

Raphael Almeida Guimarães dos Santos

**Digital Twins and Augmented Observability: An
Architecture Model for Experimental Oil & Gas
Separator Vessel Digitization**

Vitória, ES

13 de setembro de 2021

Raphael Almeida Guimarães dos Santos

**Digital Twins and Augmented Observability: An
Architecture Model for Experimental Oil & Gas Separator
Vessel Digitization**

Master's Dissertation submitted and presented to the Programa de Pós-Graduação em Engenharia Elétrica at the Universidade Federal do Espírito Santo, as a partial requirement to obtain the Master's Degree in Electrical Engineering.

Universidade Federal do Espírito Santo – UFES

Centro Tecnológico

Programa de Pós-Graduação em Engenharia Elétrica

Supervisor: Prof. Dr. Moisés R. N. Ribeiro

Co-supervisor: Prof. Dr. Arnaldo G. Leal-Junior

Vitória, ES

13 de setembro de 2021

Ficha catalográfica disponibilizada pelo Sistema Integrado de Bibliotecas - SIBI/UFES e elaborada pelo autor

S237d Santos, Raphael Almeida Guimarães dos, 1992-
Digital Twins and Augmented Observability : An
Architecture Model for Experimental Oil & Gas Separator
Vessel Digitization / Raphael Almeida Guimarães dos Santos. -
2021.
105 f. : il.

Orientador: Moisés Renato Nunes Ribeiro.
Coorientador: Arnaldo Gomes Leal-Junior.
Dissertação (Mestrado em Engenharia Elétrica) -
Universidade Federal do Espírito Santo, Centro Tecnológico.

1. Gêmeos Digitais. 2. Sistemas Ciberfísicos. 3. Indústria
4.0. 4. Vasos Separadores. 5. Clustering. 6. Visão
Computacional. I. Ribeiro, Moisés Renato Nunes. II. Leal
Junior, Arnaldo Gomes. III. Universidade Federal do Espírito
Santo. Centro Tecnológico. IV. Título.

CDU: 621.3

Raphael Almeida Guimarães dos Santos

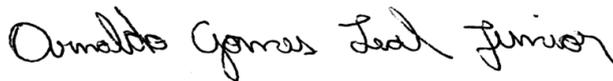
Digital Twins and Augmented Observability: An Architecture Model for Experimental Oil & Gas Separator Vessel Digitization

Master's Dissertation submitted and presented to the Programa de Pós-Graduação em Engenharia Elétrica at the Universidade Federal do Espírito Santo, as a partial requirement to obtain the Master's Degree in Electrical Engineering.

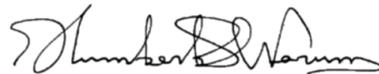
Dissertation approved on **July 22, 2021**. Vitória, ES.



Prof. Dr. Moisés R. N. Ribeiro
Supervisor



Prof. Dr. Arnaldo G. Leal-Junior
Co-Supervisor



Prof. Dr. Humberto S. A. Varum
External Member



Prof. Dr. Celso José Munaro
Internal Member

This dissertation work is dedicated to my parents, Dulce and Samuel, whose effort, love and values brought me up to be a better individual and have directed me to the right choices.

Acknowledgements

First and foremost, I praise **God**, for providing me strength and granting me the capability to proceed with faith and perseverance.

I want to express my deep gratitude to my parents, **Dulce** and **Samuel**, for all the love and support offered to me throughout my academic and life trajectory. Without their teachings I could not have come this far and become what I am today. I thank my sister, **Lorena**, for all the daily support given and partnership, which means a lot to me. I also thank my fiancée, **Thais**, for the company, for the comprehension and for always be on my side supporting me on all the paths I decided to take. Sincerely, thank you all for everything.

I would like to thank my esteemed supervisor Prof. Dr **Moises R. N. Ribeiro** for the trust, the insightful teaching, the valuable advises, for the detailed work reviews, for his patience and guidance during the entire master's course. Thanks to my co-supervisor Prof. Dr. **Arnaldo G. Leal-Junior** for the thoughtful guidance and for the support during the whole period of the study. Also, I must thank Prof. Dr. **Anselmo Frizera Neto** and Prof. Dr. **Maria José Pontes** for all the assistance and opportunity offered to me since the beginning of this course. I am grateful to the Prof. Dr. **Lucas Frizera Encarnação** and to the Prof. Dr. **André Ferreira** for having helped me during my ingress into the course. Finally, thanks to all my colleagues with whom I have the chance to learn and share edifying knowledge during classes and work at NERDS and LabTel.

I also appreciate the funding that helped this work to be realized. This financial support was provided by Petróleo Brasileiro S.A. (Petrobras) (2017/00702-6) that funds the project “Fibra Óptica na Medição de Nível e de Interface Água-Óleo em Tanques de Produção” that has been lead by the Programa de Pós-Graduação em Engenharia Elétrica of Universidade Federal do Espírito Santo.

*“Success is born out of wanting, determination and persistence
in reaching a goal. Even if you don’t reach the target,
whoever seeks and overcomes obstacles will,
at the very least, do admirable things.
(José de Alencar)*

Resumo

A transformação industrial digital tem alavancado o conceito de Gêmeos Digitais e Sistemas Ciberfísicos no contexto da Indústria 4.0. Enquanto os sistemas ciberfísicos são sistemas multidimensionais e complexos que integram computação, comunicação e controle de sistemas físicos dinâmicos, os gêmeos digitais estão relacionados a modelos de alta fidelidade de elementos físicos em um espaço virtual. Seu objetivo é simular o mundo físico e prover *feedbacks* quase em tempo real para cenários hipotéticos a fim de auxiliar na tomada de decisões operacionais. Vasos separadores são elementos-chave no processo de separação água-óleo na indústria de óleo e gás. Basicamente, são vasos pressurizados que podem estar sujeitos a falhas estruturais como fadiga se não forem monitorados, mantidos e operados de modo adequado. O sensoriamento convencional unicamente não pode atender às diversas necessidades dos gêmeos digitais, e a visão computacional vem em mãos para fornecer uma maneira moderna de estimar múltiplos parâmetros. Assim, este trabalho tem como objetivo propor um modelo de arquitetura para um gêmeo digital de um vaso separador experimental englobando técnicas modernas de sensoriamento e processamento de dados. O vaso separador experimental é um protótipo multimaterial construído com aço carbono e policarbonato monitorado por uma técnica de medição de perfil de densidade sem contato que baseia-se em abordagens de visão computacional e regressão não-linear. Esta informação é consumida pelos modelos digitais estruturais do gêmeo digital, que são customizados para este vaso separador experimental. Dessa forma, os modelos digitais permitem um sensoriamento aprimorado extrapolando dados de poucos sensores, para compor um conjunto de informações ricas a serem fornecidas aos operadores. Isso desencadeia cenários hipotéticos a serem testados rapidamente além da inferência de indicadores de manutenção. Com o intuito de ilustrar essas funcionalidades, um cenário pressurizado é considerado para o vaso separador experimental originalmente projetado para operação a 1 atm. As contribuições do trabalho são apresentadas por meio de uma simples abordagem de predição de falhas estruturais sob cargas estáticas e dinâmicas como também por meio da rastreabilidade de pontos críticos. Também, é proposto um novo método para estimar interfaces de nível em líquidos multifásicos utilizando técnicas de *clustering*. Portanto, conclui-se que, durante a operação, os gêmeos digitais possuem potencial para agregar novas dimensões de observabilidade ao controle de processo de vasos separadores.

Palavras-chave: gêmeos digitais; sistemas ciberfísicos; Indústria 4.0; vasos separadores; *clustering*; visão computacional; observabilidade.

Abstract

Digital industrial transformation has leveraged the concept of Digital Twins (DTs) and Cyber-Physical Systems (CPSs) in the Industry 4.0 context. While CPS are multidimensional and complex systems that integrate computation, communication and control of dynamic physical systems, DTs are related to high-fidelity models of physical elements in a virtual space. Its goal is to simulate the physical world and provide near real-time feedbacks to "what if" scenarios in order to assist operational decision-making. Separator vessels are key elements in water-oil separation process in oil & gas industry. Basically, they are pressurized vessels that can be subjected to structural failures like fatigue if not adequately monitored, maintained and operated. Conventional sensing alone cannot match the diverse needs of DTs, and computer vision and machine learning come in hand to provide a modern way to estimate multiple parameters. Thus, this work aims at proposing an architectural model to an experimental separator vessel DT encompassing modern sensing and data processing techniques. The experimental separator vessel is a multi-material prototype built with carbon steel and polycarbonate monitored by a non-contact sensing technique for density profiling based on computer vision technique and non-linear regression approaches. This information is consumed by the DT's structural digital models, which are custom built for this experimental separator vessel. This way, digital models enable an improved sensing, i.e., extrapolating data from few sensors, to compose a rich information set to be provided to operators. This unleashes "what if" scenarios to be quickly tested and maintenance indicators to be inferred. In order to illustrate these functionalities, a pressurized scenario is considered for this experimental separator vessel originally designed for 1 atm operation. The work contributions are presented through a simple structural failure prediction approach under static and dynamic loads as well as through critical points traceability. Also, it is proposed a new method for estimate level interfaces in multiphase liquids by using clustering techniques. Thus, it is concluded that, at operation time, DTs has the potential of aggregating to separator vessel process control new dimensions of observability.

Key-words: Digital Twins; Cyber-Physical Systems; Industry 4.0; separator vessels; clustering; computer vision; observability.

List of Figures

Figure 1.1.1–General DT architecture model.	20
Figure 1.1.2–Schematic of a horizontal three-phase separator vessel with interface level control and weir (STEWART; ARNOLD, 2011).	22
Figure 1.1.3–Nucleonic density profiler (TRACERCO, 2019).	23
Figure 1.4.1–Methodology overview for the case study.	26
Figure 2.1.1–Internal loads in free body diagrams: (a) full body, (b) sectioned body, (c) resultant force and moment (HIBBELER, 2018).	28
Figure 2.1.2–Stress definition: (a) resultant force acting on small area, (b) stresses acting on a volume, (c) stresses acting on a volume (HIBBELER, 2018).	29
Figure 2.1.3–General state of stress (HIBBELER, 2018).	30
Figure 2.1.4–Deformation of a body due to an axial load (HIBBELER, 2018).	32
Figure 2.1.5–Deformation of a body due to a tangential load (HIBBELER, 2018).	32
Figure 2.1.6–Stress-strain diagram of low-carbon steel (BEER et al., 2020).	33
Figure 2.2.1–Pressure profile of a fluid at rest (A.; CIMBALA, 2020).	34
Figure 2.3.1–Constant amplitude cycling and the associated nomenclature: (a) completely reversed stressing, (b) nonzero mean stress, and (c) zero-to-tension stressing (DOWLING; PRASAD; NARAYANASAMY, 2013).	38
Figure 2.3.2–Stress versus life (S-N) diagram plotted from the results of completely reversed axial fatigue tests. Material: UNS G41300 steel. (BUDYNAS; NISBETT; SHIGLEY, 2016).	39
Figure 3.1.1–Multiphase level measurement system and its resulting density profile (BERTHOLD, 2021).	46
Figure 3.1.2–Density profile samples measured from the production separator vessel: (a) sample #1, (b) sample #2, (c) sample #3, and (d) sample #4.	47
Figure 3.1.3–Principle of interface modeling through s-shaped curve.	49
Figure 3.1.4–Non-linear regression models subplots. The ordinate axis is normalized density and the abscissa axis is normalized level.	51
Figure 3.1.5–Root mean square errors for each s-shaped curve fitting.	52
Figure 3.1.6–Coefficients of determination for each s-shaped curve fitting.	53
Figure 3.1.7–Proposed density profile modeling workflow.	54
Figure 3.1.8–Resultant model for the entire density profile data.	55
Figure 3.2.1–Density profile data sample: (a) raw data, and (b) hard-coded grouped data.	57
Figure 3.2.2–Clustering results: (a) K-means, (b) Hierarchical Agglomerative, (c) Spectral, (d) DBSCAN.	58

Figure 3.2.3–Clustering performance test results: (a) accuracy, and (b) execution time.	59
Figure 3.2.4–DBSCAN accuracy values in function of δ and ν .	59
Figure 3.3.1–Proposed computer vision workflow diagram.	61
Figure 3.3.2–Oil and water content within the experimental separator vessel: (a) original image frame, (b) gray-scale image frame, and (c) thresholded gray-scale image frame.	62
Figure 3.3.3–Gray-scale image frame windowing, and averaging.	63
Figure 3.3.4–Gray-scale average array derivative.	64
Figure 3.3.5–Density profiling: (a) emulsion layer non-linear regression model, and (b) full density profile analytical model.	65
Figure 4.1.1–Separator prototype used as the physical element of the DT: (a) equipment design (b) equipment photography.	67
Figure 4.2.1–Proposed DT architecture model.	68
Figure 4.2.2–Operational technology integrated with cloud and edge infrastructures for implementing the pipeline from the DT architecture model.	69
Figure 4.4.1–Oil and water content image and thresholded frames: (a) at minimum load, (b) at maximum load.	71
Figure 4.4.2–Gray-scale average arrays: (a) at minimum load, (b) at maximum load.	71
Figure 4.4.3–Non-linear regression models: (a) at minimum load, (b) at maximum load.	72
Figure 4.4.4–Gray-scale first derivatives: (a) at minimum load, (b) at maximum load.	72
Figure 4.4.5–Density profiling results: (a) at minimum load, (b) at maximum load.	73
Figure 4.4.6–Pressure profile outcomes: (a) at minimum load, (b) at maximum load.	74
Figure 4.5.1–Separator prototype orthographic views for (a) geometric and (b) structural simplified models.	75
Figure 4.5.2–Integration between structural and load models.	76
Figure 4.5.3–Depressurized separator: (a) stress distribution, and (b) strain distribution. The stress scale is in MPa.	77
Figure 4.5.4–Pressurized separator: (a) stress distribution, and (b) strain distribution. The stress scale is in MPa.	78
Figure 4.6.1–Density profiling clustering results: (a) at minimum load, (b) at maximum load.	80
Figure 4.6.2–Von Mises effective stress distributions: (a) in carbon steel (depressurized), (b) in polycarbonate (depressurized), (c) in carbon steel (pressurized), (d) in polycarbonate (pressurized).	81
Figure A.0.1–Extraction from CAD design of the experimental separator vessel (page 2) (2SOLVE, 2019).	103
Figure A.0.2–Extraction from CAD design of the experimental separator vessel (page 5) (2SOLVE, 2019).	104

List of Tables

Table 1 – Material parameters of carbon steel and polycarbonate. 75

List of abbreviations and acronyms

ICT	Information and Communication Technologies
IoT	Internet of Things
CPS	Cyber-Physical Systems
RMS	Reconfigurable Manufacturing Systems
MES	Manufacturing Execution Systems
SCADA	Supervisory Control and Data Acquisition
DT	Digital Twin
PLM	Product Life Cycle Management
NASA	National Aeronautics and Space Administration
HTHA	High Temperature Hydrogen Attack
SI	International System of Units
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
W/O	Water-in-Oil
O/W	Oil-in-Water
W/O/W	Water-in-Oil-in Water
O/W/O	Oil-in-Water-in-Oil
RMSE	Root Mean Square Error
FEM	Finite Element Method
NTP	Normal Temperature and Pressure
CAD	Computer-Aided Design
FWKO	Free Water Knockout Vessel

List of symbols

σ	Normal stress
τ	Shear stress
ϵ	Normal strain
γ	Shear strain
σ_Y	Tensile yield strength
σ_U	Ultimate strength
σ_B	Breaking strength
σ_F	Fatigue strength
σ'_f	Fatigue strength coefficient
σ_m	Mean stress
σ_{max}	Maximum stress
σ_{min}	Minimum stress
σ_a	Stress amplitude
σ_{ar}	Reversed stress
b	Fatigue strength exponent
ρ	Density
γ_s	Specific weight
g	Gravitational force
$\bar{\sigma}$	Effective stress of von Mises
N	Safety factor on static load
N_f	Number of cycles to fatigue
R^2	Coefficient of determination
ν	Minimum number of points in a dense region
δ	Maximum distance between neighbors in a single cluster

Contents

1	INTRODUCTION	17
1.1	Background	17
1.1.1	The Industry 4.0	17
1.1.2	Cyber-physical Systems and Digital Twins	19
1.1.3	Separator Vessels in Oil & Gas Industry	22
1.2	Motivations for Augmented Observability	24
1.3	Objectives	25
1.4	Dissertation Overview	26
2	FOUNDATIONS ON MECHANICS OF MATERIALS AND CLUSTERING ANALYSIS	27
2.1	Mechanics of Materials	27
2.1.1	Internal Resultant Loadings	27
2.1.2	Stress	28
2.1.2.1	Normal Stress	28
2.1.2.2	Shear Stress	29
2.1.2.3	General State of Stress	29
2.1.3	Principal Stresses	30
2.1.4	Strain	31
2.1.4.1	Normal Strain	31
2.1.4.2	Shear Strain	32
2.1.5	Mechanical Properties of Materials	33
2.2	Hydrostatic Pressure Distribution	34
2.3	Theories of Failure	35
2.3.1	Static Failure Theory	36
2.3.2	Distortion-Energy Theory	36
2.3.3	Dynamic Failure Theory	37
2.3.4	Stress Versus Life (S-N) Curve	38
2.3.5	Life Estimates with Mean Stress	40
2.3.6	One Approach to Multiaxial Fatigue	40
2.4	Clustering	41
2.4.1	Centroid-based Clustering	41
2.4.2	Hierarchical Agglomerative-based Clustering	42
2.4.3	Spectral-based Clustering	43
2.4.4	Density-based Clustering	43

2.5	Final Remarks	44
3	DENSITY PROFILING, CLUSTERING METHODS AND A COMPUTER VISION BASED PROFILER	45
3.1	Density Profiling Through Non-Linear Regression Modeling	45
3.1.1	Crude Oil Density Profile Experimental Data	46
3.1.1.1	Crude Oil Emulsion	47
3.1.1.2	Crude Oil Foam	48
3.1.2	Regression Modeling Hypothesis	48
3.1.3	Validation of Non-linear Models	50
3.1.4	Proposed Density Profile Model	53
3.2	Interface Level Identification Through Density Profile Clustering	55
3.2.1	Clustering Density Profiles	57
3.3	Non-contact Density Profiling Through Computer Vision	60
3.3.1	Proposed Computer Vision Workflow	61
3.3.2	Pre-processing the Image Data	62
3.3.3	Thresholding the Gray-scale Frame	62
3.3.4	Differentiating the Data Sequence	63
3.3.5	Density Profiling	64
3.4	Final Remarks	65
4	EXPERIMENTAL SEPARATOR VESSEL DIGITAL TWIN ARCHITECTURE MODEL	66
4.1	The Physical Element	66
4.2	Proposed Architecture Model	67
4.3	Experiment Methodology	70
4.4	Pre-processing Image Data	70
4.5	Processing Load Model Data	74
4.6	Post-processing Stress and Strain Data	79
4.6.1	Predictions and Insights	80
4.6.1.1	Interface Levels Identification	80
4.6.1.2	Structural Failures	81
4.6.1.2.1	Static Load	82
4.6.1.2.2	Dynamic Load	82
5	CONCLUSIONS AND FUTURE WORK	84
5.1	Conclusions and Contributions	84
5.2	Publications	84
5.3	Future Work	85

BIBLIOGRAPHY	86
APPENDIX	98
APPENDIX A – CLUSTERING PSEUDOCODES	99
ANNEX	101
ANNEX A – EXPERIMENTAL SEPARATOR VESSEL’S CAD DE- SIGN EXTRACTIONS	102

1 Introduction

This dissertation brings an early effort of digital twins (DTs) implementation by proposing an architectural model for the digitization of an experimental separator vessel. This is intended to augment the observability of this equipment by providing extra knowledge by using a data processing pipeline that begins from sensors and could end as a feedback for the gas-oil-water separation process control. This chapter will present the context in which DTs are inserted as well as its concept, the introduction to the separator vessel as the object of study, the motivations behind this work and the established objectives to towards the experimental separator vessel digitization.

1.1 Background

1.1.1 The Industry 4.0

The integration between manufacturing operations systems and information and communication technologies (ICT), e.g., Internet of Things (IoT) - also known as Cyber-Physical Systems (CPS) - describes the new model of manufacturing organization and digital industrial transformation called Industry 4.0 (JESCHKE et al., 2017; WANG; TÖRNGREN; ONORI, 2015). This new industrial stage is changing competition rules, the structure of manufacturing and also redirecting the customers' demands (GILCHRIST, 2016; BARTODZIEJ, 2017). Such transformation are motivated by the adoption of IoT concepts and digitization of factories (Dregger et al., 2016; LASI et al., 2014; WANG; TÖRNGREN; ONORI, 2015). From the market perspective, digital technologies allow companies to offer new digital solutions for customers, such as internet-based services embedded in products (AYALA et al., 2017; COREYNEN; MATTHYSSENS; Van Bockhaven, 2017). On the other hand, from the operational context, digital technologies such as CPS, are proposed to reduce set-up and processing times, labor and material costs, resulting in a improved productivity of production process (BRETTEL et al., 2014; JESCHKE et al., 2017).

Regarding the industrial revolutions chronology, some scholars and practitioners have considered four main industry changes, while the Industry 4.0 is the last one and the current industry transformation (QIN; LIU; GROSVENOR, 2016). Starting in the period between 1760 and 1840, the first industry revolution was characterized by the steam machine creation (DALENOGARE et al., 2018). In the end of XIX century, the utilization of electricity in industrial processes delimited second industry revolution (DALENOGARE et al., 2018). After, during the 1960 decade, the third revolution was marked by the ICT

and industrial automation advent (DALENOGARE et al., 2018). Finally, in the recent time, the fourth industrial revolution, or Industry 4.0, emerged from several developed countries and it was consolidated in a German public-private initiative to build smart factories by the integration of physical objects with digital technologies (BRETTEL et al., 2014; Hermann; Pentek; Otto, 2016). The broad change in the manufacturing systems connectivity due to the integration of ICT, IoT and machines in CPS is the fundamental element that describes this contemporary industrial stage (KAGERMANN, 2015). The way how platforms are coming to be connected to each other expresses the new industrial age called Industry 4.0 (LASI et al., 2014; REISCHAUER, 2018), as it considers integration of several different dimensions of business, with a main concern on manufacturing issues, based on advance manufacturing technologies (Saldivar et al., 2015; FATORACHIAN; KAZEMI, 2018). In this sense, Industry 4.0 can be comprehended as an outcome of the increasing digitization of companies, in particular concerning to manufacturing methods. (KAGERMANN, 2015; SCHUMACHER; EROL; SIHN, 2016).

The essence of Industry 4.0 concepts resides in allowing organizations to provide more agile production processes and to analyze massive volumes of data in real time, thus enhancing the strategic and operational decision-making (KAGERMANN, 2015; PORTER; SIMONS; HARRIS, 2014). This new industrial stage can be achieved thanks to the use of ICTs in industrial environments (KAGERMANN, 2015) and the cheapening of sensors, which increases their installation in physical objects (BRETTEL et al., 2014; PORTER; SIMONS; HARRIS, 2014; Bangemann et al., 2016). These advances paved the way for the development of embedded and connected systems (Jazdi, 2014; KAGERMANN, 2015; BRETTEL et al., 2014) that seek to monitor and control equipment, conveyors, and products through a feedback loop that gathers a large amount of data (big data) and updates virtual models with information from physical processes (WANG; TÖRNGREN; ONORI, 2015; WANG et al., 2016; GILCHRIST, 2016). Thus, various developments have evolved and have been implemented in production processes since the development of digital manufacturing in the 1980s, including the following technologies: computer-aided design and manufacturing for development projects based on computerized systems; integrated engineering systems for information exchange in product development and manufacturing; digital automation with sensors for monitoring through data gathering; flexible manufacturing lines to promote Reconfigurable Manufacturing Systems (RMS) and to enable integration and rearrangement between product and the industrial environment in a cost-effective way; Manufacturing Execution Systems (MES) and Supervisory Control and Data Acquisition (SCADA) for the monitoring of shop floor and remote control of production with real time data collection; simulations and analysis of virtual models for engineering projects and commissioning model-based systems; big data collection and analysis for applications in predictive analytics; digital product-service systems for the incorporation of digital services in products; additive manufacturing for modular manufac-

turing systems; cloud computing for on-demand manufacturing facilities (DALENOGARE et al., 2018; KAGERMANN, 2015; WANG et al., 2016; YU; XU; LU, 2015).

These innovations endorse three key benefits of Industry 4.0: vertical integration, horizontal integration, and end-to-end engineering (KAGERMANN, 2015; WANG; TÖRNGREN; ONORI, 2015). The integration of ICT systems in various hierarchical layers of an organization, reflecting the association between the manufacturing and management levels of a factory, is referred to as vertical integration (KAGERMANN, 2015). Horizontal integration, on the other hand, entails cooperation between businesses, as well as the sharing of resources and real time information exchange (BRETTEL et al., 2014). End-to-end engineering is the integration of engineering in a product's supply chain, from development and production to after-sales (KAGERMANN, 2015; BRETTEL et al., 2014; GILCHRIST, 2016).

