**P ARDUINOS CHECKLIST TEMPLATE FOR
LABORATORY**

**Q ARDUINOS CHECKLIST FOR
LABORATORY**

Checklist in Laboratory for Arduino's

Sensor/Actuator	Name	Arduino	Voltage	Lot				Signal	Calibration?	Circuit Board?
				1	2	3	4			
Sensor	Illuminance	1	5.1	X				Analog	YES	YES
			5.1		X					
			5.1			X				
			5.1				X			
Sensor	pH	1	4.9	X				Analog	YES	YES
			4.9		X					
			4.9			X				
			4.9				X			
Sensor	Temp (soil)	1	5	X				Digital	NO	YES
			5		X					
			5			X				
			5				X			
Sensor	Moisture (soil)	1	4.89	X				Analog	YES	YES
			4.89		X					
			4.89			X				
			4.89				X			
Sensor	Temp (env)	1	5	X				Digital	NO	NO
			5							
			5							
			5							
Sensor	Humidity (env)	1	5	X				Digital	NO	NO
			5							
			5							
			5							
Actuator	Dynamixel	1	12	X				Serial	NO	YES
			12		X					
			12			X				
			12				X			

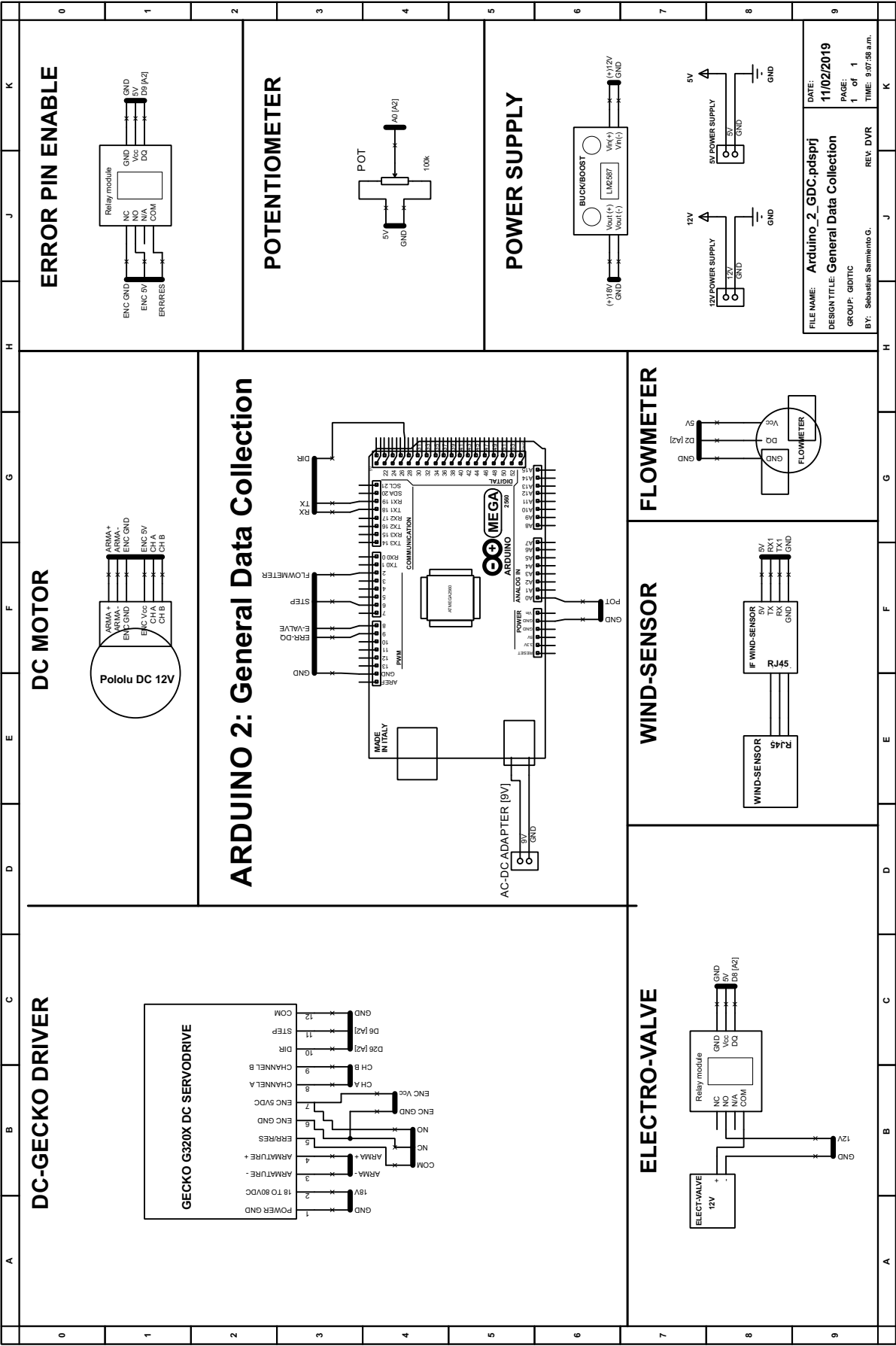
Actuator	Mini-DC	1	5	X						PWM/Digital	NO	YES
			5		X							
			5				X					
			5						X			
Actuator	Flow-rate	2	5.1	X					Digital	YES	NO	
Actuator	Pololu-DC motor	2	12	X					PWM/Digital	YES	YES	
Actuator	Wind sensor	2	4.9	X					Serial	NO	YES	
Actuator	Solenoid-valve	2	12	X					PWM/Digital	NO	YES	