1.1.2 Cyber-physical Systems and Digital Twins

Smart manufacturing refers to a future state of production, in which real time transmission and analysis of data from around the product life cycle, together with model-based simulation and optimization, generate intelligence to provide positive effects on all facets of manufacturing. In this context, knowledge-based intelligent manufacturing is giving way to data-driven and knowledge-enabled smart manufacturing (ZHONG et al., 2017). CPS convergence is both a requirement and a central component of smart manufacturing (TAO et al., 2018b; O'DONOVAN et al., 2015). CPS and DTs have gotten a lot of recognition from academia, business, and government as the desired methods of such convergence (TAO et al., 2019a). Despite both CPS and DTs are used to describe cyber-physical integration, their definitions are not the same. CPS are multidimensional and complex structures that bridge the gap between the cyber and physical worlds. CPS provides real-time sensing, information feedback, dynamic control, and other services through the convergence and coordination of computation, networking, and control (HU et al., 2012; Liu et al., 2017). DTs, in turn, creates high-fidelity simulated representations of physical structures in virtual space to simulate their behaviors in the real world and provide feedback (GRIEVES, 2014). A DT is a bi-directional dynamic mapping process that breaks down obstacles in the product life cycle and provide a full digital footprint of products (TAO et al., 2018a). Sensors strategically located in an equipment or even in a manufacturing process produce signals that enable the DT to collect operational and environmental data from the real-world process (PARROTT; WARSHAW, 2017). Therefore, DTs allow companies to predict and promptly detect physical issues more accurately, optimize manufacturing processes, and enhance products (TAO et al., 2019a).

In summary, CPS are characterized as the combination of computational and physical processes (La; Kim, 2010), whereas DT refers to the idea of using a digital copy

of physical systems to perform real time optimization (SÖDERBERG et al., 2017). Both CPS and DTs have two components: the physical part and the cyber/digital part. The physical part senses and collects data, and executes decisions from the cyber/digital part, while the cyber/digital part analyzes and processes data, and then makes decisions. Each part can influence and affect each other through this intensive connection (LEE, 2006).

The purpose of DTs consists of providing a concise physical and functional description of a component, product, or system (SÖDERBERG et al., 2017). The main step is to create high-fidelity virtual models in order to realistically replicate the geometries, physical properties, behaviors and rules pertaining to the physical world. (Tao; Zhang, 2017). This is possible thanks to the sensing data collected from the physical element in order to continuously create status snapshots that feeds the virtual entity. These computer models are not only geometrically and structurally compatible with the physical elements, but they can also simulate their spatio-temporal status, actions and features (SCHLEICH et al., 2017; DEBROY et al., 2017). The virtual models and physical entities tend to have a similar appearance, like twins, and equivalent behaviors, such as a mirror reflection. Furthermore, models in the digital space may use feedback to directly optimize operations and modify physical process (Vachálek et al., 2017). Through the bi-directional mapping, the physical and virtual entities co-evolve (GLAESSGEN; STARGEL, 2012). A general architecture model of DTs based on this conception is shown in Figure 1.1.1. Essentially, DTs have two components: the physical and digital spaces. The physical space senses and collects data, and executes decisions from the digital space, while the digital space analyzes and processes data, and then makes decisions. Each part can influence and affect each other through this intensive connection (LEE, 2006).

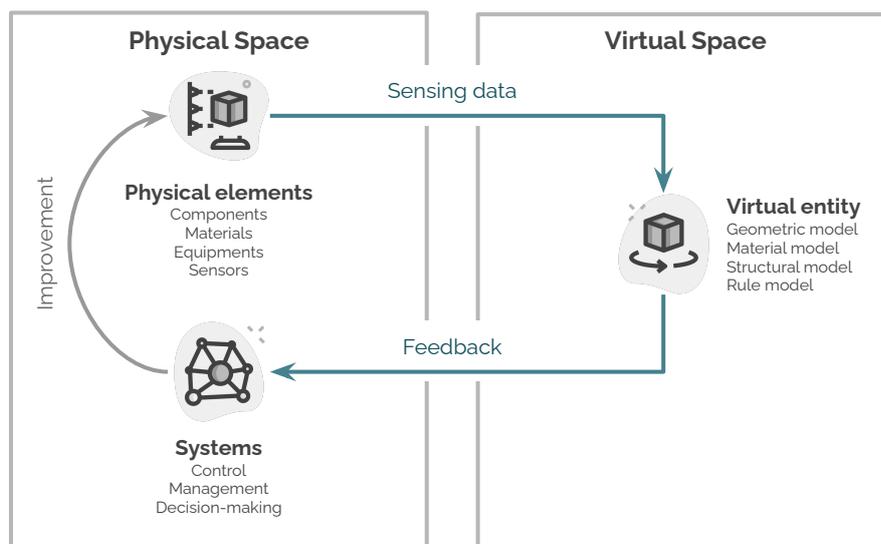


Figure 1.1.1 – General DT architecture model.

At the time in which the DT term was coined by Michael Grieves in a 2003 presentation on product life cycle management (PLM) at the University of Michigan (GRIEVES, 2014), almost no relevant studies or applications were available, due to the limitations and immaturity of the technology (GRIEVES, 2014). In order to solve increasingly complex engineering system problems, the National Aeronautics and Space Administration (NASA) and the US Air Force have used DTs in their health maintenance and for predicting the remaining useful life of aerospace vehicles (NEGRI; FUMAGALLI; MACCHI, 2017; GLAESSGEN; STARGEL, 2012). Recently, developments in new ICT are helping DTs to extend. DTs have become a common research subject since they provide a new way to synchronize physical actions with the virtual space. DTs have been used in a variety of industrial segments for product design (TAO et al., 2019b), production line design (ZHANG et al., 2017), DT shop floors (Tao; Zhang, 2017), production process optimization (UHLEMANN; LEHMANN; STEINHILPER, 2017), and prognosis and health management (TAO et al., 2018). DT industrial proceedings can also be used in a variety of major companies, including General Electric, Siemens, PTC, Dassault Systems, and Tesla (SCHLEICH et al., 2017), which use DTs to improve product performance, manufacturing flexibility and competitiveness (SCHLEICH et al., 2017).

Since companies can more reliably anticipate and diagnose physical problems with the aid of DTs, as stated before, the product reliability can be supported by this technology. The term reliability can be considered in a probabilistic or in a deterministic approach (MODARRES; KAMINSKIY; KRIVTSOV, 2017). The ability of an element, i.e., a product or a system, to perform under designated operating conditions for a specified duration of time or number of cycles is the most commonly agreed concept of reliability (MODARRES; KAMINSKIY; KRIVTSOV, 2017). An element's ability can be defined either probabilistically or deterministically (MODARRES; KAMINSKIY; KRIVTSOV, 2017). The deterministic approach is concerned with determining how and why an object fails, as well as how it can be configured and evaluated to prevent those failures from occurrence or recurrence (MODARRES; KAMINSKIY; KRIVTSOV, 2017). This includes deterministic analysis and study of field failure reports, understanding the physics of failure, the role and degree of testing and inspection, performing redesign (MODARRES; KAMINSKIY; KRIVTSOV, 2017). In this work, the methodology is based on the deterministic approach.

One of the most important maintenance techniques today is the advent of technologies to effectively verify the state of an equipment, which contributes to improvements in predictive models (Hoffmann Souza et al., 2020). Detecting device faults or risks when they emerge, and recommending repairs to minimize the risk of failure, is an relevant addition to productive processes (KIANGALA; WANG, 2018; Terrissa et al., 2016). To keep up with the progression of technology embedded in products and manufacturing equipment, maintenance strategies associated with reliability must evolve on a daily basis

(SONY; NAIK, 2020). Quality in manufacturing, when paired with emerging technology, is a necessity for successfully delivering reliable products, and diligent product reliability assurance is often a crucial routine task for production. In this scenario, Industry 4.0 establishes a desirable environment for the development of reliability models (HE et al., 2019; LI et al., 2019), as well as a background provided by technologies such as DTs, where the manufacturing system can be simulated and, thus, the equipment health monitored (CHENG et al., 2018; ROSEN et al., 2015).

1.1.3 Separator Vessels in Oil & Gas Industry

In oil & gas industry, one of the many production facility's roles is to separate the well stream into three components, generally referred to as phases (oil, gas, and water), in order to process these phases into some marketable product or dispose them in an environmentally acceptable manner (STEWART; ARNOLD, 2011). In mechanical devices called separators, gas is flashed from the liquid components, whereas the free water is separated from the oil (STEWART; ARNOLD, 2011). Separators are pressure vessels that may be installed either in an onshore processing station or on an offshore platform (PETROWIKI, 2021). Depending on the vessel configuration, they can be divided into horizontal, vertical or even spherical, as well as in terms of fluids to be separated, they may be classified as gas/liquid two-phase separators or oil/gas/water three-phase separators (PETROWIKI, 2021). Figure 1.1.2 illustrates a schematic of the horizontal three-phase separator vessel.

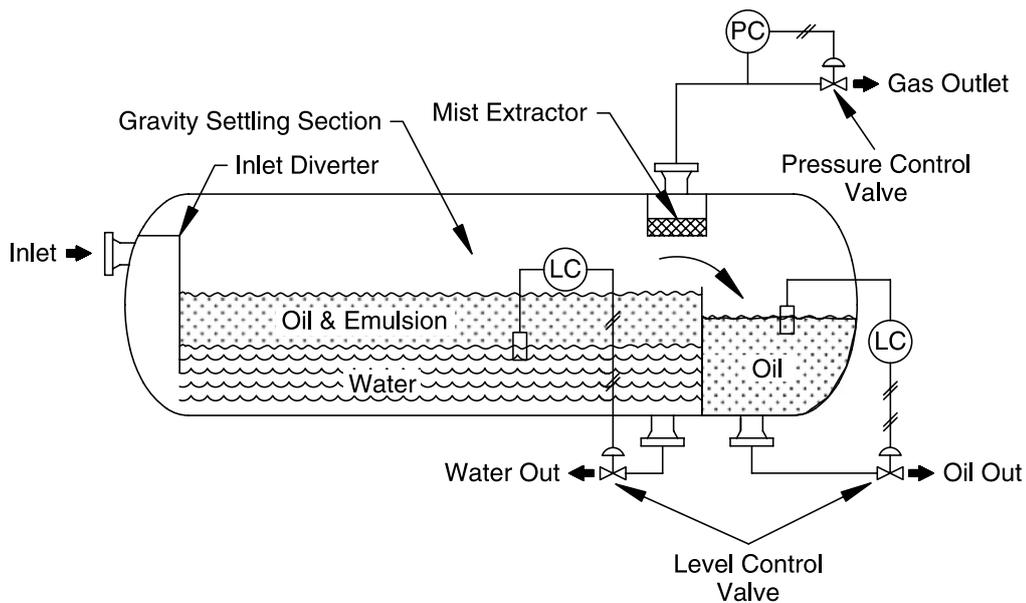


Figure 1.1.2 – Schematic of a horizontal three-phase separator vessel with interface level control and weir (STEWART; ARNOLD, 2011).

Throughout the separation process in separator vessels, the exhaust valves must be properly controlled in order to maintain liquid and pressure levels in accordance with the pre-established setpoints. To monitor fluid levels, profiling instruments are often employed, which measures the vertical density distribution inside the equipment. Depending on the type of the profiler, measurement principles may include inductive, electrical conductivity, radioactive, and electromagnetic approach ([SULZER, 2021](#); [ROCSOLE, 2021](#); [TRACERCO, 2019](#); [BERTHOLD, 2021](#)). Figure 1.1.3 shows a nucleonic density profiler, whose measurement principle consists in the emission and detection of low intensity gamma rays.



Figure 1.1.3 – Nucleonic density profiler ([TRACERCO, 2019](#)).

Separator vessels operate over a wide range of pressures and temperatures and are exposed to a wide variety of operating environments ([ASM, 2002](#)). Therefore, this equipment can be affected by several failures mechanisms (e.g., corrosion, creep, fatigue, cracking, brittle and ductile fractures) caused by defective design, fabrication, and inspection, especially for shell components ([ASM, 2002](#)). Operating pressure is an important design consideration because pressure combined with structural constraints determines the stresses in pressure boundaries both at steady state operating conditions and during transient state ([ASM, 2002](#)). Verifying whether failures occurred as a result of, or during, abnormal circumstances is an important aspect of the analysis ([ASM, 2002](#)). The service life of pressure boundaries is profoundly influenced by cyclic operation and designers are not always able to predict the equipment operational regime imposed by the owner and operators ([ASM, 2002](#)). Fatigue fractures are caused by cyclic stressing, which progressively propagates cracks until the part can no longer support the applied load ([ASM, 2002](#)). Pressure vessels are subject to high static stresses arising from the pressure of contained liquids or gases ([ASM, 2002](#)).

1.2 Motivations for Augmented Observability

Efficient operation of the separation process is of crucial relevance in the production of crude oil. Crude oil is combined with water, gas, and sediment at the wellhead and must be separated from these components before it can be sold (CHARLTON; LEES, 2002). In order to separate each component from crude oil mixture, the liquid and interface levels must be properly measured to provide feedback to the level control system. Nowadays, there are several instruments of density profile applied to this specific field. Generally, they are classified as nucleonic (TRACERCO, 2019) and non-nucleonic (ROCSOLE, 2021) profilers. Although they are already consolidated in the oil & gas industry, there are drawbacks due to the high cost of acquisition and maintainability, their installation intrusive aspect, and risks concerned to the radiation for nucleonic profilers. Therefore, it is necessary to explore an alternative approach that provides, at the same time, low cost of acquisition and maintainability, non-contact measurement and the absence of required specialized licensing and further safety precautions.

Throughout manufacturing history there have been serious incidents related to pressure vessel failures. On June 11, 2008, in the city of Houston, Texas, USA, a heat exchanger ruptured at a Goodyear facility due to ammonia over-pressurization (CSB, 2008). On February 4, 2010, in the city of Anacorte, Washington, USA, there was an explosion in an almost forty years old heat exchanger at the Tesoro refinery. There were damages in the heat exchanger due to a mechanism known as high temperature hydrogen attack (HTHA), which severely cracked and weakened carbon steel tubing leading to a rupture (CSB, 2010b; CSB, 2010a). On April 3, 2017, in St. Louis, Missouri, USA, a rupture of a pressure vessel caused by corrosion occurred at the Loy Lange Box Company (CSB, 2017; CLANCY, 2019). There was a sudden mechanical integrity failure of the equipment. This rupture separated the bottom of the tank from the rest of the pressure vessel, creating conditions for a steam explosion and launching the vessel through the building about 520 feet before landing in a neighbor company's site.

These incidents may be avoided by doing deeply structural analysis through simulations during the equipment design stage. However, the physical agents and parameters considered at design stage are not guaranteed that will remain the same at the time of operation, since the product would be subjected to diversified working conditions in an industrial plant. In addition, liquid levels in separator vessels must be monitored to ensure optimal performance and to prevent inadvertent discharges of oil in water. This monitoring comes to reduce costs and contribute to the protection of the environment (Holstad et al., 2005).

Thus, there is a need to augment the observability of separator vessels, in the context of petrochemical processes, intending to conduct simulations techniques from design to operational stage. The term observability comes from control system theory, and

it refers to how well a system's internal state can be predicted from its output (MADAN, 1984). In parallel, observability in DT context refers to the ability to monitor and predict the equipment state through the aggregation between sensors data and digital models. Thus, this work presents a way to offer optimization for operation and extra information for gas-oil-water separation process control.

1.3 Objectives

This work proposes an architecture model that serves as a guidance tool to implement an experimental separator vessel DT in order to analyze and predict its structural state and life in operational scenarios. Those structural states will be obtained from a novel methodology consisting in density profile estimate and modeling using computer vision and non-linear regression. This methodology also comes as an alternative method to automatically identify interface levels of oil-water mixture through clustering algorithms. This architecture follows a data driven structure in which sensor data collected from the equipment is transformed across a pipeline. It is thought to be integrated in real processes in order to augment the equipment's observability in terms of visibility, predictability and traceability. It is intended to conduct the simulations from the design stage to the operational stage in order to optimize predictions by analyzing real processes data.

Moreover, the specific objectives of this work can be detailed listed as:

- To perform a literature study regarding structural mechanics to understand failures under static and dynamic loads applied to ductile materials;
- To implement a density and pressure profiling method using a simple computer vision technique and non-linear regression;
- To perform an exploratory study regarding clustering algorithms applied to automatic identification of interface levels present in three-phase liquids containing oil, gas and water;
- To design a DT architecture model of an experimental separator vessel including the proposed profiling and clustering methodologies;
- To validate the proposed architecture using experimental data by simulating operational loading scenarios in order to predict safety factors and fatigue life under static and dynamic loads.

1.4 Dissertation Overview

In this work, the density profiling modeling and interface level identification methods are validated using density measurement data collected from a real separation process by a nucleonic profiler. Then, the density and pressure profiles, estimated by a computer vision approach, will feed digital models of an experimental separator vessel. The density profile is analytically modeled in order to make possible the aggregation between equipment's structural model and sensor measurements. Since the DT architecture knowledge will cover a full density profile, a clustering algorithm will be used to separate each phase into distinct clusters to provide automatic identification of liquid level interfaces. Finally, the experimental separator vessel behavior are simulated under dynamic and static loading operations, observing what-if scenarios and forward-looking predictive analysis. Figure 1.4.1 depicts the methodology overview that will be presented throughout this work.

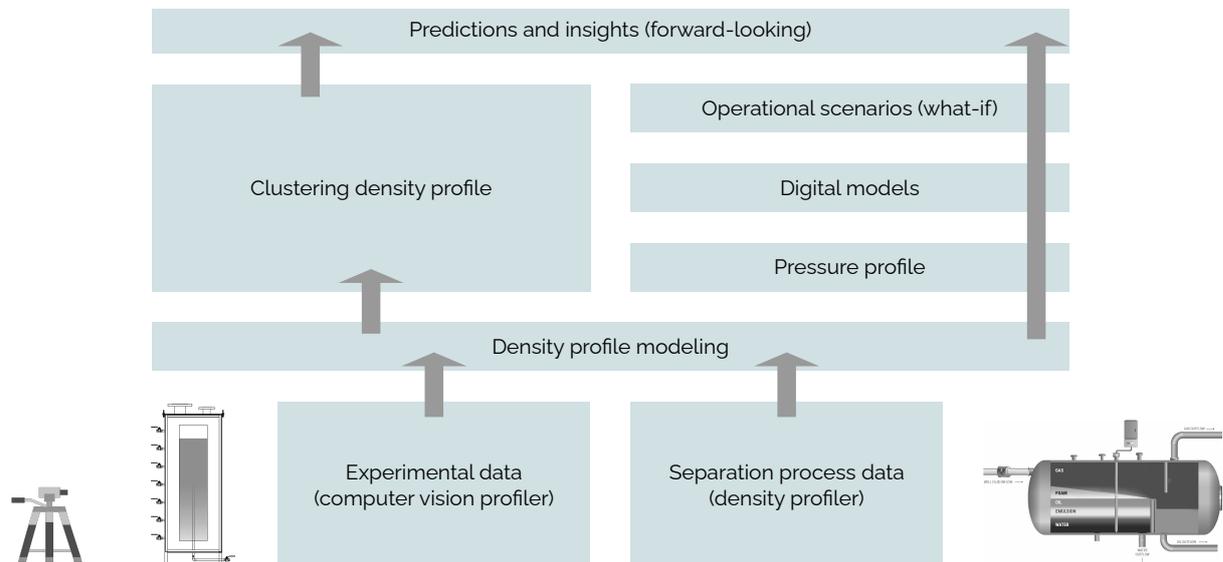


Figure 1.4.1 – Methodology overview for the case study.

This text is organized in five chapters. Chapter 2 refers to the theoretical review where will be explored physical concepts such as mechanics of materials, hydrostatic pressure distribution and failure theories. Also, some clustering techniques will be presented. In Chapter 3, the density profiling and interfaces identification methods will be detailed. The DT architecture model of an experimental separator vessel will be proposed in Chapter 4. Finally, the conclusions and contributions from this work are presented in Chapter 5 as well as additional developments and increments as future work.

2 Foundations on Mechanics of Materials and Clustering Analysis

This chapter explores the fundamentals of mechanics of materials and clustering analysis. The DT architecture model will employ mechanical analysis of the experimental separator vessel by using its structural model loaded by pressure profiles. Therefore, it is important to comprehend the concepts behind mechanical challenges. Also, clustering techniques will be presented, since they are related to the proposed interface level identification approach.

2.1 Mechanics of Materials

Separator vessels are subjected to pressure loads during its operation. Thus, in order to comprehend the effects of stress and strain in such equipment's material due to load application and so to implement its DT, it is necessary to know the fundamental concepts behind structural mechanics analysis.

The effects of stress and pressure in a rigid body is studied in material mechanics, which is a branch of mechanics (HIBBELER, 2018). Strain is a measure of the body's deformation, while stress is related to the strength of the material from which it is produced. Since many of the formulas and construction rules quoted in engineering codes are founded on the principles of this topic, a detailed knowledge of the basics is critical for the design of any machine or structure (HIBBELER, 2018).

2.1.1 Internal Resultant Loadings

Statics is mainly used in material mechanics to assess the resulting loadings that work inside a body. The method of sections is used to do this. Consider the body depicted in Figure 2.1.1 (a), which is maintained in equilibrium by four external forces. It is necessary to move an imaginary segment or break through the area where the internal loadings are acting in order to obtain the internal loadings working on that region within the body (HIBBELER, 2018). After that, the two body parts are splitted, and a free-body diagram of one of them is formed. As this is achieved, an internal force distribution will work on the section's exposed portion, as seen in Figure 2.1.1 (b). These forces are simply the results of the top portion of the body's substance working on the bottom section (HIBBELER, 2018).

The resultant force and moment, \mathbf{F}_R and \mathbf{M}_{RO} , are calculated by applying the

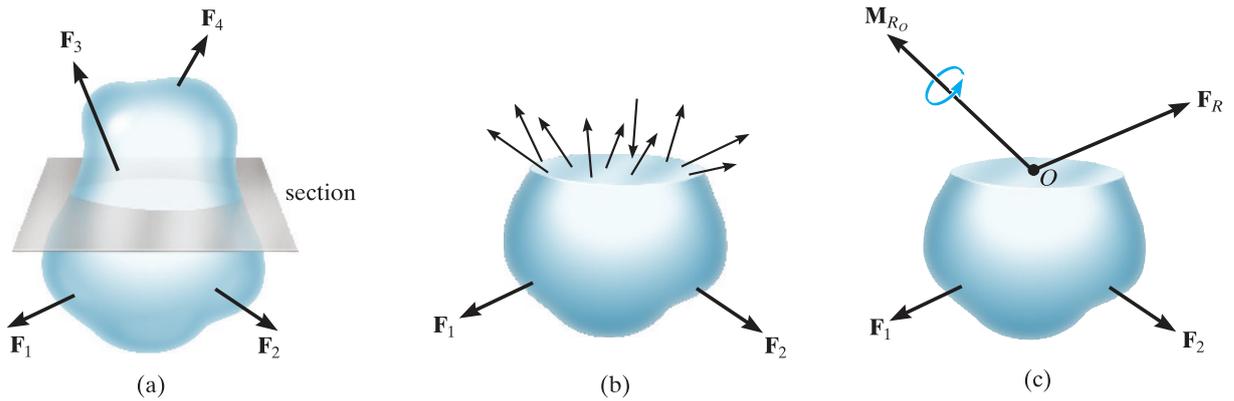


Figure 2.1.1 – Internal loads in free body diagrams: (a) full body, (b) sectioned body, (c) resultant force and moment (HIBBELER, 2018).

equations of equilibrium to the section, as shown in Figure 2.1.1 (c), despite the fact that the precise distribution of this internal loading may be unknown. These loadings operate at point O . In this case, however, the centroid of the sectioned region is often chosen as this point.

2.1.2 Stress

To define stress, firstly, the sectioned area is divided into small sections, ΔA , as shown in Figure 2.1.2 (a). As ΔA becomes infinitesimal, it is considered that the material is continuous, meaning that it's made up of a continuous or consistent distribution of matter with no voids. Also, the material must be cohesive, ensuring that no fractures, cracks, or separations exist within it. Figure 2.1.2 (a) depicts a finite but very small force $\Delta \mathbf{F}$ acting on ΔA . This force, like all the others, has a specific direction, so it is replaced with its three components, namely $\Delta \mathbf{F}_x$, $\Delta \mathbf{F}_y$, and $\Delta \mathbf{F}_z$. The force $\Delta \mathbf{F}$ and its components approach zero, as well as ΔA does; nevertheless, the quotient of force and area approaches a finite limit. The strength of the internal force acting on a particular plane, i.e., area, passing through a point is defined by this quotient, which is called stress (HIBBELER, 2018).

2.1.2.1 Normal Stress

The magnitude of the force $\Delta \mathbf{F}_z$ acting normal to ΔA is referred to as the normal stress, σ , as defined by Equation 2.1:

$$\sigma_z = \lim_{\Delta A \rightarrow 0} \frac{\Delta F_z}{\Delta A} \quad (2.1)$$

If the normal force or stress pulls on ΔA , as shown in Figure 2.1.2 (a), it is tensile stress, whereas if it pushes on ΔA it is compressive stress.

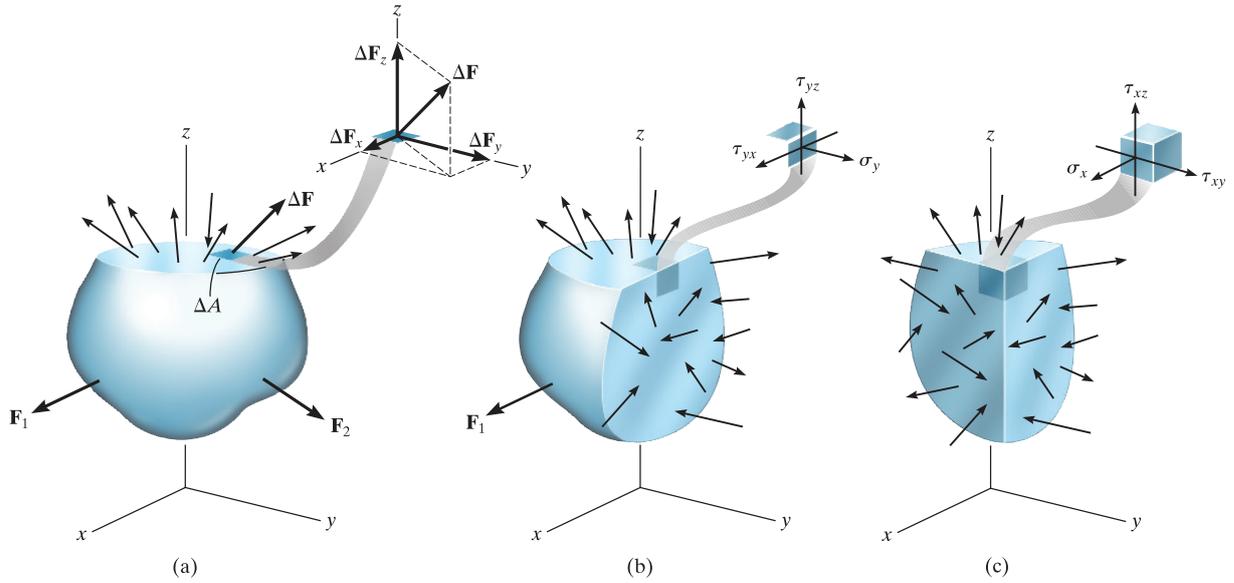


Figure 2.1.2 – Stress definition: (a) resultant force acting on small area, (b) stresses acting on a volume, (c) stresses acting on a volume (HIBBELER, 2018).

2.1.2.2 Shear Stress

The intensity of force acting tangent to ΔA , as shown in Figure 2.1.2 (b), is called the shear stress τ , as defined by Equation 2.2:

$$\begin{aligned}\tau_{yx} &= \lim_{\Delta A \rightarrow 0} \frac{\Delta F_x}{\Delta A} \\ \tau_{yz} &= \lim_{\Delta A \rightarrow 0} \frac{\Delta F_z}{\Delta A}\end{aligned}\tag{2.2}$$

where the subscript notation z specifies the orientation of the area ΔA , and x and y indicate the axes along which each shear stress acts.

2.1.2.3 General State of Stress

If the body is further separated by planes parallel to the x - z plane, as seen in Figure 2.1.2 (b), and the plane y - z , as seen in Figure 2.1.2 (c), one can then extract a cubic volume of material which reflects the state of stress that acts in the body at a selected point. Then, three components on each side of the element, as depicted in Figure 2.1.3 (b), define this state of stress (HIBBELER, 2018).

In the International System of Units (SI), the stress is computed or measured in Pascal (Pa).

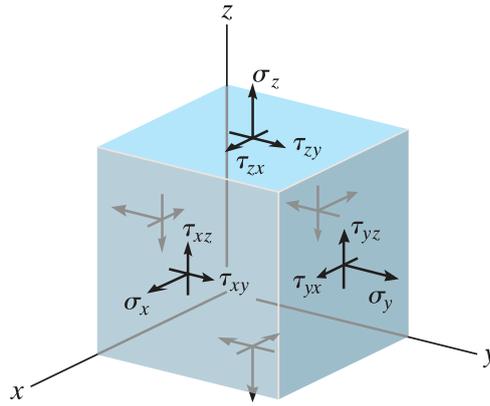


Figure 2.1.3 – General state of stress (HIBBELER, 2018).

2.1.3 Principal Stresses

The normal and shear components of stress acting on an infinitesimal element make up the terms of a tensor, as described by Equation 2.3:

$$\begin{bmatrix} \sigma_{xx} & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \sigma_{yy} & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \sigma_{zz} \end{bmatrix} \quad (2.3)$$

where the notation for each stress component contains three elements, a magnitude (either σ or τ), the direction of a normal to the reference surface (first subscript), and a direction of action (second subscript) (NORTON, 2011).

There will be a continuous distribution of the stress field at any point evaluated for any combination of applied stresses (NORTON, 2011). Any coordinate system chosen will cause the normal and shear stresses at the point to change with direction. Shear-stress components will still be zero on certain planes. Principal stresses are the normal stresses acting on these planes. The principal planes are the planes on which these major stresses act. The main axes are the directions of the surface normals to the principal planes, and the principal normal stresses are the normal stresses acting in those directions (NORTON, 2011). There will also be a second set of mutually perpendicular axes where shear stresses will be greatest. The principal shear stresses work on a set of planes that are at a 45° angle to the principal normal stresses planes (NORTON, 2011).

The expression relating the applied stresses to the principal stresses is described by Equation 2.4:

$$\begin{bmatrix} \sigma_x - \sigma & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \sigma_y - \sigma & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \sigma_z - \sigma \end{bmatrix} = \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (2.4)$$

where σ is the principal stress magnitude and n_x , n_y , and n_z are the direction cosines of the unit vector \mathbf{n} , which is normal to the principal plane. The determinant of the coefficient

matrix must be zero for the solution of Equation 2.4 to exist. Expanding this determinant and set it to zero, Equation 2.5 is obtained:

$$\sigma^3 - C_2\sigma^2 - C_1\sigma - C_0 = 0 \quad (2.5)$$

where

$$C_2 = \sigma_x + \sigma_y + \sigma_z,$$

$$C_1 = \tau_{xy}^2 + \tau_{yz}^2 + \tau_{zx}^2 - \sigma_x\sigma_y - \sigma_y\sigma_z - \sigma_z\sigma_x \text{ and}$$

$$C_0 = \sigma_x\sigma_y\sigma_z + 2\tau_{xy}\tau_{yz}\tau_{zx} - \sigma_x\tau_{yz}^2 - \sigma_y\tau_{zx}^2 - \sigma_z\tau_{xy}^2.$$

The three roots of this cubic polynomial are the three principal (normal) stresses σ_1 , σ_2 , and σ_3 . This polynomial's roots are always real and are normally arranged in the order $\sigma_1 > \sigma_2 > \sigma_3$. By substituting each root of Equation 2.5 into Equation 2.4 and solving for n_x , n_y , and n_z for each of the three principal stresses, the directions of the principal stress vectors can be determined. The three principal stresses have mutually orthogonal directions.

The values of the principal normal stresses can be used to calculate the principal shear stresses, as expressed by Equation 2.6.

$$\begin{aligned} \tau_{13} &= \frac{|\sigma_1 - \sigma_3|}{2} \\ \tau_{21} &= \frac{|\sigma_2 - \sigma_1|}{2} \\ \tau_{32} &= \frac{|\sigma_3 - \sigma_2|}{2} \end{aligned} \quad (2.6)$$

The planes of the principal shear stresses are 45° from the planes of the principal regular stresses, and they are also mutually orthogonal.

2.1.4 Strain

When a force is applied to a body, the body's form and size can be changed. These modifications are known as deformation and can be extremely apparent or almost unnoticeable. Body shape can also occur as the body's temperature is altered. The concept of strain is developed around the deformation of a body by changes in its dimensions (HIBBELER, 2018).

2.1.4.1 Normal Strain

The normal strain is defined by the change of body's length related to the its initial length, due to to an applied load. Figure 2.1.4 depicts the deformation of body in response to an axial load.

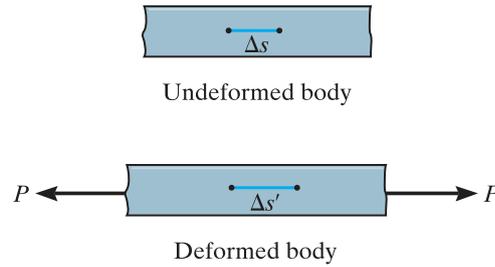


Figure 2.1.4 – Deformation of a body due to an axial load (HIBBELER, 2018).

If the initial length is getting smaller and smaller and approaches to zero, the normal strain in an arbitrary body's point is expressed by Equation 2.7:

$$\epsilon = \lim_{\Delta s \rightarrow 0} \frac{\Delta s' - \Delta s}{\Delta s} \quad (2.7)$$

where ϵ is the normal strain, $\Delta s'$ is the length after deformation and Δs is the initial length.

2.1.4.2 Shear Strain

Besides the elongation or contraction deformations described by normal strains, there are also angular deformations described by the shear strains (HIBBELER, 2018). Shear strain is defined as the change in angular orientation related to the perpendicular referential, i.e., $\frac{\pi}{2}$ due to an applied tangent load in the body's surface, as depicted in Figure 2.1.5.

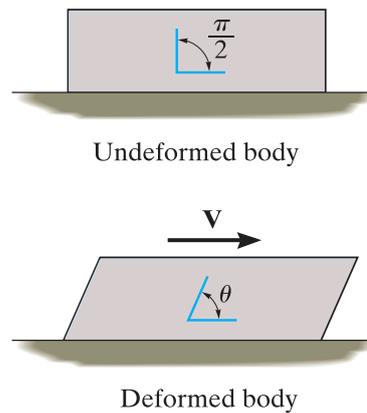


Figure 2.1.5 – Deformation of a body due to a tangential load (HIBBELER, 2018).

If the added load allows the body to deform as seen, the shear strain at the point can be described by Equation 2.8:

$$\gamma = \frac{\pi}{2} - \theta \quad (2.8)$$

where ϵ is the normal strain, $\Delta s'$ is the length after deformation and Δs is the initial length. where γ is the shear strain and θ is the angular change. It is worth noting that the strain is a dimensionless quantity.

2.1.5 Mechanical Properties of Materials

The strength of a material is determined by its ability to withstand a load without deforming or failing. This resilience is built into the material and must be determined by trial and error (HIBBELER, 2018). The strain or stress test is one of the most critical experiments to perform in this respect. The mechanical properties of materials, e.g., metals, ceramics, polymers, and composites, are characterized by the relationship between average normal stress and average normal strain after performing this test (HIBBELER, 2018).

Since the stress and strain data from the test are known, the effects can be drawn in a curve called the stress-strain diagram, as illustrated in Figure 2.1.6. This graph refers to a material specimen of any dimension (HIBBELER, 2018). Stress-strain diagrams vary a lot, and different tensile tests on the same material will produce different results depending on the temperature of the specimen and the loading speed (BEER et al., 2020). On the other hand, it can be used to differentiate between ductile and brittle materials by identifying certain common characteristics. The ability to subject to significant plastic deformation under tensile stress, before rupture, characterizes ductile materials, i.e., structural steel and several alloys of other materials. The length of the specimen increases linearly with the load and at a very slow rate as it is subjected to a rising elongation load.

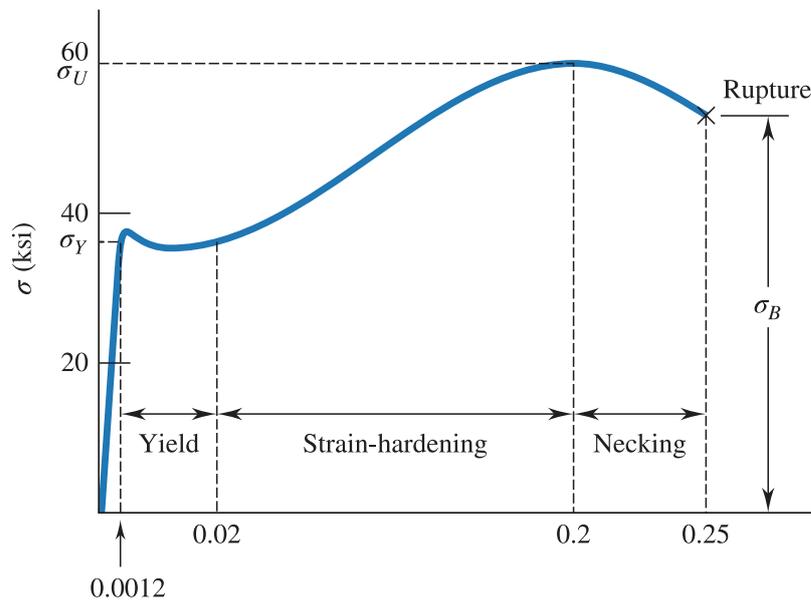


Figure 2.1.6 – Stress-strain diagram of low-carbon steel (BEER et al., 2020).

As a result, the stress-strain diagram's initial section is a straight line that corresponds to the material's linear response up to elastic limit. The specimen, on the other hand, undergoes a major deformation with a comparatively slight rise in the applied load until a critical value, i.e., tensile yield strength σ_Y or S_y , of the stress is achieved. Slippage over oblique surfaces causes this deformation due mainly to shearing stresses. Owing to

local instability, after a full load has been achieved, the diameter of a part of the specimen tends to decrease. Necking is the term for this condition (BEER et al., 2020). Lower loads are necessary for the specimen to elongate further after necking has began, before it eventually ruptures. This means that shear is the primary cause of ductile material loss. The elongation of a ductile specimen, after it has ruptured, can be 200 times greater than its deformation at yield, as seen in Figure 2.1.6. The ultimate strength is defined as the stress σ_U or S_{ut} corresponding to the full load applied. The rupture stress, σ_B , corresponds to the breaking strength (BEER et al., 2020).

2.2 Hydrostatic Pressure Distribution

Pressure in a fluid at rest does not change in the horizontal direction (A.; CIMBALA, 2020). This can be demonstrated by considering a thin horizontal layer of fluid and doing a force balance in every horizontal direction. In a gravity field, however, this is not the case in the vertical direction. Since more fluid rests on deeper layers, pressure rises with depth, and the impact of this extra weight on a deeper layer is compensated by a rise in pressure. This situation is illustrated in Figure 2.2.1.

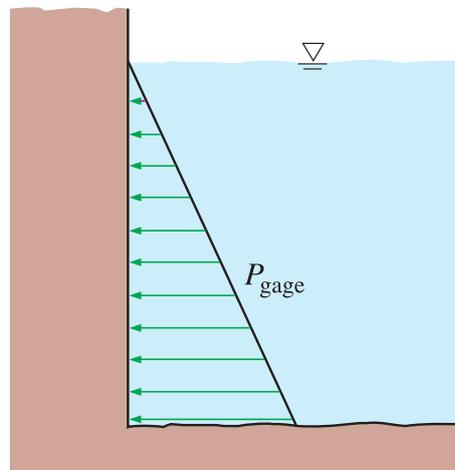


Figure 2.2.1 – Pressure profile of a fluid at rest (A.; CIMBALA, 2020).

Considering the density of fluid ρ to be constant, the pressure change with height is described in Equation 2.9:

$$\Delta P = \Delta P_2 - \Delta P_1 = -\rho g \Delta z = -\gamma_s \Delta z \quad (2.9)$$

where γ_s is the specific weight of the fluid. As a result, the pressure difference between two points of a constant density fluid is equal to the vertical distance Δz between the points and the fluid density. It is worth noting that the negative symbol denotes that pressure in a static fluid increases linearly with depth.

One can observe from Equation 2.9, that the difference of pressure with height for gases is negligible for small to moderate distances, due to their low density. Since the weight of the gas is too small to make a significant difference, the pressure in a tank containing a gas, for example, should be considered uniform. In addition, the pressure in an room filled space can be approximated by a constant (A.; CIMBALA, 2020). Since liquids are practically incompressible, the difference in density with depth is negligible. When the height difference is not very large, this is also true for gases. The difference in density of liquids or gases with temperature, on the other hand, can be substantial and should be taken into account where high precision is needed. Also, because of the enormous volume of liquid weight above, the difference in density of a liquid at great depths can be significant (A.; CIMBALA, 2020).

A relationship of pressure variation with elevation can be obtained by dividing the Equation 2.9 by Δz for liquids whose density varies dramatically at the elevation, and by taking the limit as $\Delta z \rightarrow 0$. This results in Equation 2.10:

$$\frac{dP}{dz} = -\rho g \quad (2.10)$$

where dP is negative when dz is positive since pressure decreases in an upward direction. The pressure difference between any two points 1 and 2 can be computed by the integration when the change of density with elevation is known, as expressed by Equation 2.11:

$$\Delta P = P_2 - P_1 = - \int_1^2 \rho g dz \quad (2.11)$$

As predicted, this relation reduces to Equation 2.9 for constant density and constant gravitational acceleration. At rest, the pressure in a fluid is regardless of the shape or cross section of the container. It varies with vertical distance but is constant in all other directions (A.; CIMBALA, 2020).

2.3 Theories of Failure

Research in failures of parts has occupied scientists and engineers for centuries (NORTON, 2011). Whenever the stresses exceeds its strength, the part tends to fail. The failure may be caused by three types of stresses, i.e., tensile, compressive and shear. To determine which stress could be the failure agent, it is necessary to comprehend the material in question and its relative strengths in compression, tension, and shear. Furthermore, the loading nature, i.e., static or dynamic, and the presence or absence of cracks in the material are relevant to be known.

Generally, ductile, isotropic materials are limited by their shears strength in static tensile loading whereas brittle materials are restricted by their tensile strengths (though there are exceptions to this rule when ductile materials can behave as if brittle) (NORTON,

2011). This means that, for the two types of materials ductile and brittle, there are different failure theories.

Above all, the failure term must be precisely described. If a component yields and distorts to the point that it no longer functions properly, it can fail (NORTON, 2011). A component can also fail by fracturing and separating. Either of these conditions are failures, but the processes, i.e., failures mechanisms, that cause them can be very different (NORTON, 2011). Only ductile materials may have the ability to yield significantly before fracturing. Brittle materials may fracture without significantly changing their form.

The nature of the loading, whether static or dynamic, is another important factor in failure. Static loads are added gradually and effectively remain stable over time. Dynamic loads are either added unexpectedly (impact loads) or vary over time (fatigue loads), or both. The difference between ductile and brittle materials' failure behavior blurs under dynamic loading, and ductile materials fail in a "brittle" manner (NORTON, 2011). Due to the major variations in failure mechanisms under static and dynamic loading, they must be treated separately.

2.3.1 Static Failure Theory

While ductile materials can fracture if statically stressed above their ultimate tensile strength, equipment parts fail when they yield under static loading (NORTON, 2011). A ductile material's yield strength is significantly less than its ultimate strength.

2.3.2 Distortion-Energy Theory

The relative slipping of the material's atoms within their lattice arrangement can be understood as the microscopic yielding process (NORTON, 2011). Shear stress causes this slipping, which is followed by a distortion of the part's form. The amount of energy accumulated in the component as a result of the distortion indicates the extent of the shear stress (NORTON, 2011).

When materials are hydrostatically loaded to generate uniform stresses in all directions, massive quantities of strain energy can be stored without failure occurrence (NORTON, 2011). Several studies have demonstrated that materials can be hydrostatically stressed to levels above their ultimate strengths in compression without failing because the specimen's volume is reduced without altering its shape (NORTON, 2011). Thus, the distortion tends to be the major responsible for tensile failure.

The total strain energy in a loaded element can be divided into two categories: hydrostatic loading, which changes the volume, and distortion, which changes the form (NORTON, 2011). The distortion-energy element will give us a measure of the shear stress present if we separate the two components.

When there are mixed tensile and shear stresses acting on the same point, it is convenient to describe an effective stress that can be used to reflect the stresses combination. For ductile materials, the distortion-energy solution is a safe way to go. The von Mises effective stress is described as a uniaxial tensile stress that produces the same amount of distortion energy as the actual combination of applied stresses (NORTON, 2011). This method allows handle cases of combined, multiaxial tension and shear stresses as if they were due to pure tensile loading. The effective stress of von Mises for the three-dimensional case is described by Equation 2.12:

$$\bar{\sigma} = \sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2 - \sigma_1\sigma_2 - \sigma_2\sigma_3 - \sigma_1\sigma_3} \quad (2.12)$$

Through the failure criterion considered for the distortion-energy theory, the yield strength S_y is the failure stress of interest (NORTON, 2011). It is convenient to use a safety factor N , as expressed by Equation 2.13, for design purposes, so that the stress condition remains in safe levels.

$$N = \frac{S_y}{\bar{\sigma}} \quad (2.13)$$

The smallest value of N for any component is of greatest concern since it estimates the most likely mode of failure. As N is reduced to 1, the stress in the component equals the material's strength, i.e., the added pressure equals the load that fails it, and failure happens. As a result, it is desirable N to be always greater than 1 (NORTON, 2011).

2.3.3 Dynamic Failure Theory

Machine components, vehicles and buildings are often subjected to repeated loads, and the resultant cyclical stresses can cause damages to the material concerned (DOWLING; PRASAD; NARAYANASAMY, 2013). Also at stresses below a given material's ultimate strength, this microscopic damage can accumulate over time and build into a crack or other macroscopic damage that contributes to component failure. Fatigue is the term used to describe the mechanism of damage and degradation caused by cyclic loading. The concept was coined because it seemed to early researchers that cyclic stresses induced an incremental, but not immediately apparent, transition in the material's ability to withstand stress.

Some practical applications, and also many fatigue tests on materials, involve cycling between constant maximum and minimum stress levels. This is known as continuous amplitude stressing, as seen in Figure 2.3.1 (a), (b), and (c).

The stress range, $\Delta\sigma$, is defined as the difference between the maximum and minimum values, σ_{max} and σ_{min} , respectively (DOWLING; PRASAD; NARAYANASAMY, 2013). The mean stress, σ_m , is calculated by averaging the maximum and minimum values,

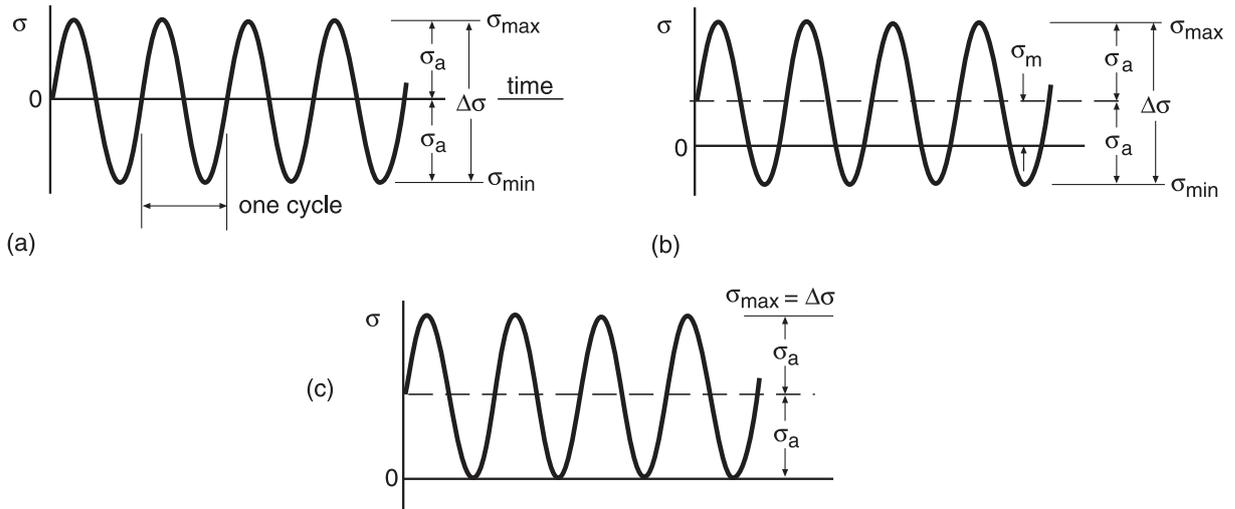


Figure 2.3.1 – Constant amplitude cycling and the associated nomenclature: (a) completely reversed stressing, (b) nonzero mean stress, and (c) zero-to-tension stressing (DOWLING; PRASAD; NARAYANASAMY, 2013).

as expressed by Equation 2.14:

$$\sigma_m = \frac{\sigma_{max} + \sigma_{min}}{2} \quad (2.14)$$

The mean stress can be zero, as seen in Figure 2.3.1 (a), but it is often not, as seen in Figure 2.3.1 (b) (DOWLING; PRASAD; NARAYANASAMY, 2013). The stress amplitude, σ_a , which is the difference across the mean, is half of the range, as expressed by Equation 2.15:

$$\sigma_a = \frac{\Delta\sigma}{2} = \frac{\sigma_{max} - \sigma_{min}}{2} \quad (2.15)$$

The stress amplitude has the same meaning of the alternating stress term. Since $\sigma_{max} > \sigma_{min}$, where tension is considered to be positive, the signs of σ_a and $\Delta\sigma$ are always positive.

2.3.4 Stress Versus Life (S-N) Curve

If there is a sufficiently extreme cyclic stress on a test specimen of a substance or component engineering, a fatigue crack or other harm develops, leading to a complete failure of the member. The number of cycles to failure would be lower if the test is repeated at a higher stress frequency (DOWLING; PRASAD; NARAYANASAMY, 2013).

The outcomes of these measurements from a variety of stress levels can be drawn to a stress-life curve, also known as the S-N diagram, as shown in Figure 2.3.2. The fatigue strength, S_f or σ_F , is the ordinate of the S-N diagram, and it must always be followed by a description of the number of cycles, N_f , to which it is referred.

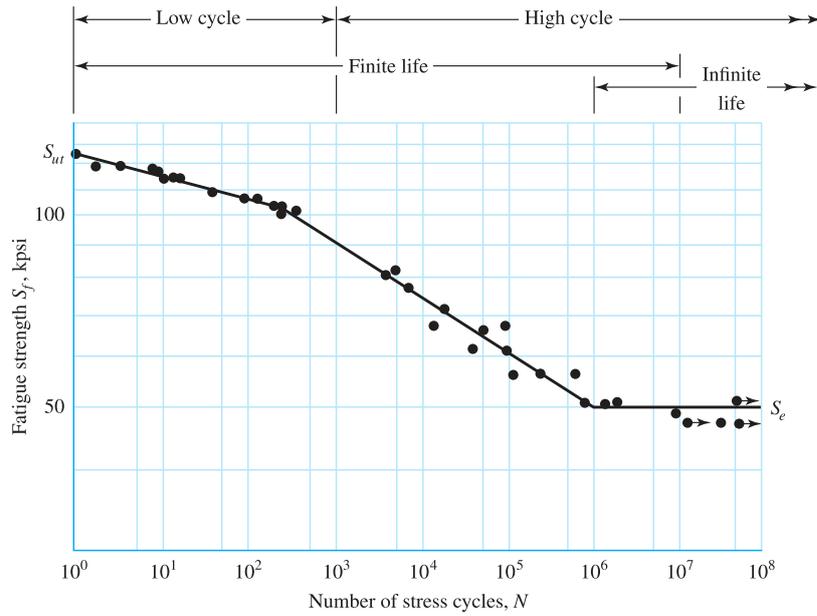


Figure 2.3.2 – Stress versus life (S-N) diagram plotted from the results of completely reversed axial fatigue tests. Material: UNS G41300 steel. (BUDYNAS; NISBETT; SHIGLEY, 2016).