R ARDUINOS CHECKLIST TEMPLATE FOR GREENHOUSE

CHECKLIST Arduino's Greenhouse

[illegible]

**S ARDUINO 2: GENERAL DATA
COLLECTION - FINAL VERSION**



**T RASPBERRY PI® CHECKLIST TEMPLATE
FOR GREENHOUSE**

CHECKLIST RASPBERRY COMPONENTS (GREENHOUSE)

[illegible]

U RASPBERRY PI® - FINAL VERSION

V SUB-MODELS' ARCHITECTURES

JSON sub-model's architecture

N	Layer	Output Shape	# of Parameters
1.	Input layer	(None, 6)	-
2.	Fully Connected	(None, 16)	112
3.	Batch Normalization	(None, 16)	64
4.	Activation	(None, 16)	0
5.	Fully Connected	(None, 64)	1088
6.	Batch Normalization	(None, 64)	256
7.	Activation	(None, 64)	0
8.	Dropout	(None, 64)	0
9.	Fully Connected	(None, 32)	2080
10.	Batch Normalization	(None, 32)	128
11.	Activation	(None, 32)	0
12.	Dropout(Rate=rate/2)	(None, 32)	0
13.	Fully Connected	(None, 4)	132
14.	Activation	(None, 4)	0

RGB sub-model's architecture

N	Layer	Output Shape	# of Parameters
1.	Input layer	(None, 96, 128, 3)	-
2.	Convolutional2D(Kernel=(5, 5))	(None, 92, 124, 18)	1.368
3.	Batch Normalization	(None, 92, 124, 18)	72
4.	Activation	(None, 92, 124, 18)	0
5.	Max Pooling(Pool=(2, 2))	(None, 46, 62, 18)	0
6.	Convolutional2D(Kernel=(5, 5))	(None, 42, 58, 36)	16.236
7.	Batch Normalization	(None, 42, 58, 36)	144
8.	Activation	(None, 42, 58, 36)	0
9.	Max Pooling(Pool=(2, 2))	(None, 21, 29, 36)	0
10.	Convolutional2D(Kernel=(3, 3))	(None, 19, 27, 54)	17.550
11.	Batch Normalization	(None, 19, 27, 54)	216
12.	Activation	(None, 19, 27, 54)	0
13.	Max Pooling(Pool=(2, 2))	(None, 9, 13, 54)	0
14.	Dropout	(None, 9, 13, 54)	0
15.	Flatten	(None, 6318)	0
16.	Fully Connected	(None, 512)	3'235.328
17.	Batch Normalization	(None, 512)	2048
18.	Activation	(None, 512)	0
19.	Dropout	(None, 512)	0
20.	Fully Connected	(None, 128)	65.664
21.	Batch Normalization	(None, 128)	512
22.	Activation	(None, 128)	0
23.	Dropout(Rate=rate/2)	(None, 128)	0
24.	Fully Connected	(None, 5)	645
25.	Activation	(None, 5)	0