If it is shown that the S-N data approximates to a straight line in a log-linear plot, a mathematical description of the curve can be given by Equation 2.16:

$$\sigma_a = \sigma'_f (2N_f)^b \quad (2.16)$$

where σ'_f is the fatigue strength coefficient and b is known as the fatigue strength exponent or Basquin exponent (SURESH, 1991).

In the case of steels, a knee appears in the graph, and beyond this knee failure will not occur, no matter how great the number of cycles (BUDYNAS; NISBETT; SHIGLEY, 2016). The strength corresponding to the knee is called the endurance limit S_e , or the fatigue limit (BUDYNAS; NISBETT; SHIGLEY, 2016). The S-N diagram is commonly obtained by completely reversed stress cycles, in which the stress level alternates between equal magnitudes of tension and compression (BUDYNAS; NISBETT; SHIGLEY, 2016).

The term fatigue strength refers to the value of a stress amplitude from an S-N curve at a specific life of interest. Hence, the fatigue strength at 10^5 cycles is simply the stress magnitude associated with $N_f = 10^5$ (DOWLING; PRASAD; NARAYANASAMY, 2013). High-cycle fatigue and low-cycle fatigue are two other concepts associated with S-N curves. The former describes conditions with a long fatigue life in which the stress level is low enough that the yielding effects do not control the behavior. The life at where high-cycle fatigue begins varies depending on the material, but it is usually in the range between 10^2 to 10^4 cycles (DOWLING; PRASAD; NARAYANASAMY, 2013).

2.3.5 Life Estimates with Mean Stress

Since the S-N curve for completely reversed loading is known and tests at $\sigma_m = 0$ are employed to obtain the constants σ'_f and b , the stress amplitude σ_a corresponds to the special case denoted as completely reversed stress, σ_{ar} (DOWLING; PRASAD; NARAYANASAMY, 2013). This is directly expressed by Equation 2.14. Considering that the relation between σ_{ar} and σ_a is defined by Equation 2.17:

$$\sigma_{ar} = \frac{\sigma_a}{1 - \frac{\sigma_m}{\sigma'_f}} \quad (2.17)$$

then it can be combined with Equation 2.14 in order to yield a more general stress-life equation that applies for nonzero mean stress, as expressed by Equation 2.18:

$$\sigma_a = (\sigma'_f - \sigma_m)(2N_f)^b \quad (2.18)$$

2.3.6 One Approach to Multiaxial Fatigue

In engineering components, cyclic loadings that cause complex states of stress are common. Examples include biaxial stresses due to cyclic strain in tubes or pipes, combined bending and torsion of the shafts and sheets or plates about more than a single axis (DOWLING; PRASAD; NARAYANASAMY, 2013). Steady applied loads that cause mean stresses may also be combined with such cyclic loads (DOWLING; PRASAD; NARAYANASAMY, 2013). An additional complexity is that cyclical load sources may differentiate between phases, frequencies or both (DOWLING; PRASAD; NARAYANASAMY, 2013). In a thin-walled tube under cyclic pressure, for example, if a constant bending stress is applied, there will be several stress amplitudes and mean stresses in two directions (DOWLING; PRASAD; NARAYANASAMY, 2013). The axial and hoop directions are the directions of principal stress and remain so as the pressure fluctuates.

The approach to multiaxial fatigue takes into account the simple situation where all cyclic loads are reversed and have the same frequency and are either in phase or 180° out of phase (DOWLING; PRASAD; NARAYANASAMY, 2013). Additionally, assume that no steady (non-cyclical) load is present at this time. In this case, it is reasonable to suppose that fatigue life is dominated by the cyclic amplitude of the octahedral shear stress, for ductile engineering metals (DOWLING; PRASAD; NARAYANASAMY, 2013). In order to calculate the effective stress amplitude, the amplitudes of the main stresses σ_{1a} , σ_{2a} , and σ_{3a} can then be employed to compute an effective stress amplitude using a relationship similar to that employed for the octahedral shear yield criterion, as expressed by Equation 2.19:

$$\bar{\sigma}_a = \frac{1}{\sqrt{2}} \sqrt{(\sigma_{1a} - \sigma_{2a})^2 + (\sigma_{2a} - \sigma_{3a})^2 + (\sigma_{3a} - \sigma_{1a})^2} \quad (2.19)$$

The life can then be calculated by entering an S-N curve for fully reversed uniaxial stress with the aid of $\bar{\sigma}_a$. It's worth noting that the most typical S-N curves come from bending or axial tests, which do entail a uniaxial state of stress and can hence be directly used with $\bar{\sigma}_a$ values (DOWLING; PRASAD; NARAYANASAMY, 2013). If steady (noncyclic) loads are present, they change the effective stress amplitude $\bar{\sigma}_a$ in a similar way to how the mean stress effect under uniaxial loading (DOWLING; PRASAD; NARAYANASAMY, 2013). One approach is to assume that the governing mean stress variable is proportional to the hydrostatic stress' steady value (DOWLING; PRASAD; NARAYANASAMY, 2013). On this basis, an efficient mean stress can be determined from the mean stresses in the three principal directions, as expressed by Equation 2.20:

$$\bar{\sigma}_m = \sigma_{1m} + \sigma_{2m} + \sigma_{3m} \quad (2.20)$$

The quantities $\bar{\sigma}_a$ and $\bar{\sigma}_m$ can be combined into an equivalent completely reversed uniaxial stress by generalizing Equation 2.15 (DOWLING; PRASAD; NARAYANASAMY, 2013).

2.4 Clustering

Clustering is one of the most widely used unsupervised machine learning technique for exploratory data analysis (HENNIG et al., 2016). Intuitively, clustering is the task of grouping a set of objects so that similar ones are contained in the same set, whereas different ones are dissociated in distinct sets (HENNIG et al., 2016). Formally, a set of data, or dataset, $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$, that contains n samples, the task of clustering consists in group those samples in " k " disjoint subsets of X , represented by C_1, C_2, \dots, C_K . For instance, a clustering is the partition obtained from $C = \{C_1, C_2, \dots, C_K\}$. If a data sample \mathbf{x}_i belongs to the cluster C_K , then the label of \mathbf{x}_i is k . Thus, the labels vary from 1 to k , that is the number of clusters. Certainly, not every C is considered as reasonable or useful clustering, and what can determine this classification depends on the particular approach in which occurs the clustering analysis. It is relevant to consider that clustering or data grouping can have different meanings because it depends on the context e analysis area where it is applied. There are not a unique concept related to what cluster is, or what should be the most reasonable grouping mode for a determined abstract dataset X . Hence, the fundamentals of any rigorous clustering analysis is a clear and suitable definition of what would be an appropriate grouping in a specific context (HENNIG et al., 2016).

2.4.1 Centroid-based Clustering

Supposing that there is a fixed number of known clusters k , the idea behind the method k-means is finding k centroid objects in order to represent them in a optimized way,

being this one of the most traditional approaches in clustering. From this perspective, the clustering task is understood as an optimization problem. The objective function has the pair of inputs (X, d) , and the proposed solution for the clustering is $C = \{C_1, C_2, \dots, C_K\}$. Given a specific objective function, determined by G , the algorithm's goal consists in finding, for a determined input (X, d) , a clustering C that minimizes $G((X, d), C)$. To reach this goal, it is necessary to apply the appropriated search algorithm. Several objective functions require the number of clusters, k , as a parameter. In practice, it is common for the user to be in charge of deciding the most convenient k for the clustering problem (HENNIG et al., 2016).

In k-means method, data is partitioned in disjoint sets C_1, C_2, \dots, C_K , where each C_i is represented by one centroid μ_i . It is assumed that the input set X is joined in a larger metric space (X', d) , such that $X \subseteq X'$, considering that the centroids belongs to X' . The objective function measures the square of distances between each point in X and the centroid of its cluster (SHALEV-SHWARTZ, 2012). The centroid related to C_i is defined by Equation 2.21:

$$\mu_i(C_i) = \underset{\mu \in X'}{\operatorname{argmin}} \sum_{x \in C_i} d(x, \mu)^2 \quad (2.21)$$

Thus, the objective function can be described by Equation 2.22:

$$G_{K\text{-means}}((X, d), (C_1, C_2, \dots, C_k)) = \sum_{i=1}^k \sum_{x \in C_i} d(x, \mu(C_i))^2 \quad (2.22)$$

Sometimes, the process to find an optimal solution is impracticable. As an alternative, the Algorithm 1 in Appendix A, that presents a iterative process, is commonly used (HENNIG et al., 2016). In this case, the clustering solution is returned by a numerical method and does not necessarily represents the result that minimizes the objective function. In Algorithm 1 the distance function $d(x_i, c_j)$ is the euclidean distance $\|x_i - c_j\|$.

2.4.2 Hierarchical Agglomerative-based Clustering

The hierarchical agglomerative clustering begins with a single object that composes its own cluster. thus there are n clusters for n objects initially. At each step, the most similar pair of clusters are joined, resulting in a new clustering setting with a reduced number of clusters. This process is repeated until the top level all objects are joined into a single cluster. Similar agglomerative hierarchical approaches have different means of computing the dissimilarity D for two clusters from the member object dissimilarities d (HENNIG et al., 2016). The most known method is *Single Leakage* or *Nearest Neighbor*, as expressed in Equation 2.23:

$$D(C_1, C_2) = \min_{\mathbf{x}_1 \in C_1, \mathbf{x}_2 \in C_2} d(\mathbf{x}_1, \mathbf{x}_2) \quad (2.23)$$

It is worth noting that this method is focused on dissimilarity data and is not linked to probability models. By cutting the hierarchy at the required stage, it is possible to obtain partitions into k clusters from hierarchies (HENNIG et al., 2016). The hierarchical agglomerative clustering algorithm is described by Algorithm 2 in Appendix A (HENNIG et al., 2016).

2.4.3 Spectral-based Clustering

The spectral clustering deals with similarity or dissimilarity data, by means of a $n \times n$ square similarity matrix. However, this method is commonly used with other data formats, after having obtained similarities from the original data. For example, a matrix of objects and variables can be transformed in a square matrix of similarities between object pairs. The advantage in using this method for data belonging to the euclidean space is that the spectral clustering allows clusters have arbitrary formats (HENNIG et al., 2016).

The spectral clustering term is related to the family of techniques based on laplacian matrix decomposition, that is derived from the similarity matrix. The k main eigenvectors of laplacian matrix provide a mapping of objects in k dimensions. In order to obtain the clusters, the resultant k dimensional vectors are grouped by standard methods, such as k -means. The spectral clustering algorithm is described by Algorithm 3 in Appendix A (HENNIG et al., 2016).

2.4.4 Density-based Clustering

Density-based clustering is a non parametric method in which clusters are called high-density regions of the density. Density-based clustering approaches don't include the number of clusters as an input parameter, and they don't make any predictions about the underlying density or the variation between the clusters that might occur in the data set. As a result, density-based clusters aren't always groups of points with low pairwise within-cluster dissimilarity as determined by the dissimilarity function, and therefore aren't always convex in form, but can be randomly formed in the data space. A density-based cluster is a collection of data objects distributed around a contiguous area of high object density in the data space (KRIEGEL et al., 2011).

An intuitive description of a cluster based on probability models is to presume a non-parametric density model for the whole dataset and classify clusters either with high-density regions or with density modes, assigning each observation to the density mode that "attracts" it. Such approaches are much more flexible in terms of cluster shapes than parametric model mixtures (although such flexibility is not always desirable, particularly in applications in which large within-cluster distances should be avoided) (HENNIG et al., 2016).

Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is a density-based clustering algorithm. It can detect clusters of various shapes and sizes in a large volume of data that is noisy and contains outliers (KRIEGEL et al., 2011). DBSCAN's model employs a straightforward minimum density level estimate based on a threshold for the number of neighbors, ν , within the radius, δ , with an arbitrary distance measure. In other words, ν is also known as the minimum number of points in a dense region, while δ can also be interpreted as the maximum distance between neighbors in a single cluster. A core point is an object that has more than ν neighbors within this radius (including the query point). DBSCAN intuition is to select regions that meet this minimum density and are separated by areas of a lower density. DBSCAN does not perform density computation between points for efficiency purposes. Instead, all neighbors within a core point's radius are considered members of the same cluster as the core point (called direct density reachable). If any of these neighbors are a core point again, their neighborhoods are transitively included (density reachable). Border points are non-core points in this group, and all points inside the same set are density related. Points that can't be reached by density from any core point are called noise and don't belong to any cluster (SCHUBERT et al., 2017). The DBSCAN clustering algorithm is described by Algorithm 4 in Appendix A (SCHUBERT et al., 2017).

2.5 Final Remarks

In this chapter, mechanics of materials foundations as well as clustering analysis were presented. This explanations are relevant in order to comprehend the methodology described in the following chapters. The centroid, agglomerative, spectral and density based clustering algorithms will be analyzed as interface level identification techniques. After, the operational scenarios simulation and the computation of structural life indicators under dynamic and static loadings will require the comprehension about principal stresses, hydrostatics and theory of failures.

3 Density Profiling, Clustering Methods and a Computer Vision based Profiler

This chapter will describe the methodology employed by the proposed DT architecture model. Here, will be presented the validation and test of a density profiling approach through non-linear regression as well as a non-parametric interface level identification technique through clustering analysis. Additionally, a computer vision profiler system will be demonstrated in order to estimate the density profile of the experimental separator vessel.

3.1 Density Profiling Through Non-Linear Regression Modeling

Similarly to linear regression, the fundamental concept behind non-linear regression is to link a variable, often called response or dependent variable, to another variable, usually called explanatory, regressor, or independent variable (SMYTH, 2006; SEBER; WILD, 2003). This independent variable is used to predict or explain the behavior of the dependent one. The assumption that the prediction equation depends non-linearly on one or more uncertain parameters characterizes nonlinear regression (SMYTH, 2006). While linear regression is regularly utilized for building a pure experimental model, non-linear regression normally emerges when there are physical reasons behind accepting that the connection between the dependent and independent variables follows a specific functional structure (SMYTH, 2006).

Sometimes, the resolution of data read by sensors is not enough to feed mathematical or digital models in DTs. Thus, instead of increasing the number of sensors, the resolution problem can be solved by interpolating the gaps between each data point. Besides the linear approach, it is possible to adapt more sophisticated non-linear regression models to predict unknown variables where sensors can not reach in the physical space. In this work's case study, the structural model of the experimental separator vessel should be fed by a load pressure profile with a spatial resolution in the the order of millimeters. Though, to physically reach such resolution is very costly or even not realizable. Analytical regression models comes as an alternative in order to predict measurements of high resolution sensing systems. Therefore, this work presents a novel approach about modeling separator vessel density profile data.

3.1.1 Crude Oil Density Profile Experimental Data

Separator vessels present a density distribution along its height when filled with crude oil mixture. Figure 3.1.1 shows that the density profile can be measured by an instrument called density profiler. The density profiler consists in a set of density sensors distributed along a rod, by which density measurements are acquired at several height levels.

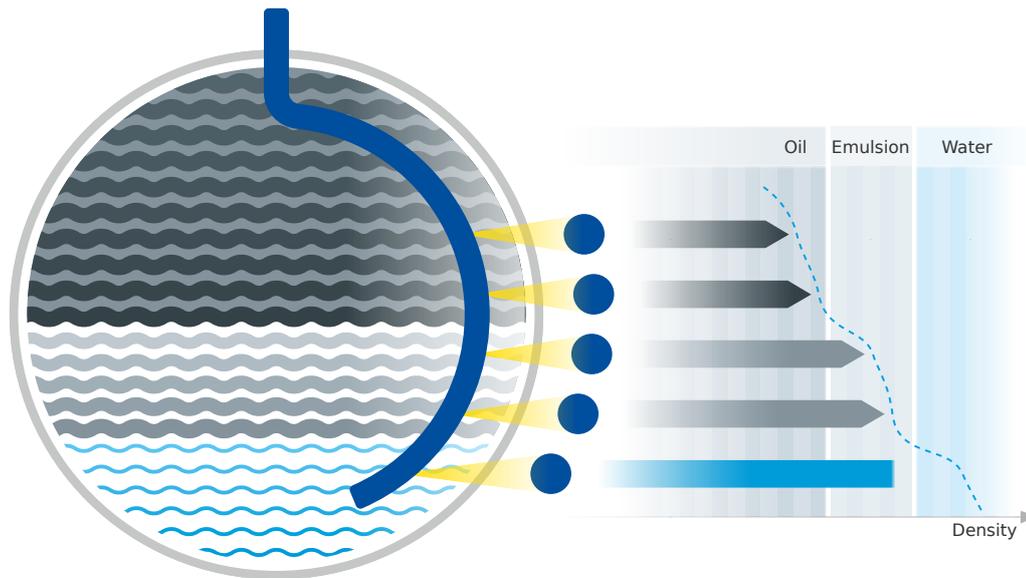


Figure 3.1.1 – Multiphase level measurement system and its resulting density profile (BERTHOLD, 2021).

Through a non-linear regression method, as introduced above, it is intended to build an analytical model of the density profile measured from a real separator vessel, as depicted in Figure 1.1.2. For this, density profile data provided by the *Petróleo Brasileiro S.A.* company (Petrobras) was used (CENPES, 2019). This data set contains density readings measured by a nucleonic density profiler with 96 data points. This instrument comprises a vertical array of small, gamma ray emitting radioactive sources from which radiation is monitored by a vertical array of detectors. The density profile readings will be presented horizontally for consistency with the methodology of this dissertation. This configuration will come in handy for the nonlinear regression curve fitting that will be discussed later in this section. Figure 3.1.2 (a), (b), (c), and (d) show four density profile samples collected from the data set. Each sample was measured in different instants from the production separator vessel.

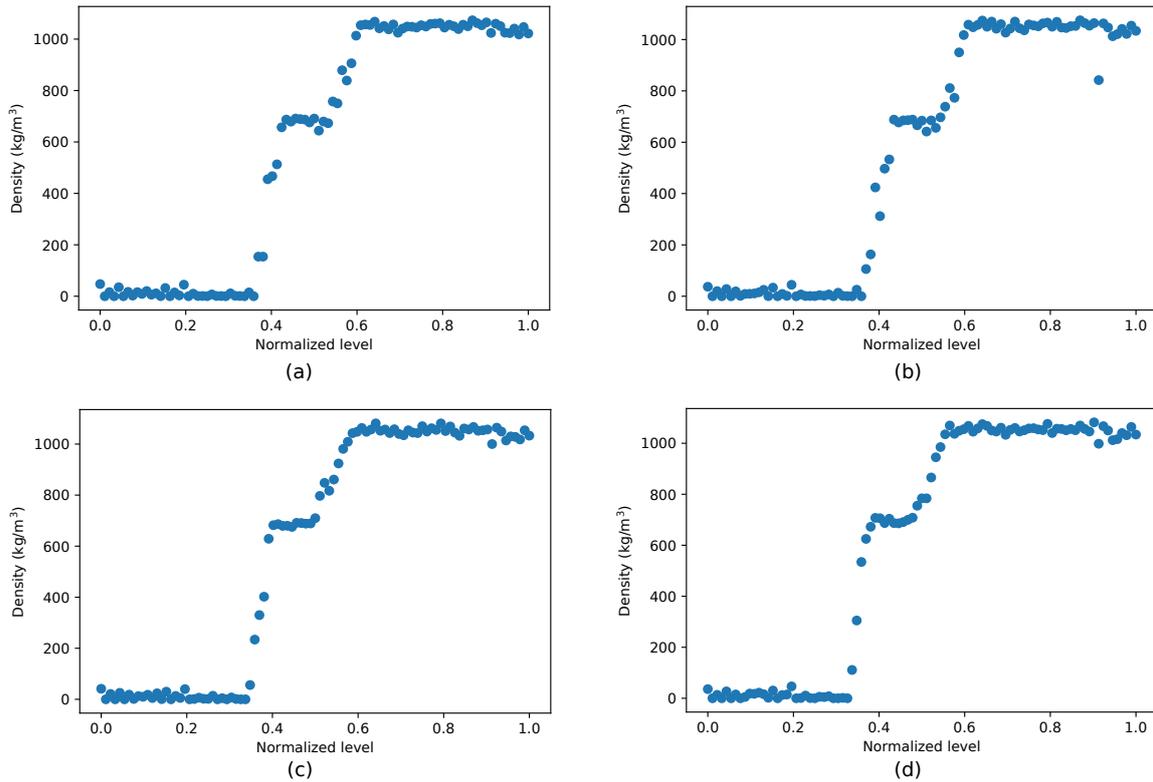


Figure 3.1.2 – Density profile samples measured from the production separator vessel: (a) sample #1, (b) sample #2, (c) sample #3, and (d) sample #4.

It is worth mentioning that these density profiles belong to a multi-phase mixture with gas, foam, oil, emulsion and water layers. Also, the horizontal axis contains a normalized level range, since the real level data was not provided with the density profile dataset. In this scale, *zero* corresponds to the maximum measurable level, while *one* corresponds to the minimum measurable level of the separator vessel. Thus, this configuration of normalized level scale is adopted as a standard approach for the density profile horizontal axis in this dissertation.

3.1.1.1 Crude Oil Emulsion

Emulsion can be characterized as a system, whereby one immiscible fluid phase is scattered as globules (dispersed stage) in the second phase of immiscible fluid (continuous stage) (LIM et al., 2015). There are three kinds of emulsions: water-in-oil (W/O) emulsion, oil-in-water (O/W) emulsion and multiple emulsion. W/O emulsion is formed when the water globules are dispersed throughout the oil continuous phase. O/W emulsion is formed when the oil globules are dispersed all through the water continuous phase. Multiple emulsion is a complex emulsion system, whereby W/O or O/W emulsions are dispersed throughout another immiscible phase (BARKAT et al., 2011). Multiple emul-

sion includes water-in-oil-in-water (W/O/W) emulsion and oil-in-water-in-oil (O/W/O) emulsion (WONG; LIM; DOL, 2015). The O/W and W/O emulsion types can be easily found in petroleum industry (SCHRAMM, 1992).

In oil production industries, although crude oil and water are initially present in two separated phases, the turbulence, the mixing, as well as agitation through downhole wellbore, surface chokes, valves, pumps and pipes will develop emulsion formation (FINGAS; FIELDHOUSE, 2003). Emulsions which can be particularly troublesome in the operation of separators vessels. Over a period of time an accumulation of emulsified materials and/or impurities may form at the interface of the water and oil phases. This accumulation will not only have a adverse effect on liquid level control, but it will also reduce the efficient oil or water retention time in the separator, lowering the water–oil separation quality (STEWART; ARNOLD, 2011). Chemicals and/or heat are often used to alleviate this issue.

3.1.1.2 Crude Oil Foam

The appearance of impurities other than water, which are impossible to extract until the stream meets the separator, is the main source of foam in crude oil. CO₂ is one impurity that nearly always results in foam. Foam can also be caused by completion and workover fluids that are incompatible with the wellbore fluids (STEWART; ARNOLD, 2011). Foaming issues can reduce crude oil production and even trigger unplanned process shutdowns, depending on the quality of the crude oil and the form of separation scheme used (POINDEXTER et al., 2002). Foam isn't an issue in a separator if the internal architecture allows enough time or sufficient coalescing surface for the foam to break (STEWART; ARNOLD, 2011).

Due to the increased residence time for foam breaking, crude oil foaming has a negative effect on separator capacity. Another drawback being the inaccuracy of level measurement (CHEN et al., 2018). Foaming can allow liquid to leak into the gas line, contaminating solvents and causing harm to scrubbers and compressors (POINDEXTER et al., 2002). Foaming may also trigger unwanted gas carry-over in the oil line, resulting in dangerous degassing and pump harm. Foam inhibition provides better control of the crude oil separation process (CHEN et al., 2018).

3.1.2 Regression Modeling Hypothesis

Observing the regions of foam and emulsion with its adjacent single-phase layers, in Figure 3.1.2 (a), (b), (c), and (d), one can see that they are similar to a step function. However, the step function does not represents the density changes accurately, since their transitions are not abrupt. Thus, a better alternative would be modeling such changes by s-shaped or sigmoid curves. Specifically, sigmoid or s-shaped functions are monotonic

increasing and bounded on their definition domain, assuming values between zero and one (ILIEV; KYURKCHIEV; MARKOV, 2017). The growth pattern of such functions could represent gradual changes in density in function of level. The s-shaped curve can be interpreted as two portions bounded by its mid point where occurs the reversion of each continuous phases, for the emulsion, or the predominance of gas or liquid, for the foam layer. This analogy is demonstrated by Figure 3.1.3.

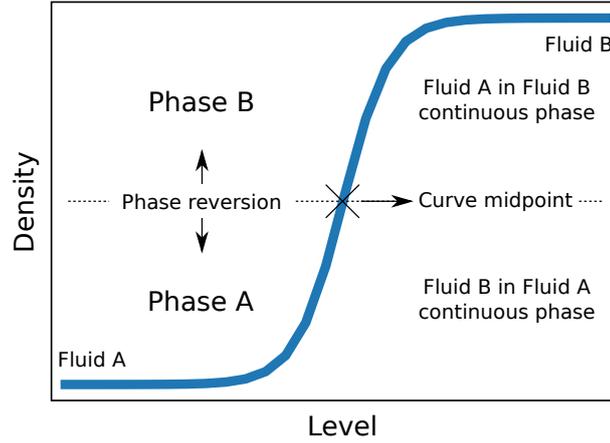


Figure 3.1.3 – Principle of interface modeling through s-shaped curve.