RGN and RE sub-models' architectures

N	Layers	Output Shape	# of Parameters
1.	Input layer	(None, 128, 96, 3)	-
2.	Convolutional2D(Kernel=(5, 5))	(None, 124, 92, 18)	1.368
3.	Batch Normalization	(None, 124, 92, 18)	72
4.	Activation	(None, 124, 92, 18)	0
5.	Max Pooling(Pool=(2, 2))	(None, 62, 46, 18)	0
6.	Convolutional2D(Kernel=(5, 5))	(None, 58, 42, 36)	16.236
7.	Batch Normalization	(None, 58, 42, 36)	144
8.	Activation	(None, 58, 42, 36)	0
9.	Max Pooling(Pool=(2, 2))	(None, 29, 21, 36)	0
10.	Convolutional2D(Kernel=(3, 3))	(None, 27, 19, 54)	17.550
11.	Batch Normalization	(None, 27, 19, 54)	216
12.	Activation	(None, 27, 19, 54)	0
13.	Max Pooling(Pool=(2, 2))	(None, 13, 9, 54)	0
14.	Dropout	(None, 13, 9, 54)	0
15.	Flatten	(None, 6318)	0
16.	Fully Connected	(None, 512)	3'235.328
17.	Batch Normalization	(None, 512)	2048
18.	Activation	(None, 512)	0
19.	Dropout	(None, 512)	0
20.	Fully Connected	(None, 128)	65.664
21.	Batch Normalization	(None, 128)	512
22.	Activation	(None, 128)	0
23.	Dropout(Rate=rate/2)	(None, 128)	0
24.	Fully Connected	(None, 4)	516
25.	Activation	(None, 4)	0

**W COMPARATIVE TABLE FOR THE
STATISTICAL ANALYSIS**

ID	Visual inspection	Technological integration	ID	Visual inspection	Technological integration	ID	Visual inspection	Technological integration	ID	Visual inspection	Technological integration	ID	Visual inspection	Technological integration	ID	Visual inspection	Technological integration	ID	Visual inspection	Technological integration	ID	Visual inspection	Technological integration
1277	0	0	192	0	0	209	0	0	384	0	0	633	0	0	229	0	0	257	2	2			
494	4	4	391	3	2	749	0	0	608	0	0	1511	0	0	449	4	4	3148	3	1			
498	0	1	1090	0	0	245	2	2	628	0	0	1517	0	0	105	0	0	268	0	0			
4217	4	3	380	3	3	854	0	2	1572	0	0	1571	0	0	527	0	0	215	2	2			
469	0	1	246	0	0	3115	3	3	181	0	0	3193	3	2	361	0	0	164	0	0			
1258	0	0	356	3	3	183	0	0	665	0	0	259	2	2	635	0	0	694	0	1			
490	0	1	4231	4	3	1567	0	1	1077	0	0	154	0	0	1342	0	0	208	0	1			
636	0	1	321	0	0	1525	0	0	204	0	0	1479	0	0	252	0	1	436	4	4			
230	2	2	425	4	4	1437	0	1	360	0	0	437	4	3	602	0	1	1221	0	0			
456	4	3	487	4	4	348	0	0	409	0	0	609	0	0	711	0	0	1559	0	0			
667	0	0	1574	0	0	489	0	1	654	0	0	1512	0	0	919	0	0	1005	0	0			
230	0	1	1499	0	0	895	0	1	1414	0	0	245	0	0	410	0	0	1463	0	0			
514	0	0	4114	4	4	3188	3	2	359	3	3	3108	3	3	879	0	1	4233	4	4			
911	0	0	695	0	0	76	0	0	4106	4	4	394	0	1	1230	0	1	488	0	1			
721	0	0	138	0	0	552	0	1	572	0	0	1281	0	0	1173	0	0	2120	2	2			
513	0	0	244	0	1	445	4	4	611	0	0	256	2	2	1386	0	0	439	0	0			
4	0	1	792	0	1	692	0	1	596	0	0	1189	0	0	687	0	0	79	0	0			
4207	4	4	777	0	0	1370	0	1	316	0	0	1462	0	0	453	0	0	387	0	0			
3135	3	2	351	0	0	356	0	0	1278	0	0	97	0	0	28	0	0	300	0	0			
723	0	0	468	0	0	599	0	1	666	0	0	4216	4	2	2119	2	2	767	0	0			
53	0	0	4150	4	4	1422	0	1	5	0	0	1497	0	0	4178	4	4	4244	4	1			
682	0	1	109	0	0	548	0	1	1078	0	0	1518	0	1	894	0	1	285	2	2			
237	2	2	78	0	0	891	0	0	3101	3	3	1302	0	0	302	0	0	443	4	4			
2114	2	2	1457	0	1	1201	0	0	4219	4	3	373	3	3	282	0	0	271	0	0			
255	0	0	807	0	0	3106	3	3	323	3	3	491	0	0	153	0	0	3162	3	2			
464	0	0	152	0	1	543	0	1	1314	0	0	0	0	0	4218	4	3	1301	0	0			
1507	0	0	318	0	1	23	0	0	657	0	0	571	0	1	338	3	3	607	0	1			
394	3	3	376	0	1	1061	0	0	861	0	0	793	0	0	386	0	0	311	3	3			
193	0	0	2101	2	2	4205	4	4	4230	4	3	850	0	1	377	0	0						