In order to validate this hypothesis, non-linear regression models were built for each interface layer using the main functions of sigmoid curves family: logistic, hyperbolic tangent, inverse tangent, and algebraic form. The logistic function is expressed by Equation 3.1:

$$f(x) = \frac{a}{1 + e^{-b(x+c)}} + d \quad (3.1)$$

where a is the curve's maximum value or scale magnitude, b is the growth rate or steepness of the curve, and (c, d) pair is the coordinate of the sigmoid's midpoint related to the curve's translational state.

The hyperbolic tangent function is expressed by Equation 3.2:

$$f(x) = a \left(\frac{e^{b(x+c)} - e^{-b(x+c)}}{e^{b(x+c)} + e^{-b(x+c)}} + d \right) \quad (3.2)$$

where a is the curve's maximum value or scale magnitude, b is the growth rate or steepness of the curve, and (c, d) pair is the coordinate of the sigmoid's midpoint related to the curve's translational state.

The inverse tangent function is expressed by Equation 3.3:

$$f(x) = a \arctan(b(x+c)) + d \quad (3.3)$$

where a is the curve's maximum value or scale magnitude, b is the growth rate or steepness of the curve, and (c, d) pair is the coordinate of the sigmoid's midpoint related to the curve's translational state.

The sigmoid's algebraic form function is expressed by Equation 3.4:

$$f(x) = a \frac{-b(x+c)}{\sqrt{1 + -b(x+c)^2}} + d \quad (3.4)$$

where a is the curve's maximum value or scale magnitude, b is the growth rate or steepness of the curve, and (c, d) pair is the coordinate of the sigmoid's midpoint related to the curve's translational state. It is worth mentioning that the parameters from each function were estimated by calling curve fitting methods from the Python library Scipy ([VIRTANEN et al., 2020](#)).

3.1.3 Validation of Non-linear Models

The non-linear regression models for the foam and emulsion layers from each density profile sample showed by Figure 3.1.2 (a), (b), (c), and (d) are depicted in Figure 3.1.4. Generally, the eight non-linear regression models returned a average root mean square error (RMSE) of $4.443 \cdot 10^{-2}$ and a mean coefficient of determination (R^2) of 0.987.

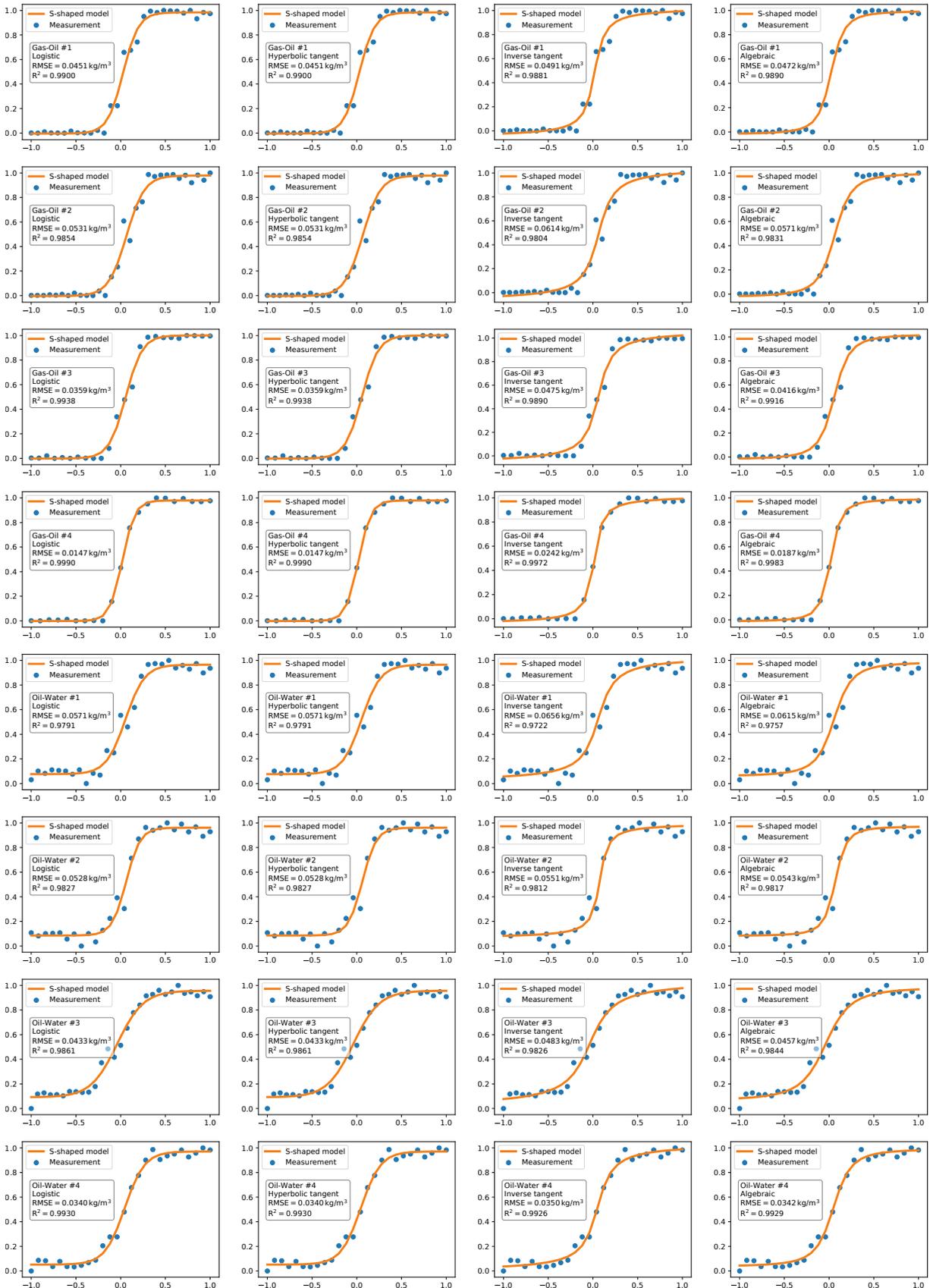


Figure 3.1.4 – Non-linear regression models subplots. The ordinate axis is normalized density and the abscissa axis is normalized level.

Observing Figure 3.1.5, it is evident that the logistic and hyperbolic tangent curve fittings presented the same and the lowest average RMSE values, $4.206 \cdot 10^{-2} \text{ kg/m}^3$, since they are very similar s-shaped curves. On the other hand, the inverse tangent curve fitting presented the highest average RMSE, $4.846 \cdot 10^{-2} \text{ kg/m}^3$. The algebraic form curve fitting presented an intermediary average RMSE of $4.515 \cdot 10^{-2} \text{ kg/m}^3$. It is worth noting that the RMSE evaluation also provides insight into residual analysis, since the RMSE can be regarded as the residuals standard deviation.

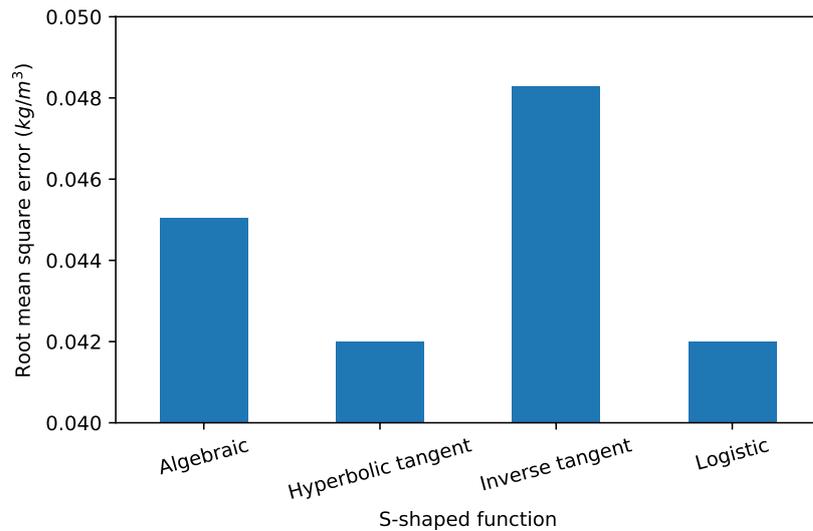


Figure 3.1.5 – Root mean square errors for each s-shaped curve fitting.

Regarding the R^2 , as depicted in Figure 3.1.6, the logistic and hyperbolic tangent curve fittings presented the same and the highest average values, ~ 0.989 . Still, the inverse tangent curve fitting presented the lowest average R^2 , ~ 0.985 . Again, the algebraic form curve fitting presented an intermediary average R^2 of ~ 0.987 . It is evident from Figures 3.1.5 and 3.1.6 the negative correlation between RMSE and R^2 values.

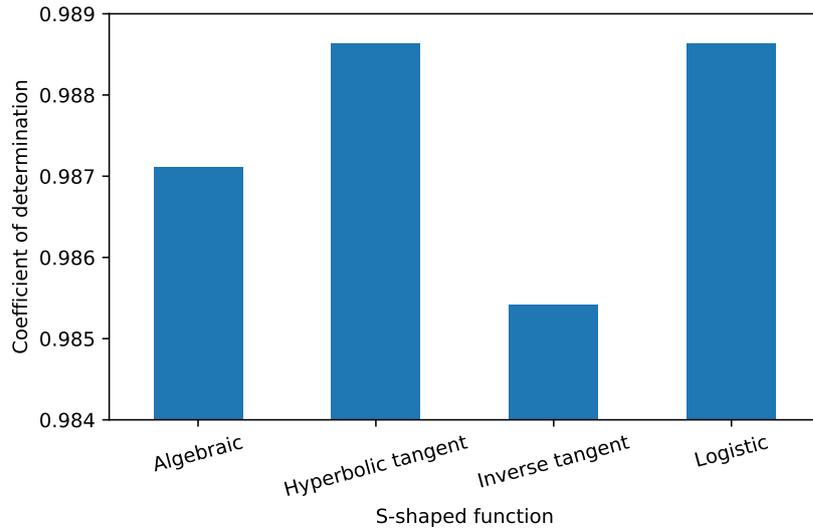


Figure 3.1.6 – Coefficients of determination for each s-shaped curve fitting.

The results showed that the non-linear regression models through s-shaped curve fitting is a promising method for analytically modeling the foam and emulsion layers by using density profile data. This achievement will be relevant for experimental separator vessel DT load modeling in Section 4.

3.1.4 Proposed Density Profile Model

In a three-phase density profile, there are three single phase levels and two interface layers to be modeled. This indicates that two non-linear regression models has to be concatenated in order to represent an entire three-phase density profile. Thus, the resultant model can be represented by the piecewise equation expressed in Equation 3.5:

$$\rho(z) = \begin{cases} S_{gas/oil} & z < z_c \\ S_{oil/water} & z \geq z_c \end{cases} \quad (3.5)$$

where $\rho(z)$ is the density profile function, z is the level height, $S_{gas/oil}$ is the s-shaped fitted curve related to the foam layer, $S_{oil/water}$ is the s-shaped fitted curve related to the emulsion layer, and z_c is the level where the concatenation is done, generally the oil layer intermediate level. The proposed density profile modeling workflow is depicted in Figure 3.1.7. Initially, the density profile measurement has to be split into the foam and emulsion layer chunks. Then, each chunk is normalized in order to guarantee regression algorithm convergence. Then, the obtained regression models have to be denormalized and concatenated in order to map them to the actual density levels and join them into a single density profile.

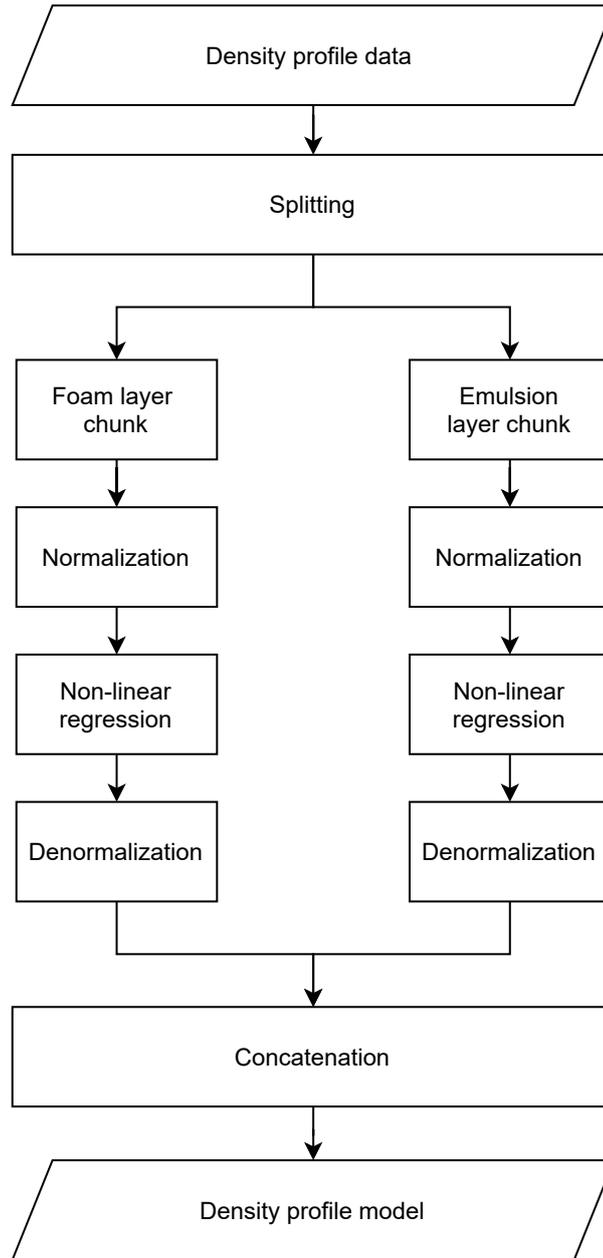


Figure 3.1.7 – Proposed density profile modeling workflow.

Figure 3.1.8 depicts the proposed workflow output model for the density profile measurement data showed in Figure 3.1.2 (a).

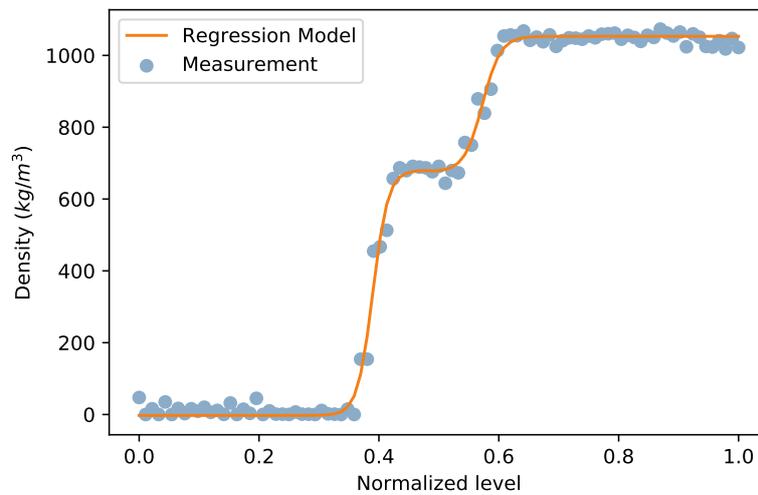


Figure 3.1.8 – Resultant model for the entire density profile data.

3.2 Interface Level Identification Through Density Profile Clustering

The level monitoring instruments are an essential part in hydrocarbon processing industrial plants, and they have been widely applied in the last decades with separator vessels. The level instrument's precision and reliability as well as the capability to monitor every liquid interface layer, including foam and emulsion lengths, are relevant attributes since the level monitoring is considered to be the main sensing element of separators vessels automatic process control (HJERTAKER; JOHANSEN; JACKSON, 2001a). Liquid level measurement has an important role in several industrial applications such as chemical processes, fuel storage and transport, oil reservoirs, residual treatment plants, etc. Multiple sensing techniques based on acoustic, mechanic, electric and optic principles are employed to measure level parameter (Díaz et al., 2018). It is worth considering that electric sensors are broadly used in level measurement, however they have limitations regarding safety requirements in hazardous areas (SENGUPTA; KISHORE, 2014).

Oil & Gas industries have directed efforts in development e optimization of process control in order to enhance production quality and to minimize chemical agents use in separation processes, since they can increase costs and environmental harm (HJERTAKER; JOHANSEN; JACKSON, 2001a). Machine learning comes as a promising and complementary methodology to the traditional monitoring techniques so that the sensor data is processed and interpreted more effectively. It is comprehended as one of the multidisciplinary fields of data science and can be widely defined as computational methods that use experience to enhance process performance by running learning algorithms (MOHRI;

ROSTAMIZADEH; TALWALKAR, 2018). Experience refers to past information available to the learning system, that consists in digital data collected and made available for analysis. This data may be a labeled digital information set or may be other types of information obtained by environment interaction. Generally, the quality and size of data are crucial for the learning system prediction success.

Since gas, oil and water are immiscible liquids with different densities, the separation process is achieved by confining the multi-phase liquid in storage tanks until gas, oil and water layers are formed at the top, middle and bottom of the recipient respectively (Guirong Lu et al., 2009). However, there may be the formation of a interface layer between oil and water called emulsion (Al-Naamany; Meribout; Al Busaidi, 2007), and also a interface layer between the gas and oil called foam due to the carbonation process (SHI et al., 1991). The multi-phase mixture of hydrocarbons may also include sand and other solids present in petroleum wells forming a clay layer at bottom of separator vessels (HJERTAKER; JOHANSEN; JACKSON, 2001a).

Recently, the diffusion and consolidation of modern measurement instruments, that are commonly connected to communication networks, big volumes of data are generated from sensors applied to several industrial processes (KHAN et al., 2020; Qiu et al., 2020). In order to extract useful information from such sets of data, machine learning algorithms can be used for refine measurement readings and to provide support in decision-making or even to the control system optimization in production processes.

Oil and water separation is one of the key processes in oil & gas industry (LEAL-JUNIOR et al., 2018). The hydrocarbons contained in production wells consists in a multi-phase mixture of oil, gas, water and solids (HJERTAKER; JOHANSEN; JACKSON, 2001b). Generally, these heterogeneous components are separated by identifying the density differences between each immiscible fluids (HJERTAKER; JOHANSEN; JACKSON, 2001b). Nonetheless, there are not laminar and well defined interfaces between each phase. Contrarily, there are emulsion layer between oil and water, in which its composition and behavior are affected by the properties of the both liquids and by the process dynamics (LANGEVIN et al., 2004). Furthermore, there may be foam formation between gas and oil layers (YANG; BRANT; BECK, 1994). These interface layers have to be monitored and restrained because they can reduce the efficiency of separation process.

Therefore, level monitoring in separator vessels is necessary to optimize the performance of the equipment in order to avoid accidental discharges of oil in water, providing cost reduction, collaborating with the environment preservation by dispensing chemical agents in separation process. Also, it is important to control the amount of sand in the equipment so that this material can be removed before problems can occur in the process (Holstad et al., 2005).

In this section, it is presented a new approach to identify interfaces levels in separator

vessels multi-phase liquids through a machine learning technique called clustering. For this study, four categories of clustering are considered (k-means, hierarchical agglomerative, spectral and density based). This techniques are applied to real process data in order to test their viability and to compare each one's performance in terms of accuracy and algorithm execution time.

3.2.1 Clustering Density Profiles

Through the machine learning techniques presented above, it is intended to automatically identify the layers across a density profile of multi-phase liquids in a real separation process. For this, density profile data provided by the Petr leo Brasileiro S.A. company (Petrobras) was used (CENPES, 2019). Figure 3.2.1 (a) shows a sample of the density profile in a separator vessel.

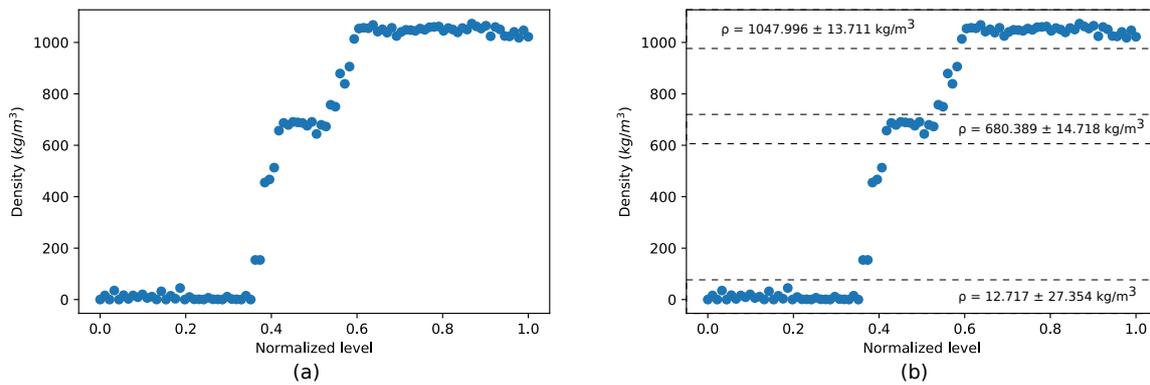


Figure 3.2.1 – Density profile data sample: (a) raw data, and (b) hard-coded grouped data.

Since clustering is an unsupervised learning technique, training the model from labeled data is unnecessary. Additionally, model evaluation by cross-validation is not suitable for this approach, once an unique model needs to be generated for each density profile sample. Thus, in order to evaluate the clustering methods performance, it was necessary to manually group each phase in the data set previously by a hard-coded algorithm. This labeled grouped will be relevant to compare and then compute accuracy of clustering results lately. The clustering algorithms were executed in the Google Colab, a cloud platform for running Jupyter notebooks. Figure 3.2.1 (b) shows the density intervals that groups each phase present in the heterogeneous liquid.

The labeled density profile data for each clustering method are depicted in Figure 3.2.2 (a), (b), (c) and (d), where the resultant clusters are identified with a color code. It is worth noting that the k parameter, i.e., the number of clusters, was considered to be

5 for the parametric methods (k-means, agglomerative hierarchical, and spectral), since there are gas, foam, oil, emulsion and water layers.

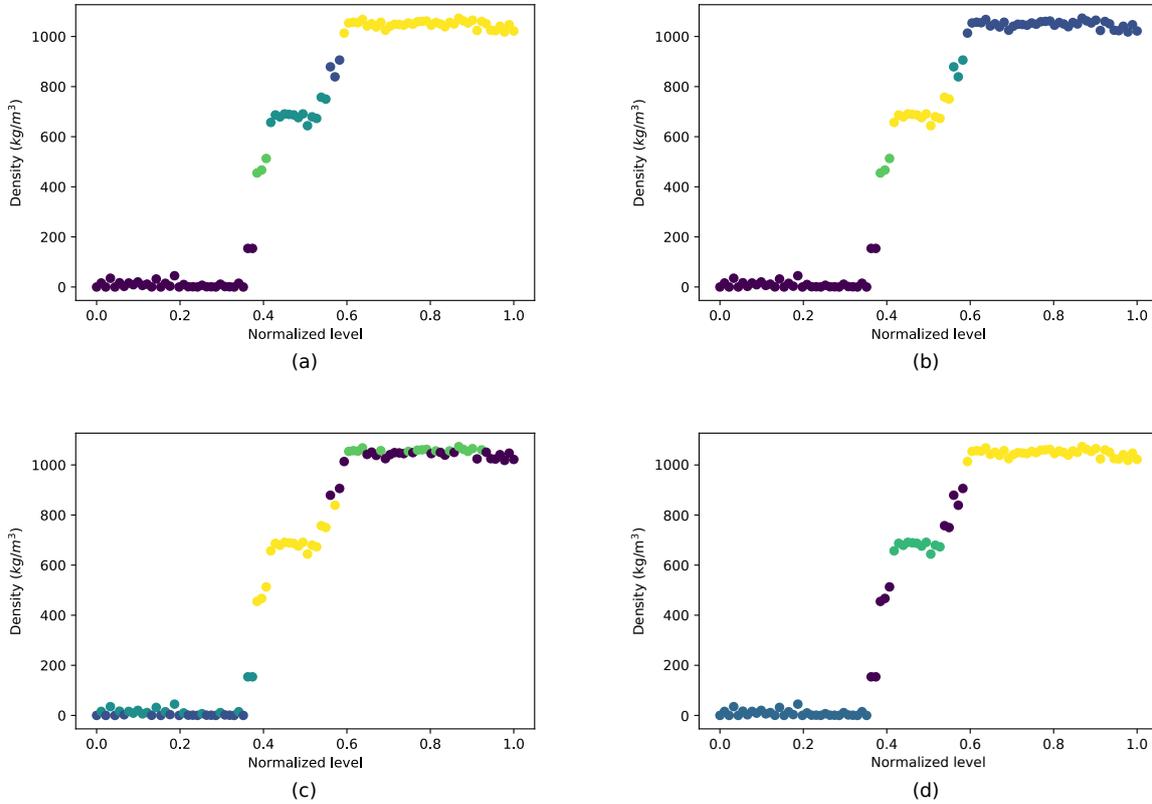


Figure 3.2.2 – Clustering results: (a) K-means, (b) Hierarchical Agglomerative, (c) Spectral, (d) DBSCAN.

Figure 3.2.2 (a) and (b) show that the k-means and hierarchical agglomerative clustering methods have similar results, both giving an accuracy of 95.652%. On the other hand, spectral clustering was not capable of properly segmenting the data sequence, as depicted in Figure 3.2.2 (c), giving an accuracy of 53.261%. Finally, DBSCAN was the only clustering method that matched with the pre-established grouping, giving an accuracy of 100%. It is worth considering that the foam and emulsions layers are interpreted as noise by DBSCAN algorithm, whereas the gas, oil and water layers were accurately identified as distinct clusters, as depicted in Figure 3.2.2 (d).

In terms of execution time, the hierarchical hierarchical clustering presented the best performance, being executed in $(1.136 \pm 0.176) \cdot 10^{-3}$ s. Then, in ascending order of elapsed time, DBSCAN, hard-coded algorithm, spectral, and k-means clustering were executed in $(1.568 \pm 0.348) \cdot 10^{-3}$ s, $(3.187 \pm 0.644) \cdot 10^{-3}$ s, $(1.177 \pm 0.214) \cdot 10^{-2}$ s, $(2.178 \pm 0.217) \cdot 10^{-2}$ s, respectively. Figure 3.2.3 (a) and (b) shows the performance test results for accuracy and execution time, respectively. This evaluation was realized in

Google Colabatory cloud computing environment with default settings.

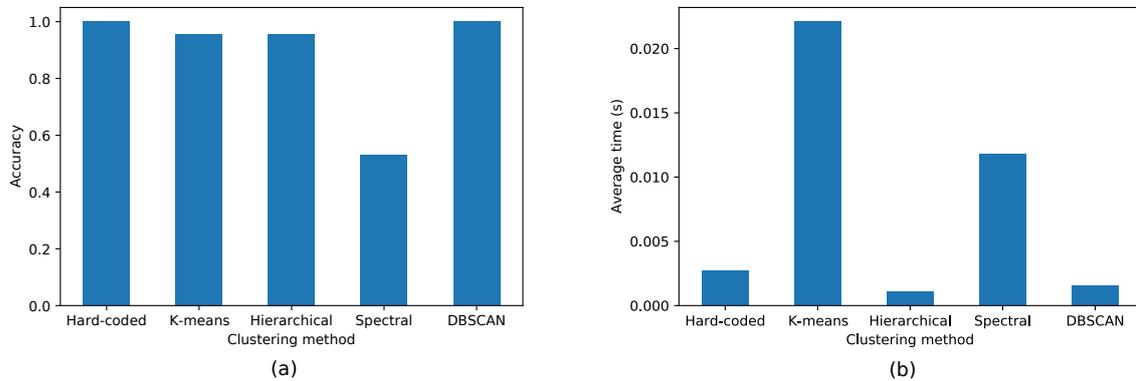


Figure 3.2.3 – Clustering performance test results: (a) accuracy, and (b) execution time.

It is worth mentioning that, besides DBSCAN is a non-parametric clustering method, its algorithm has two main arguments, δ and ν , not linked to the number of clusters, in which its values have to be optimized in order to improve the clustering process. Figure 3.2.4 shows the DBSCAN clustering accuracy values for different values of δ and ν . In this performance analysis, it was considered $\delta = 34$ and $\nu = 5$, since these parameters combination resulted in 100% of accuracy.

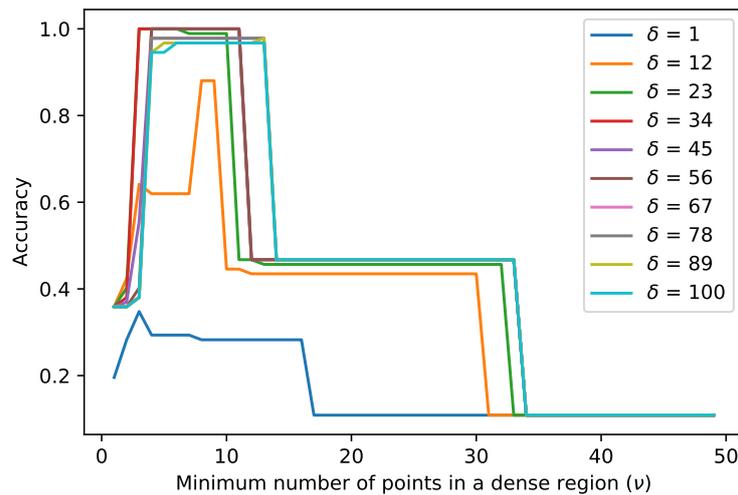


Figure 3.2.4 – DBSCAN accuracy values in function of δ and ν .

The results showed that clustering can be a useful technique for a non conventional grouping approach by segmenting density profile data sequences in order to automatically identify interface levels. The executed tests showed that the DBSCAN and hierarchical agglomerative clustering are the most accurate and the fastest methods. Still, the advantage

of DBSCAN over the others is that it is a non-parametric method, so the number of clusters k does not need to be known in advance. Therefore, the conventional hard-coded approach can be adequately replaced by the DBSCAN or hierarchical agglomerative methods for automatically density clustering and interface levels identification purposes, since the interface layers, i.e., foam and emulsion, presents high variability and are irregularly distributed along the level. Also, the clustering methods are not vulnerable to pressure and temperature variations, because the actual densities of each liquid phase do not need to be known before grouping, unlike the conventional hard-coded approach.

3.3 Non-contact Density Profiling Through Computer Vision

The moisture content of crude oil is an important parameter in the petroleum exploration and petrochemical industries, and it has a lot of significance in terms of crude oil refining, storage, transportation, and pricing (LIU *et al.*, 2019). There are several studies about measurement of water in crude oil mixtures using different methodologies, such as microwave technology (MAKEYEV; LIFANOV; SOVLOUKOV, 2009; MAKEEV; LIFANOV; SOVLUKOV, 2014; SHARMA; LAO; FALCONE, 2018), conductive principle (GARCÍA-GOLDING *et al.*, 1995), ultra-short-wave technology (LI; GUOYI, 2008), spectral absorption method (ZHAO *et al.*, 2004), and capacitive sensing system (ASLAM; TANG, 2014; ASLAM; BOON, 2014). Those methods can reliably quantify water content in crude oil samples in real time, reducing labor pressure in the manufacturing process (LIU *et al.*, 2019). Nonetheless, they have strict requirements for the measurement environment and instruments, such as alkanes and hydrocarbons in crude oil, and the environment's temperature has a negative impact on measurement performance (LIU *et al.*, 2019).

Sensing structural and physical parameters (e.g, density, pressure, temperature, vibration, etc) of separator vessels can provide updated versions of its digital representations in order to run simulations analysis recurrently. Recently, digital image processing has become a prominent and enticing research field (LIU *et al.*, 2019). Artificial intelligence, industrial detection, and other fields have benefited from the technology, as well as the advancement of related disciplines (DELIGIANNIS *et al.*, 2017; DIXIT; NASKAR, 2017). Computer vision techniques have become a trendy method for non-contact liquid level measurement. Throughout the years, there were works in which computer vision techniques have been explored for liquid level detection in closed and transparent recipients through optical imaging processing systems (TAKAGI *et al.*, 1998; Batagelj *et al.*, 2000; Chakravarthy; Sharma; Kasturi, 2002; YU, 2015; WANG *et al.*, 2009; EPPEL, 2016). Horizontal liquid level detection using computer vision is becoming more widespread as a common non-contact liquid level measurement method (LIU *et al.*, 2019). Compared to other contact liquid measurements, this approach can be adequately applied to the liquid level measurement of corrosive and explosive liquids, as well as other special application

scenarios (LIU et al., 2019).

The above-mentioned methods for determining water content are based on the premise that the physical or chemical properties of each product of a crude oil sample are different (LIU et al., 2019). Visualization technology has become a relevant tool for the growth of the oil & gas industry (EVANS et al., 2002). However, image processing-based approaches for calculating crude oil water content are scarcely found in established references (Liu et al., 2017). Hence, the approach proposed in this work employs computer vision to estimate not only the water content in a multi-phase mixture, but also the full density profile of oil and water content stored in a experimental separator vessel.

3.3.1 Proposed Computer Vision Workflow

The proposed algorithm that process the oil and water mixture image is depicted by Figure 3.3.1. It consists in a four-stage chain in which the acquired image data is sequentially pre-processed or filtered, thresholded, and differentiated.

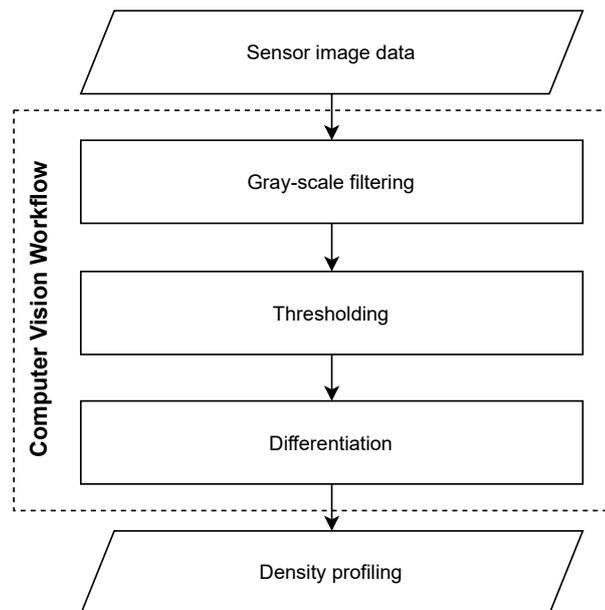


Figure 3.3.1 – Proposed computer vision workflow diagram.

In order to capture the separator’s display image, a Samsung SAK2L4 ISOCELL CMOS image sensor with 12 megapixels of resolution, and a pixel size of 1.40 μm was used. Initially, the separator was filled with Mobil AW68 oil and fresh water. In the captured separator’s display image frame, as shown in Figure 3.3.2 (a), it is observed that the oil-water interface is not well defined due to emulsion formation. In contrast, the foam layer was not observed.

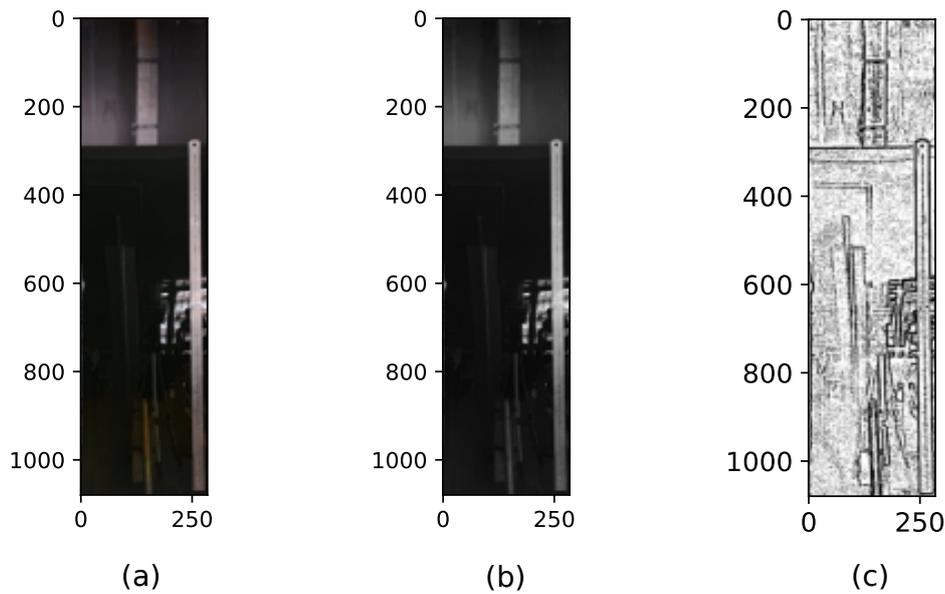


Figure 3.3.2 – Oil and water content within the experimental separator vessel: (a) original image frame, (b) gray-scale image frame, and (c) thresholded gray-scale image frame.

3.3.2 Pre-processing the Image Data

After receiving the image frame from the sensor, the first step of the proposed workflow is convert or filter the image data from RGB color space to the gray-scale color space. Each pixel in a gray-scale image is equivalent to a one-dimensional vector (KUHNS; OLIVEIRA; FERNANDES, 2008), with no chrominance or chroma details. The main advantage in processing gray-scale images is that they have less redundant information and image data to handle than RGB images, and their storage memory requirements are smaller. Figure 3.3.2 (b) shows the image frame in gray-scale color space. It is worth mentioning that the gray-scale can represent the continuous changes in liquid density along the vertical axis, i.e., the height, through a uniform color code interval with values between 0 and 255. Considering a 256-level gray-scale image, the white value, which is 255, is the highest, and the black value, which is 0, is the smallest (GONZALEZ; WOODS, 2018).

3.3.3 Thresholding the Gray-scale Frame

It is evident that the image frames are polluted by ambient reflections as well as light refractions due to the display's reflexive and translucency material. Thus, in order to detect the less polluted areas in the image, the adaptive thresholding method was applied in the gray-scale image frame, as illustrated in Figure 3.3.2 (c). This method was chosen because it deals well with variations of illumination, whereas fixed thresholding

often fails with light variation in a image. In adaptive threshold method, the threshold value is computed for each single pixel in the image, providing increased robustness to illumination changes (BRADLEY; ROTH, 2007). This technique helps to identify by highlighting reflection and refraction contents through a discrete image segmentation algorithm (BRADLEY; ROTH, 2007). Then, the image frames are windowed to the region of interest, i.e., the continuous portion with less segments between 0 and 255, and its respective column array, with the display height, that contains the mean color code values along the horizontal axis are obtained, as showed by Figure 3.3.3. These sequences have data related to the single-phase layers, emulsion layer and also oil-air interface level.

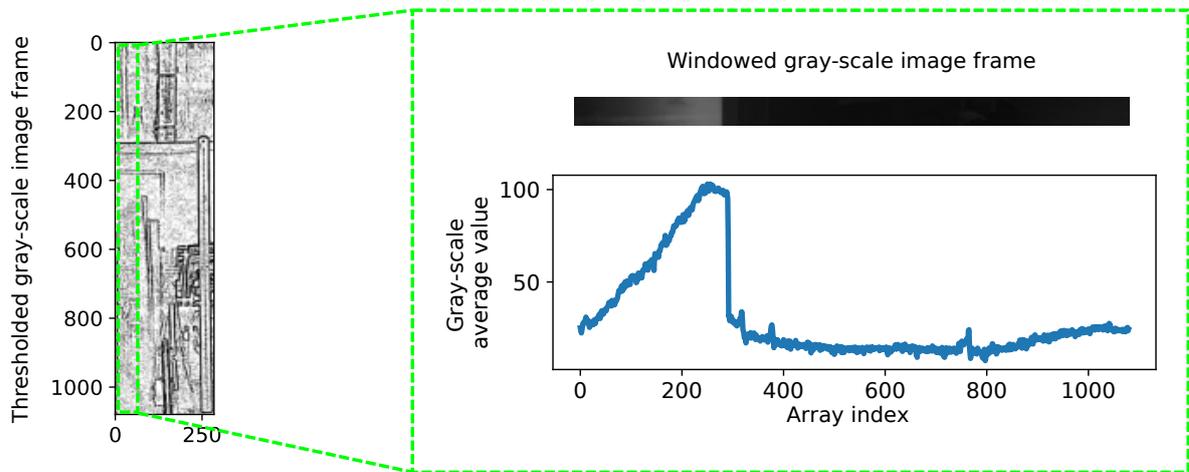


Figure 3.3.3 – Gray-scale image frame windowing, and averaging.

3.3.4 Differentiating the Data Sequence

In order to track the oil-air interface layer position, it is computed the minimum value of the obtained gray-scale average array first derivative, as described in Equation 3.6:

$$\min \left(\frac{dG}{dx} \right) = \min (G[x + 1] - G[x]) \quad (3.6)$$

where G is the average gray-scale value array.

Since the gray-scale value is abruptly reduced from air to oil, the oil-air interface level can be easily highlighted through its derivative, as shown in Figure 3.3.4.

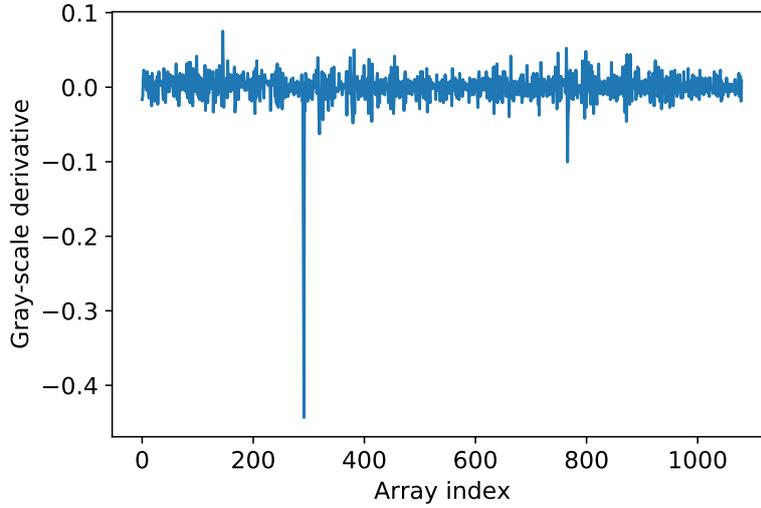


Figure 3.3.4 – Gray-scale average array derivative.

Then, this data array can be linearly mapped to the density profile scale based on the Mobil AW68 oil (880 kg/m^3), fresh water (998.21 kg/m^3) and air (1.204 kg/m^3) densities at NTP as the referential values. Once there is not enough material deformation to be detected by the image sensor, the linear mapping approach could be applied directly, only considering the original equipment dimensions as reference. Since the linear mapping needs known referential densities, this methodology can be considered as an indirect density profile estimate.

3.3.5 Density Profiling

In Section 3.1, the foam and emulsion layers from several density profile samples were modeled by s-shaped curves. Thus, with the gray-scale average value array and its first derivative sequence, one can observe that this data can be analytically expressed by a piecewise function derived from Equation 3.5, as described by Equation 3.7:

$$\rho(z) = \begin{cases} S & z < z_{air} \\ \rho_{air} & z \geq z_{air} \end{cases} \quad (3.7)$$

where $\rho(z)$ is the density profile function, z is the level height, S is the s-shaped fitted curve, ρ_{air} is the air density, and z_{air} is the oil-air interface level. Hence, the emulsion and its adjacent single-phase layers are modeled by the s-shaped curve due to their smooth density transition, whereas the oil-air interface is represented by a step density shift, since the transition is abrupt there and the foam layer is not observed in the profile. The obtained s-shaped regression model returned a RMSE of $7.805 \cdot 10^{-3} \text{ kg/m}^3$ and a R^2 of 0.865. Figure 3.3.5 (a) and (b) show the emulsion layer non-linear regression model and the full density profile analytical model, respectively.

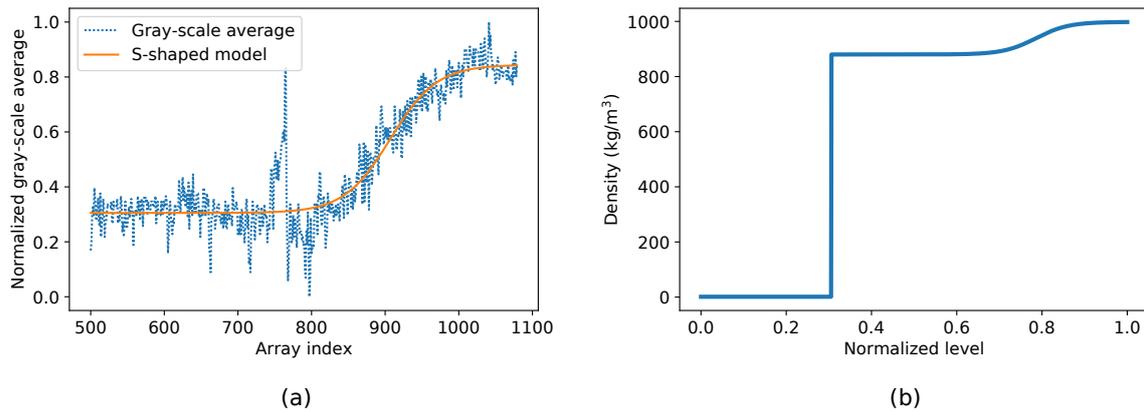


Figure 3.3.5 – Density profiling: (a) emulsion layer non-linear regression model, and (b) full density profile analytical model.

3.4 Final Remarks

Once the density profiling, clustering, and computer vision profiler methods were validated, it is possible to test them through the DT architecture model of the experimental separator vessel presented in the next chapter. Thus, the density profile model, obtained by the computer vision approach, will be converted into the pressure profile and it will feed the separator structural model. From there, the operational scenarios can be simulated and the interface levels for each one can be predicted.

4 Experimental Separator Vessel Digital Twin Architecture Model

This chapter explores the proposed architecture model for a separator prototype DT implementation in order to be a tool for supporting liquid level control systems and equipment structural monitoring. Through this, the maintenance could be improved if structural failures are minimized by predicting the occurrence of failures under static and dynamic loads due to cyclic operations in different working scenarios. Then, through data processing across a pipeline, the equipment observability can be augmented in terms of visibility, predicatability and traceability.

4.1 The Physical Element

The experimental separator vessel, i.e., the physical element, has unique structural characteristics, since it is made by heterogeneous materials. This equipment, as shown in Figure 4.1.1 (a) and (b), is built with carbon steel sheets and has a polycarbonate display on the front side in order to monitor the liquid content within it. On the top, the lid has two flanges for density and pressure profiler instruments installation. On the left side, there are seven pressure outlets. On the bottom, there are two connections for liquid input and output as well as four support legs. Some extractions from the experimental separator vessel's CAD design are presented in Figures A.0.1 and A.0.2 at Annex A (2SOLVE, 2019).

The proposed DT architecture model is based on the "what-if" and "forward-looking" keywords (MAYANI et al., 2018). That means it is intended to build a tool that allow us to continuously pre-establish possible operational scenarios and then run simulations and predictive analysis in order to enhance safety, reduce risks, prevent losses and increase the equipment performance and reliability.

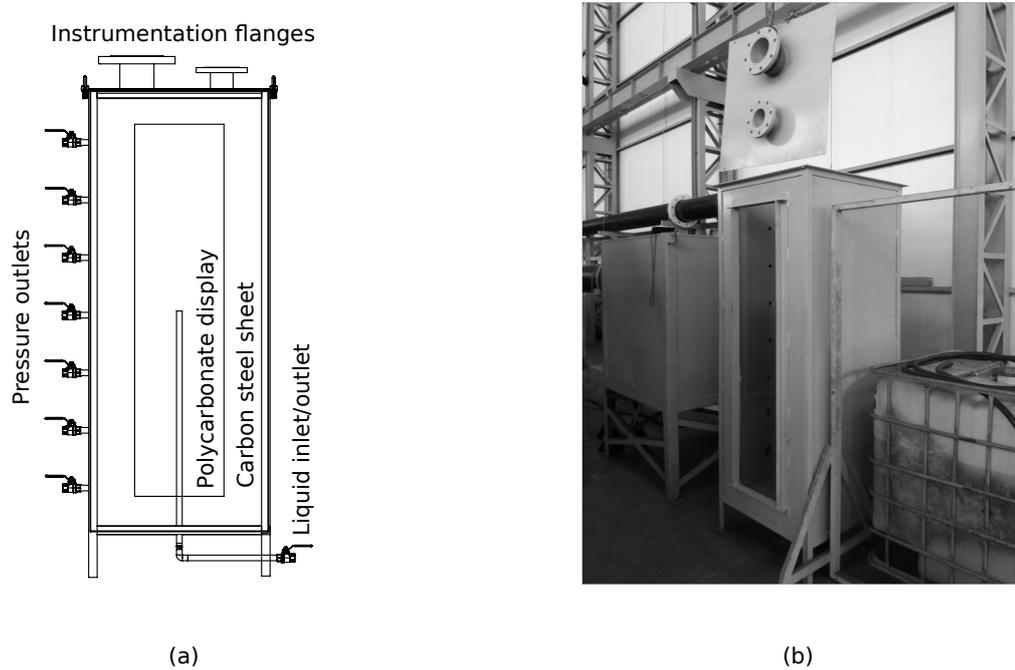


Figure 4.1.1 – Separator prototype used as the physical element of the DT: (a) equipment design (b) equipment photography.

4.2 Proposed Architecture Model

The architecture, shown in Figure 4.2.1, follows a pipeline model in which data flows through a chain of transformation procedures, since density/pressure sensing and data preparation until computational structural simulation and predictive analysis. This pipeline starts with data acquisition through a data source represented by the sensor to measure the current separator's density profile subjected to a physical challenge. The adopted operational scenarios consist in static load conditions and a deterministic cyclic load provided by the variation of oil and water levels inside the separator, whose maximum and minimum are read by the sensor. For this work, we used a image sensor that captures the display's reflected and refracted light in order to obtain the density profile respective to separator's load state. After the image sensing, there are the pre-processing, processing and post-processing stages. In the pre-processing stage, the data collected is used to model the current separator load in which image data is mapped to the density and pressure profiles.

In related works, computer vision approaches has already been employed to detect liquids interface levels (EPPEL; KACHMAN, 2014; LIU et al., 2019), but not yet for process density profiles data. Then, in the processing stage, the modeled and mapped data is incorporated to the separator's virtual entity composed by the geometric, material, structural and rule digital models. When updating the virtual entity with the current pre-

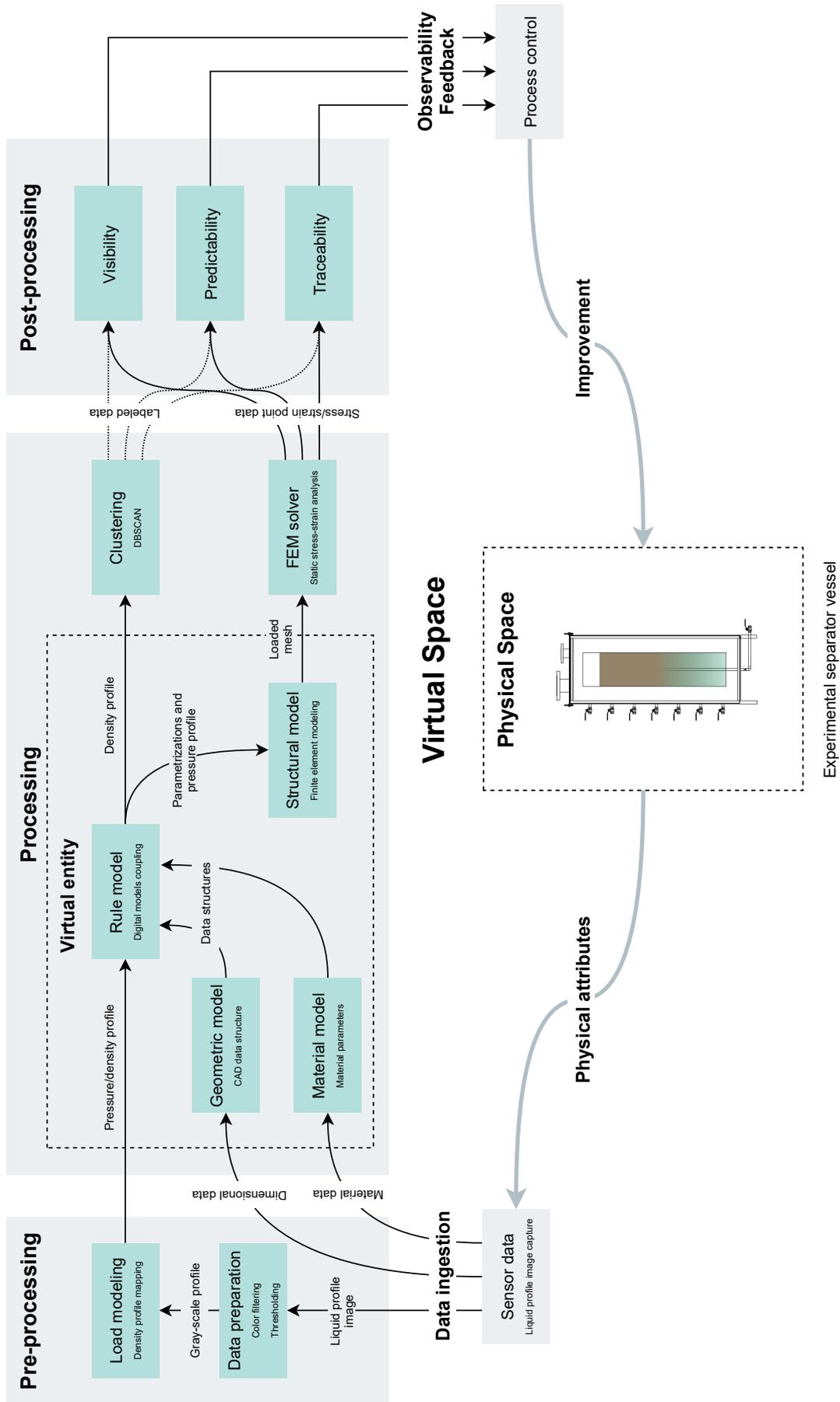


Figure 4.2.1 – Proposed DT architecture model.

processed load, a finite element method (FEM) simulation is performed in order to compute stress and strain values on each node of the model's structural mesh. Furthermore, the pre-processed load is clustered by DBSCAN algorithm in order to automatically group each layer. This clustering approach can be an effective feedback for enhancing the separator level control system by defining accurate setpoints for actuating valves at the liquid inlet and outlet. Finally, in the post-processing stage, at the end of the pipeline, the FEM simulation and clustering results are used to identify the interface levels as well as the separator's safety factor and fatigue life under the pre-established operation regime to trace the spots of major stress and strain concentrations across the experimental separator's body. Providing means of visibility, predictability and traceability, it is possible to augment the separator observability at the process context.

It is worth mentioning that the pipeline stages must be distributed across ICT infrastructure for being functional. For computational capability reasons, edge and cloud computing can share distinct workloads present in the architecture. Since the pre-processing stage requires less computational resources, i.e., processing power and memory, it could lie on the edge, where the physical space meets the digital space. On the other hand, the processing and post-processing stages could be set at cloud, since these steps contains heavier workloads. Figure 4.2.2 depicts the edge and cloud information technology infrastructure integration with the operational technology in oil & gas industry.

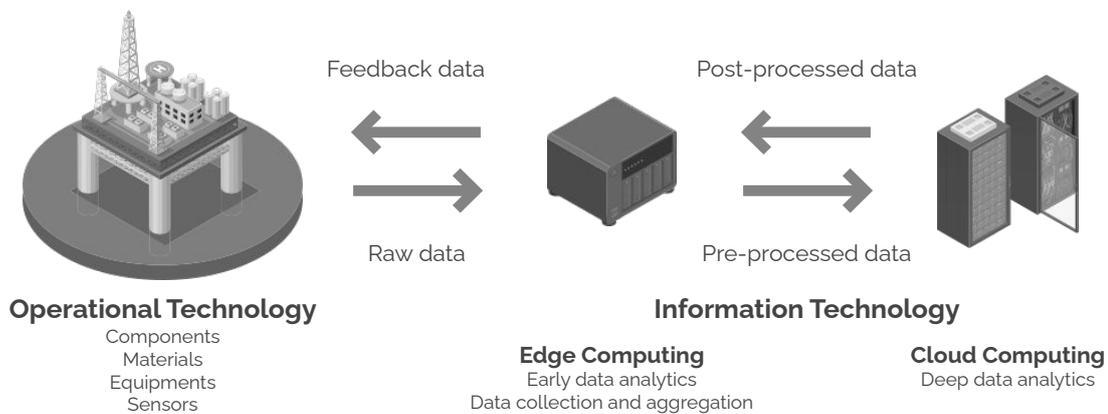


Figure 4.2.2 – Operational technology integrated with cloud and edge infrastructures for implementing the pipeline from the DT architecture model.

By establishing the integration between operating technology and information technology, it is possible to guarantee the communication of data read by the sensors referring to the interface levels clustering and the density profile. After dynamically integrating this information to the separator vessel digital models, its states and behaviors can be predicted in diversified operational scenarios. Thus, with the density profile, it is possible to analyze the effects of hydrostatic pressure load on the separator vessel walls.

Additionally, with the clustered interface levels, the pressure and level control system can be optimized with extra information about foam and emulsion transition layers.

4.3 Experiment Methodology

Initially, the separator were filled with Mobil AW68 oil and fresh water until the oil-air interface level have reached ~ 1.4 m of height (minimum level). Then, additional fresh water was added into the equipment until the oil-air interface level reached ~ 1.7 m of height (maximum level). The difference between the minimum and maximum levels (~ 0.3 m) represents the load range that characterizes the amplitude of a cyclic operation of separator loading and unloading. This operational profile is the "what-if" dynamic load scenario considered for this work. Besides that, the static load operation scenario for each level condition will also be considered. The initial assumptions considered for this case study are: Normal Temperature and Pressure (NTP) conditions (20 °C and 1 atm), gravitational acceleration of 9.81 m/s², incompressible liquids, air as an ideal gas, neglected material hysteresis, and neglected environment effects (e.g, corrosion, oxidation, friction, etc.).

4.4 Pre-processing Image Data

In order to capture the separator's display image, a Samsung SAK2L4 ISOCELL CMOS image sensor with 12 megapixels of resolution, and a pixel size of 1.40 μm was used. The sensor captured the two frames related to minimum and maximum separator's load states in which one can observe a gradient gray-scale color range that represents the continuous changes in liquid density along the vertical axis, i.e., the height, as depicted in Figure 4.4.1 (a) and (b).

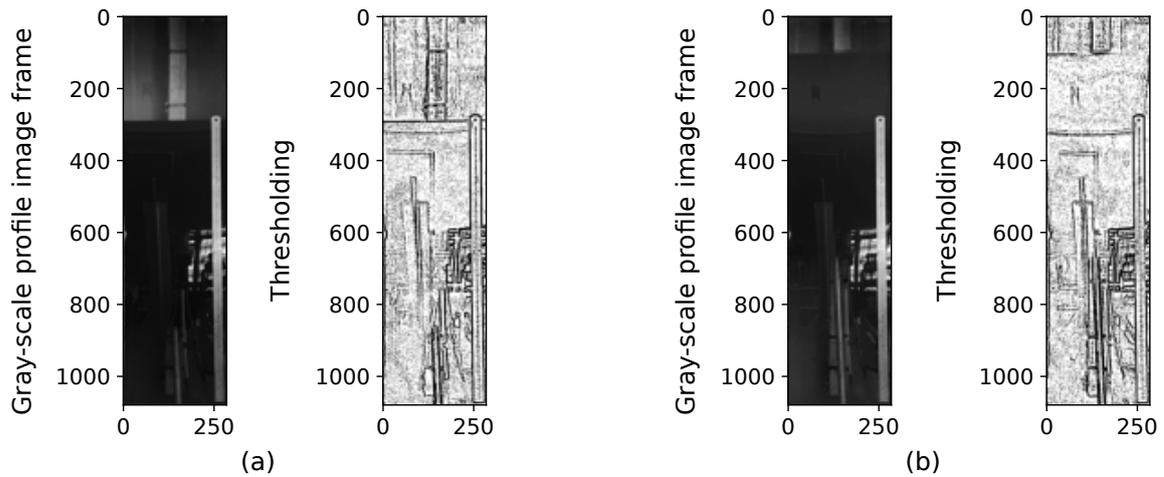


Figure 4.4.1 – Oil and water content image and thresholded frames: (a) at minimum load, (b) at maximum load.

According to Section 3.3, the numerical color code series for this gray-scale gradient can be mapped to the density profile for the separator load state. However, due to the display's reflexive and translucency content, the picture frames are clearly polluted by ambient reflections as well as light refractions. As a result, the adaptive thresholding technique was used to detect the image's least contaminated regions, as shown in Figure 4.4.1 (a) and (b). Using this discrete image segmentation algorithm, this method aids in the identification of reflection and refraction contents. The image frames are then windowed to the area of interest, and the corresponding arrays of the image frame height are collected, which carry the mean color code values along the horizontal axis, as illustrated in 4.4.2 (a) and (b).

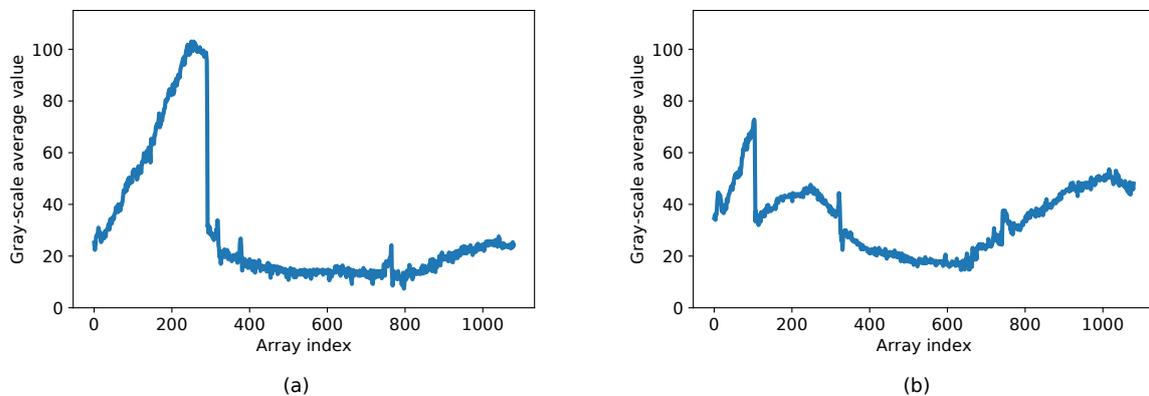


Figure 4.4.2 – Gray-scale average arrays: (a) at minimum load, (b) at maximum load.

These sequences provide information about the single-phase layers, the emulsion

layer, and the oil-air interface layers. During separator loading, both oil-in-water and water-in-oil emulsion are produced as a result of chemical interactions between the liquids particles. Thus, there are no laminar and well-defined layers due to emulsion formation, which is influenced by the characteristics of water and oil in terms of composition and behavior (LANGEVIN et al., 2004; LEAL-JUNIOR et al., 2018). Recalling from Section 3.1, the density profile data can be modeled by a s-shaped or sigmoid curve, by observing the region of emulsion with its adjacent layers. Then, the single-phase and emulsion layers array chunks are modeled using non-linear regression through logistic curve fitting. The obtained models, as depicted in Figure 4.4.3 (a) and (b), returned RMSE values of $7.805 \cdot 10^{-2} \text{ kg/m}^3$ and $5.9 \cdot 10^{-2} \text{ kg/m}^3$, respectively, and R^2 s of 0.865 and 0.964, respectively.

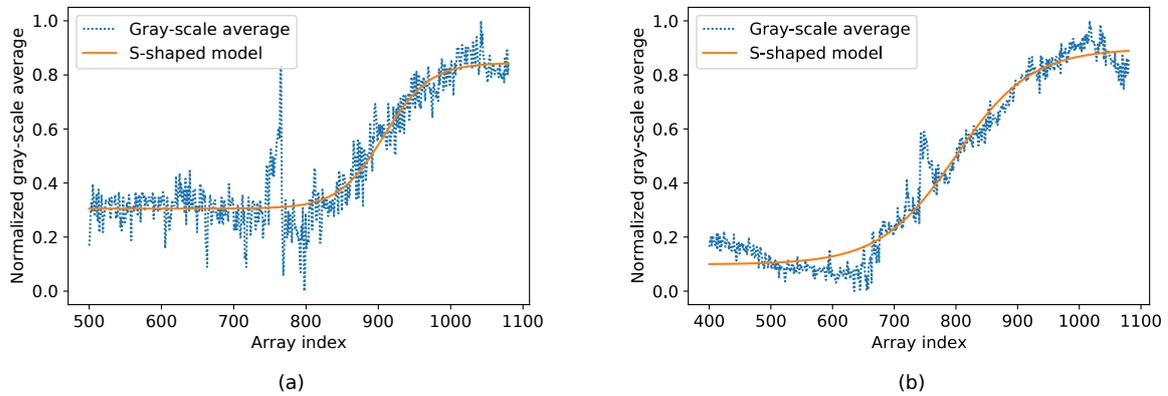


Figure 4.4.3 – Non-linear regression models: (a) at minimum load, (b) at maximum load.

The minimum values of each gray-scale code series first derivative are used to identify the location of the oil-air interface layer, as demonstrated in Section 3.1. Figure 4.4.4 (a) and (b) detail the derivatives.

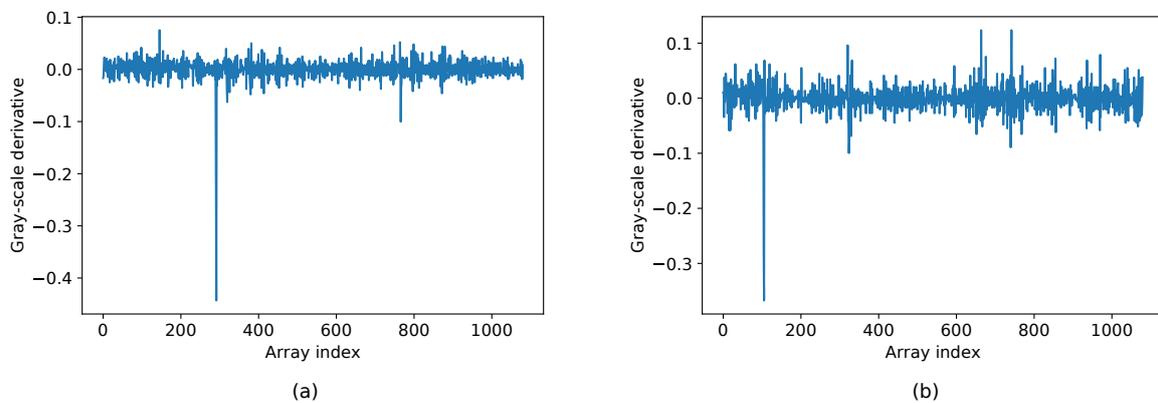


Figure 4.4.4 – Gray-scale first derivatives: (a) at minimum load, (b) at maximum load.

Then, this data array is linearly mapped to the density profile scale based on the Mobil AW68 oil (880 kg/m³), fresh water (998.21 kg/m³) and air (1.204 kg/m³) densities at NTP as the referential values. The obtained model for separator's density profile is characterized by the piecewise function expressed in Equation 4.1 and derived from Equation 3.5.

$$\rho(z) = \begin{cases} S & z < z_{air} \\ \rho_{air} & z \geq z_{air} \end{cases} \quad (4.1)$$

where $\rho(z)$ is the density profile function, z is the level height, S is the s-shaped fitted curve, ρ_{air} is the air density, and z_{air} is the oil-air interface level. Hence, the emulsion and its adjacent single-phase layers are modeled by the s-shaped curve due to their smooth density transition, whereas the oil-air interface is represented by a step density shift, since the transition is abrupt there and the foam layer is not observed in the profile. Figure 4.4.5 (a) and (b) illustrate the resultant load models.

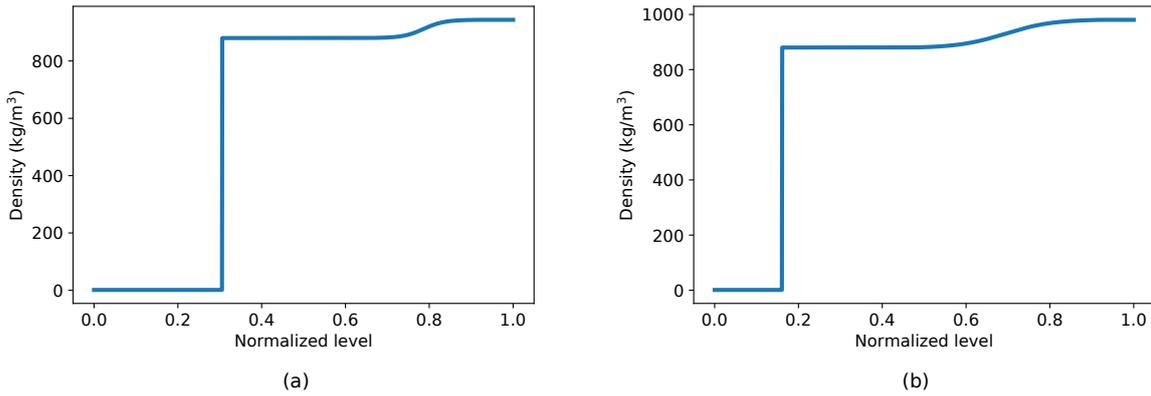


Figure 4.4.5 – Density profiling results: (a) at minimum load, (b) at maximum load.

The obtained density pressure profile is then integrated over the vertical axis to get its correspondent separator's pressure profile, as derived from Equation 2.11 and depicted by Equation 4.2:

$$P(z) = - \int_{z_0}^z \rho(z)g dz + P_0 \quad (4.2)$$

where $P(z)$ is the pressure profile function, z is an arbitrary level, z_0 is the initial level, ρ is the density function expressed in Equation 4.1, g is the gravitational acceleration, and P_0 is the initial pressure. Figure 4.4.6 details the pressure profiles for the minimum and maximum load states. This outcomes are the inputs for the processing stage in the pipeline.

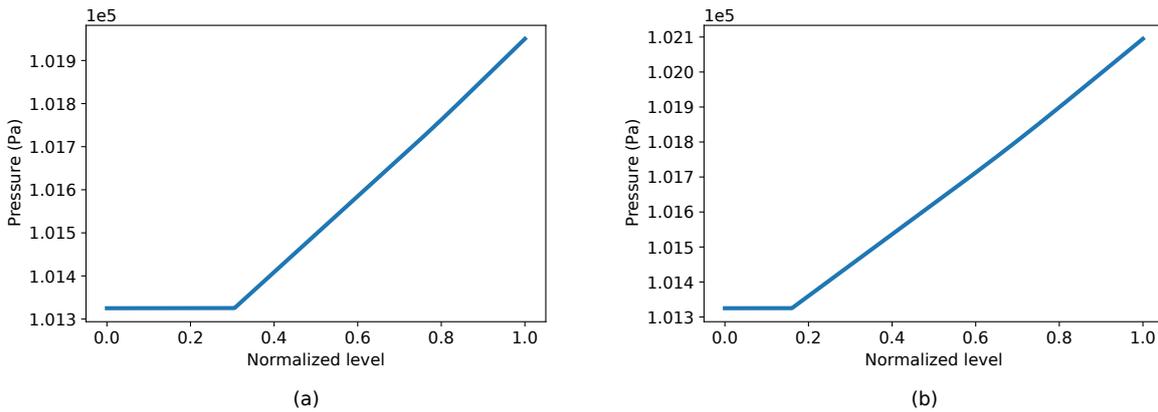


Figure 4.4.6 – Pressure profile outcomes: (a) at minimum load, (b) at maximum load.

4.5 Processing Load Model Data

This stage is the core of the proposed separator DT architectural model. Here lies the separator virtual entity that is the composition of digital models that are the geometric, structural, material and rule models. At this point, the virtual element is updated with the pre-processed data, i.e., pressure profile load model, in order to assume the separator current load state. After that, the pipeline runs a FEM solver to process the updated virtual entity in order to run static analysis and compute nodal stress and strain values distributed across the separator shell material continuum. Geometric model is a data structure that describes how the physical element is defined and how it occupies the physical space. Generally, this model consists in a computer-aided design (CAD) file.

Structural model is another data structure that describes how the geometry defined by the geometric model is subdivided. This model is derived from the geometric model and expresses the discretized separator's body in order to be the direct input of the FEM solver (LOGAN; IMRAN, 2017). Commonly, structural model is represented by a mesh that is composed by small interconnected subregions called elements or cells (RAO, 2018). In order to optimize FEM simulation, the CAD and structural models were simplified by removing the flanges, pressure outlets and liquid inlet/outlet from them, since these parts adds several superfluous elements to the separator's topology and do not compromise the overall structural behavior. Thus, the adopted separator's mesh contains 7602 quadratic tetrahedron elements (non-linear cells) and 15250 nodes. Geometric and structural models are depicted in Figure 4.5.1 (a) and (b), respectively.

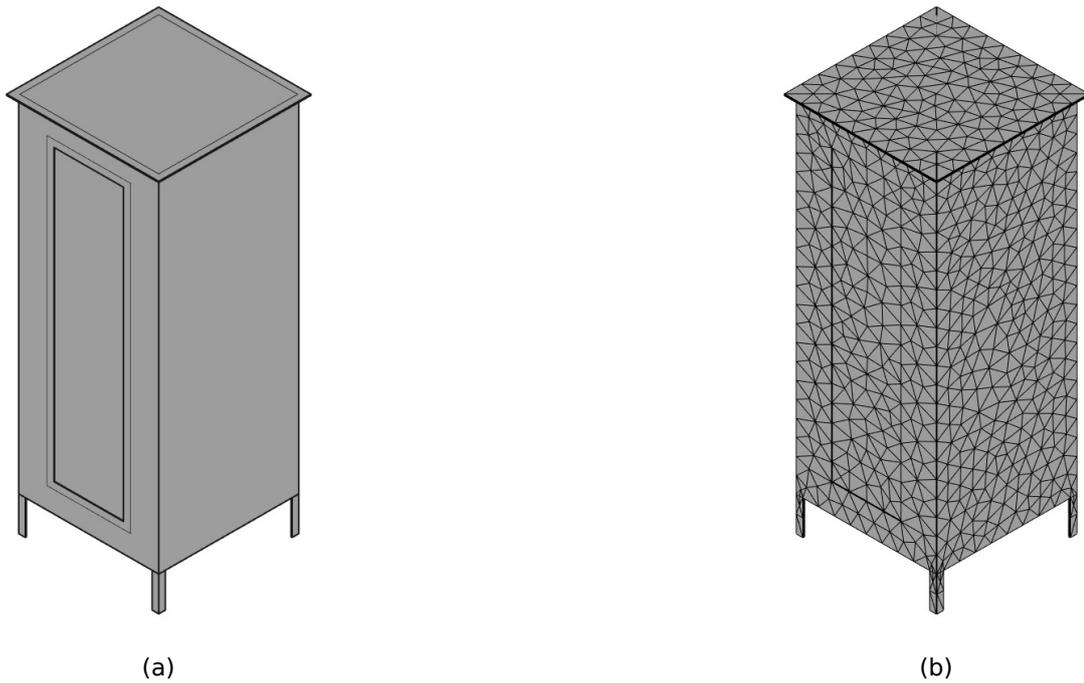


Figure 4.5.1 – Separator prototype orthographic views for (a) geometric and (b) structural simplified models.

Material model is a parametric database that gathers the separator’s multi-material physical properties. This model complements the structural one and defines the materials dynamics during simulation execution. Table 1 lists the material properties considered for the carbon steel and polycarbonate (THE... , 2021a; 2013, 2013; THE... , 2021b; SOLUTIONS et al., 2021; MATERIAL, 2021; HASSANIFARD; HASHEMI, 2020).

Table 1 – Material parameters of carbon steel and polycarbonate.

Parameter	Carbon steel	Polycarbonate
Ultimate tensile strength (MPa)	420	67.5
Density (kg/m ³)	$1.87 \cdot 10^3$	$1.26 \cdot 10^3$
Yield strength (MPa)	350	94.6
Poisson ratio	0.29	0.33
Ultimate strain	22	$7.5 \cdot 10^{-2}$
Shear modulus (MPa)	$8 \cdot 10^4$	6.6
Compressive strength (MPa)	350	140
Young’s modulus (MPa)	$2.05 \cdot 10^5$	$4.05 \cdot 10^3$
Fatigue strength coefficient (MPa)	850	835
Fatigue strength exponent	-0.12	-0.42

Lastly, the rule model consists of an algorithmic rule set that ties together the other models. It describes how the physical interactions happen as well as the procedures

related to couple the pressure load to the separator's digital models together. Here, the importance of the analytical load profile obtained in the previous stage becomes evident. The analytical load model, composed by 4.4.3 and 4.4.4, ensures that pressure values, measured by the computer vision technique and modeled by non-linear regression, are distributed across all the structural model elements.

Thus, the obtained pressure load model needs to be integrated to the structural model, since it is acting on the experimental separator surfaces. The hydrostatic pressure is acting inside the equipment, whereas the atmospheric pressure is acting outside and inside as well. The process that aggregates the load model to the structural model consists in sweeping every element external face contained in its mesh. On each iteration, the z component vector, related to the level height, is obtained from the centroid position of the element face. Then, the z component enters in the analytical model in order to estimate the pressure applied on the element face centroid. This process is recursively executed for all faces in the structural model that are directly exposed to pressure in order to distribute the hydrostatic forces along the equipment surface. Figure 4.5.2 depicts the element face centroid from structural model mesh surface, where the hydrostatic force acts.

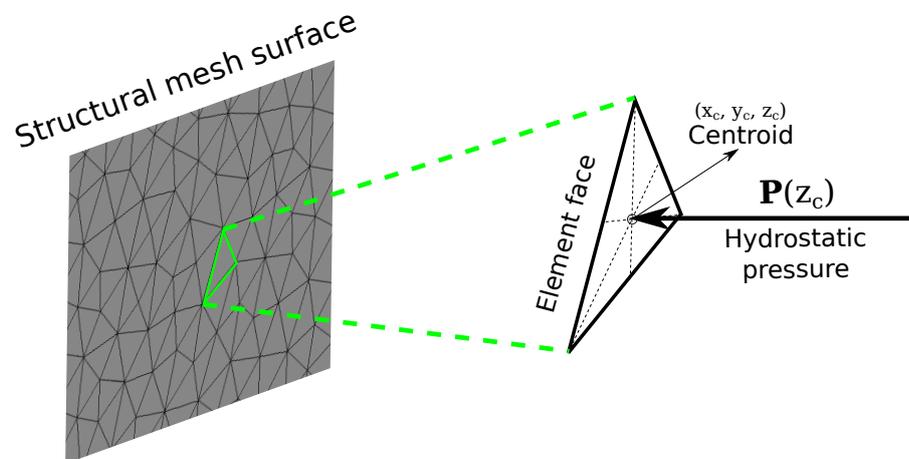


Figure 4.5.2 – Integration between structural and load models.

Separator's geometric, material and structural models were built on Freecad design software, its rule model was implemented in Python programming language, and CalculiX was used as the FEM solver to run stress and strain static analysis (RIEGEL; MAYER; HAVRE, 2021; DHONDT; WITTIG, 2021).

The separator's DT prototype was tested in two distinct scenarios. In the first scenario, the separator was considered to be depressurized, i.e., the experimental real case, in which the initial pressure at oil-air interface is the atmospheric pressure. Figure 4.5.3

(a) and (b) show depressurized separator stress and strain distributions at maximum load, respectively.

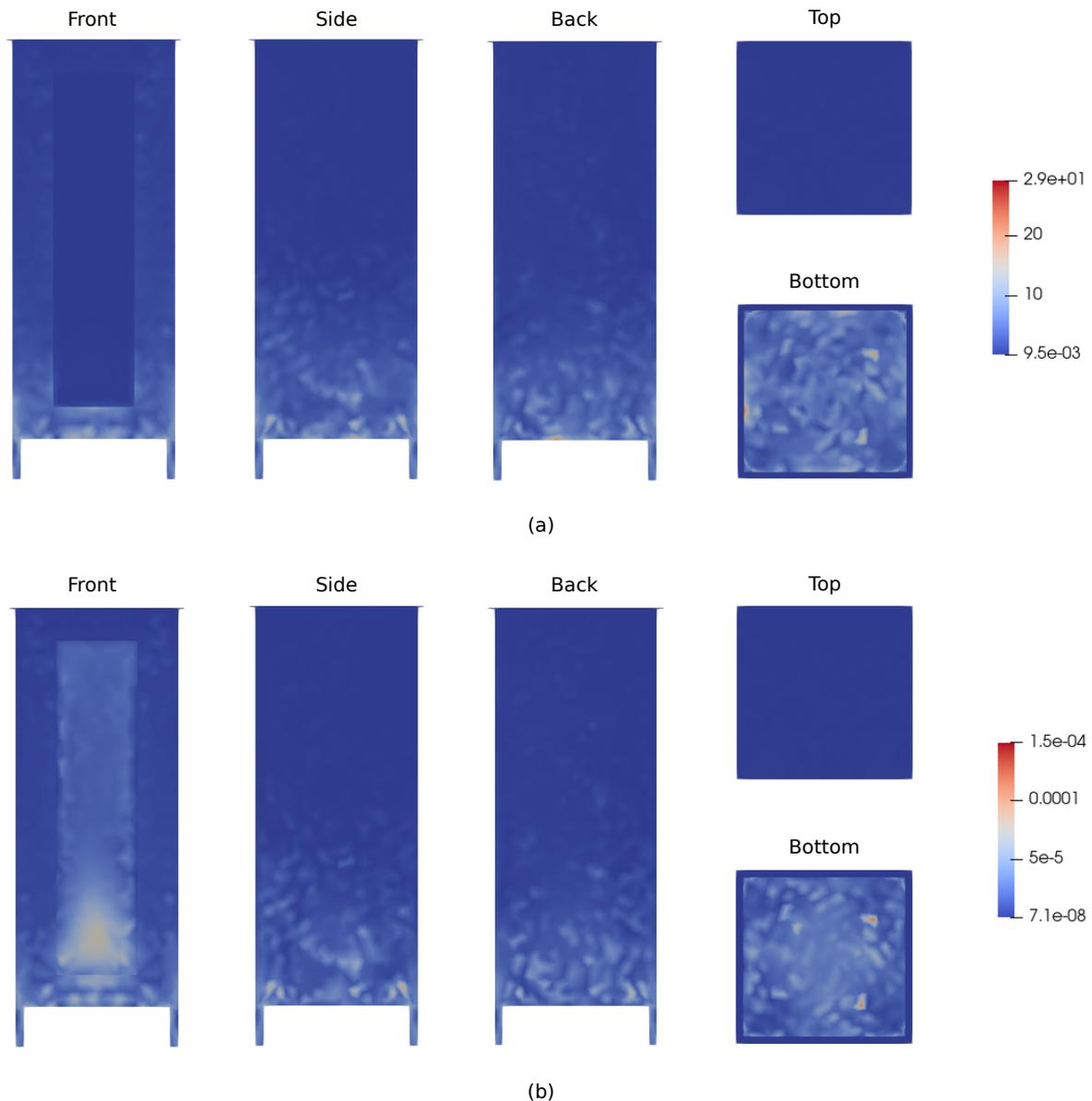


Figure 4.5.3 – Depressurized separator: (a) stress distribution, and (b) strain distribution. The stress scale is in MPa.

In the second scenario, the separator was considered to be pressurized, i.e., the "what-if" case, in which the internal pressure was considered to be equivalent to the free water knockout vessel (FWKO) operational pressure (50 psig) (PETROWIKI, 2021). Figure 4.5.4 (a) and (b) show pressurized separator stress and strain distributions at maximum load, respectively. The stress and strain results from the FEM solver were graphically visualized by Paraview software tool (AHRENS; GEVECI; LAW, 2005).

By inspecting the spots across the experimental separator's body with the highest

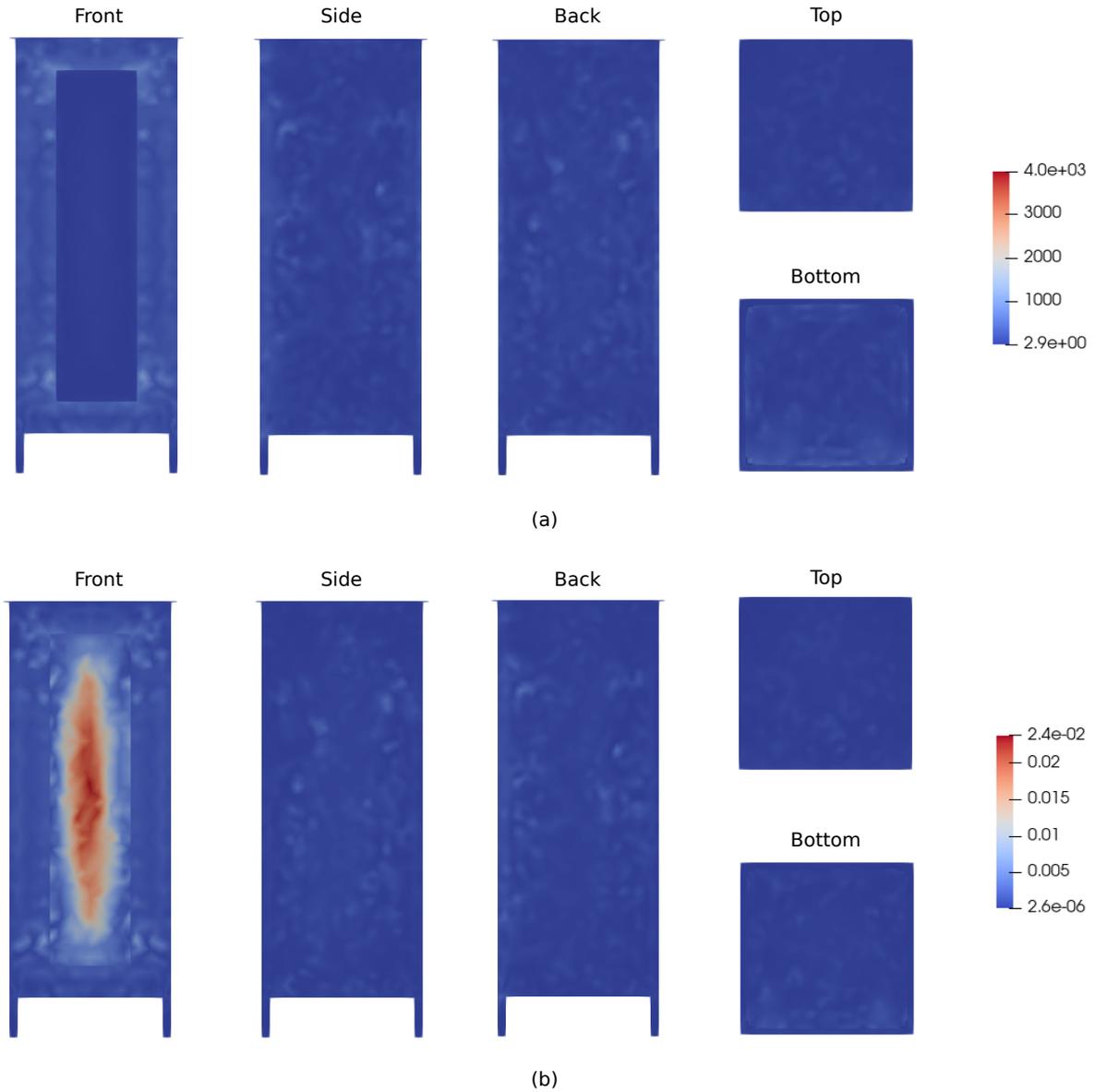


Figure 4.5.4 – Pressurized separator: (a) stress distribution, and (b) strain distribution. The stress scale is in MPa.

mean stress and strain values, one can track the most critical points of each material where the failure can initially occur. In the first scenario, the highest stress values are located at equipment's bottom in the carbon steel part of separator's body, whereas the highest strain values are observed at the bottom of polycarbonate display and carbon steel portion as well. It is relevant noting that there is not stress concentration in the polycarbonate display for this case. In the second scenario, the stress is more uniformly distributed across the separator's body, whereas the highest strain values are observed along polycarbonate display longitudinal axis. It is relevant noting that there some punctual stress concentration locations in the carbon steel portions next to polycarbonate display boundaries. Through Figure 4.5.3 and Figure 4.5.4, one can observe that the pressurization state directly

affects the pattern of stress and strain distribution on separator's body, especially in the polycarbonate display. It happens because the effect of pressure profile along the height due to liquid weight is more prominent when depressurized than pressurized.

In real processes, the liquid levels tend to be constant because of separators' level control system. Nevertheless, the pressure loads can remain variable due to the gas pressure fluctuations over time. Thus, the proposed architecture model can redefine the traditional observability's concept for the separator's process control system adding a controlled variable not yet considered so far: the reliability. Through the separator DT, the process control would not only be sensitive to liquid levels, but also to the entire pressure profile, to the current structural life cycle and to critical points of stress and strain concentrations in the equipment's body. Observing the changes in the predicted life and the most critical stress locations during operation, the process parameters can be preventively adjusted in order to preserve separator's integrity and then optimize production. Therefore, by ensuring that separator operates the way it was specified in the design for a longer time, reliability is enriched by its DT.

It is worth mentioning that the second scenario in the analysis suggests that the separator DT could be expanded if other sensing systems are coupled to the equipment. Pressure, temperature, chemical, vibration and optical sensors can increase the dimensionality of separator's physical monitored attributes in order to allow additional structural predictive failure analysis (i.e., overheating, leakage, stress-corrosion crack, hydrogen embrittlement, brittle fractures, ductile fractures, stress rupture, creep rupture ([ASM, 2002](#))), besides fatigue. Thus, with a wide variety of sensor data updating the digital models, the DT will become more complex, it will mirror the physical element with more accuracy and fidelity, and the predictive analysis will be extended, further improving decision-making and process control.

4.6 Post-processing Stress and Strain Data

The "forward-looking" aspect lies in the DT's ability to provide predictive analytics and insights. For this case study, it will be necessary to evaluate the reliability of the experimental separator through the analysis of structural failures under operational conditions of static and dynamic loading. Hence, the nodal stress-strain data obtained at previous pipeline stage has to be post-processed in order to predict separator's safety factor, in case of static load, as well as fatigue life under the considered dynamic load operation.

The majority of equipment failures are caused by time-varying loads rather than static loads. These failures typically happen at stress thresholds that are much smaller than the materials' yield strengths ([NORTON, 2011](#)). Often, fatigue failures begin at a

crack. The crack may have existed in the material since its manufacture, or it may have formed as a result of cyclic straining across stress concentrations over time (NORTON, 2011). Whereas the safety factor, N , expressed in Equation 2.13 has to be computed for static loads, the dynamic loads, in which mean stresses are non-zero, are characterized by estimating the fatigue life in number of cycles, as defined by Equation 4.3:

$$N_f = \frac{1}{2} \left(\frac{\sigma_a}{\sigma'_f - \sigma_m} \right)^{1/b} \quad (4.3)$$

where N_f is the fatigue cycles, σ_a is stress amplitude, σ_m is the mean stress, σ'_f is the material's fatigue strength coefficient, and b is the material's fatigue strength exponent or Basquin exponent (SURESH, 1991; DOWLING; PRASAD; NARAYANASAMY, 2013). In separator's case, multiaxial stresses are applied, so it is necessary to compute von Mises effective amplitude and mean stresses using the principal stresses (DOWLING; PRASAD; NARAYANASAMY, 2013). The fatigue life implies the number of cycles, N_f , that the equipment will operate until the event of a crack initiation occurs (SURESH, 1991). Besides the fatigue life prediction, this pipeline stage can also sort the full stress-strain datasets to trace the spots in separator's structure with the highest stress or strain amplitudes, since there would be the most probable point to firstly initiate a crack.

4.6.1 Predictions and Insights

4.6.1.1 Interface Levels Identification

In order to identify the interface levels, DBSCAN algorithm was used to cluster each density profile model, as illustrated in Figure 4.6.1 (a) and (b).

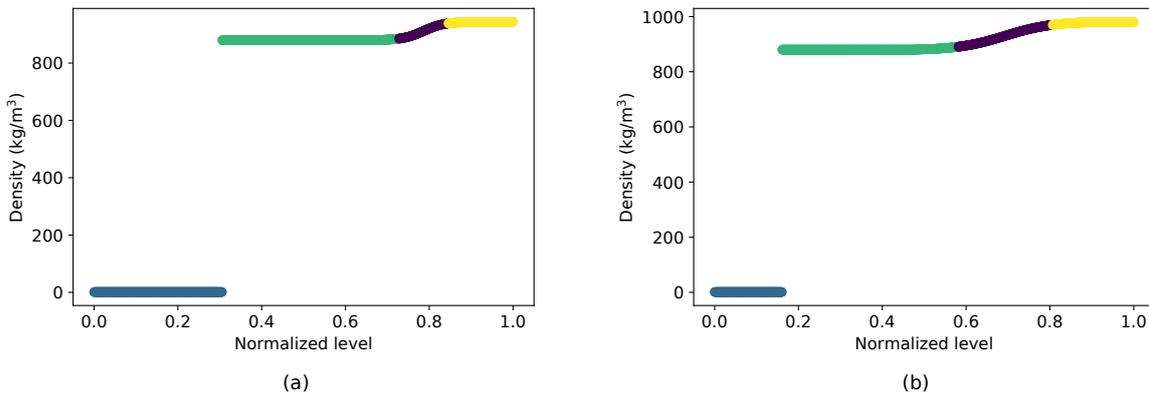


Figure 4.6.1 – Density profiling clustering results: (a) at minimum load, (b) at maximum load.

Observing the obtained clustered density profiles, it is evident that the emulsion layer almost double its length having an increase of 91.391% after additional fresh water

was added into the experimental separator vessel. Its level range changed from $\sim 0.713 - 0.822$ to $\sim 0.575 - 0.8$. The oil layer did not change, and its level range was shifted from $\sim 0.313 - 0.713$ to $\sim 0.163 - 0.575$. The water layer has been increased by 25.126%, and its level range was shifted from $\sim 0.838 - 1$ to $\sim 0.8 - 1$. The air layer has been reduced by 47.328%, and its level range was shifted from $\sim 0 - 0.313$ to $\sim 0 - 0.163$. The maximum distance between neighbors in a single cluster, δ , and the minimum number of points in a dense region, ν , considered for this analysis were 0.3 and 5, respectively.

4.6.1.2 Structural Failures

Since the stresses acting in the experimental separator vessel material are multiaxial, it is necessary to compute the von Mises effective stress values through Equation 2.12 in order to predict safety factors and fatigue life. Figures 4.6.2 (a), (b), (c) and (d) show the von Mises effective stresses distributions for equipment's carbon steel and polycarbonate parts at minimum and maximum loads, respectively.

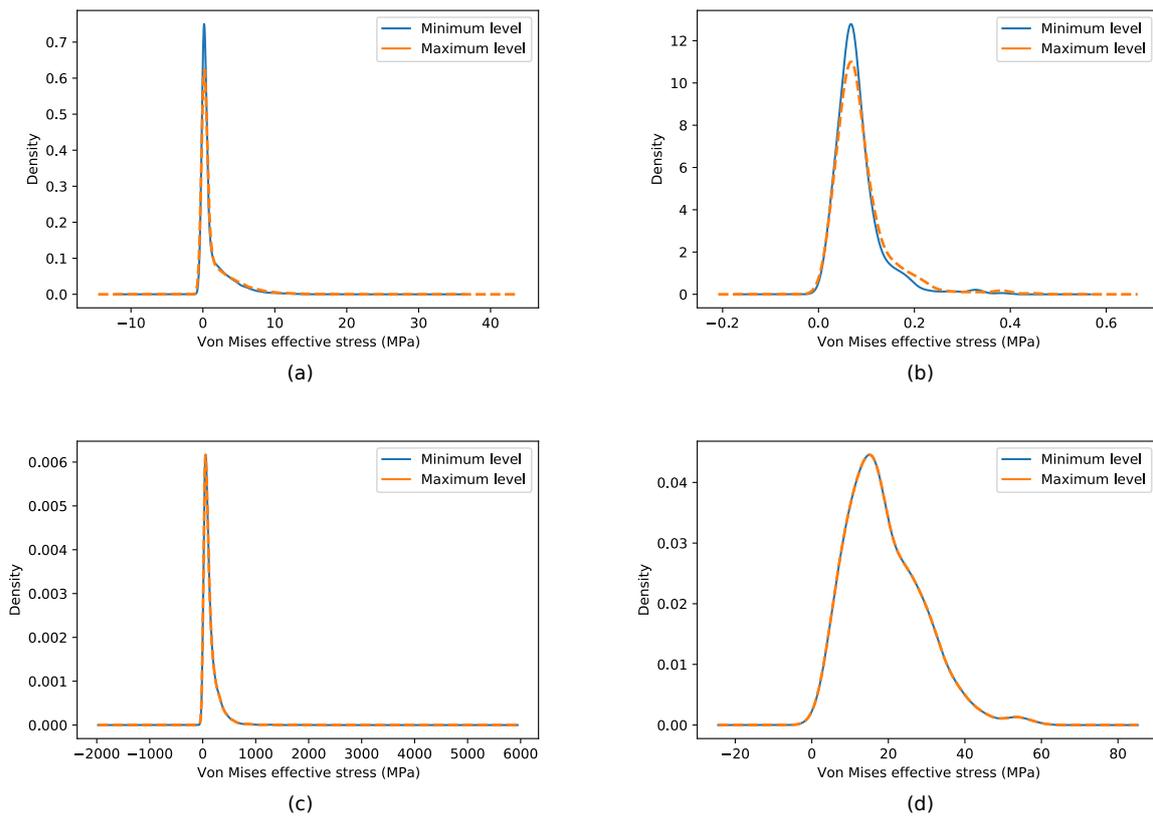


Figure 4.6.2 – Von Mises effective stress distributions: (a) in carbon steel (depressurized), (b) in polycarbonate (depressurized), (c) in carbon steel (pressurized), (d) in polycarbonate (pressurized).

Observing Figures 4.6.2 (a) and (b), in the scenario in which the equipment is depressurized, it is evident that the effective stresses in carbon steel mostly vary in the

range between 0 and 10 MPa, while in polycarbonate this interval is much smaller being constrained from 0 to ~ 0.3 MPa. Additionally, in the Figures 4.6.2 (c) and (d), in the scenario in which the equipment is considered to be pressurized, one can see that the effective stresses in carbon steel mostly vary in the range between 0 and ~ 500 MPa, whereas, for polycarbonate, this interval covers from 0 to ~ 60 MPa. It is worth noting that the differences between the distributions at minimum and maximum levels are more evident when the experimental separator vessel is depressurized than pressurized.

4.6.1.2.1 Static Load

The safety factors were computed for the static load analysis based on the worst case, i.e., the highest obtained effective stress on structural mesh. In the depressurized scenario, at minimum level, the safety factors are 14.41 and 247.786 for carbon steel and polycarbonate, respectively. On the other hand, at maximum level, the safety factors are 11.998 and 211.722 for carbon steel and polycarbonate, respectively. In the pressurized scenario, at minimum level, the safety factors are $8.832 \cdot 10^{-2}$ and 1.638 for carbon steel and polycarbonate, respectively. Then, at maximum level, the safety factors remain the same values for both materials. Therefore, it is observed that the safety factor variability is drastically reduced at higher hydrostatic pressure loads. Since the safety factor of carbon steel is decreased to $8.832 \cdot 10^{-2}$, i.e., lower than 1, in the pressurized scenario, it indicates that the carbon steel part would fail if subjected to this working condition.

4.6.1.2.2 Dynamic Load

It is worth mentioning that the predictive analysis is based on the nodes with the highest effective amplitude stress in each separator's material in this case study. If the separator is depressurized under the considered cyclic load conditions, the predicted fatigue life of carbon steel and polycarbonate are $8.579 \cdot 10^{21}$ and $8.402 \cdot 10^{10}$ cycles, respectively. The highest effective stress amplitude and mean values for carbon steel are 1.768 MPa and 26.391 MPa, respectively. The highest effective stress amplitude and mean values for polycarbonate are $1.611 \cdot 10^{-2}$ MPa and $2.97 \cdot 10^{-2}$ MPa, respectively.

If the separator is pressurized under the same cyclic load conditions, the predicted fatigue life of carbon steel and polycarbonate are $1.078 \cdot 10^5$ and $1.288 \cdot 10^{11}$ cycles, respectively. The highest effective stress amplitude and mean values for carbon steel are 1.771 MPa and $8.423 \cdot 10^3$ MPa, respectively. The highest effective stress amplitude and mean values for polycarbonate are $1.292 \cdot 10^{-2}$ MPa and 33.37 MPa, respectively.

The results showed that the carbon steel and polycarbonate have distinct behaviors in each scenario. Whereas the carbon steel cycles were drastically decreased by $7.961 \cdot 10^{16}$ times, the polycarbonate cycles were increased by 1.534 time. The high cycle value of

carbon steel in the depressurized scenario indicates that it would not entail risks of failure due to fatigue in the short and medium term. This means that, if the life of the carbon steel part only depends on the effects of cyclic stress, its physical integrity would remain intact for a long period of time. Observing Figure 2.3.2 as a reference, in the depressurized scenario, the cycles number intercepts the S-N curve beyond its knee, at the asymptotic portion of the diagram.

One can see that the separator's carbon steel is suitable for long cyclic operation in depressurized scenario, while it would have a shortened life for cyclic operation in pressurized scenario. The anomaly behavior on polycarbonate computed fatigue life, whose value was increased, turns evident that the adopted dynamic load analysis approach is suitable only for ductile engineering metals, as mentioned in Section 2.3.

5 Conclusions and Future Work

5.1 Conclusions and Contributions

This work presented a DT architecture model of an experimental separator vessel capable of augment the observability of the equipment as an element of liquid separation process. This architecture includes a computer vision profiling measurement technique combined with clustering for liquid level interface identification.

Through the employed computer vision method, it was possible to estimate and process not only liquid level interfaces but also the entire liquid density and pressure profile over the experimental separator height. This allowed creating a detailed load model for the separator DT even if emulsion layer are present in oil-water mixture. Since the proposed computer vision approach indirectly estimates the density level profile, it is necessary to employ extra thermal and density sensors in order to complement such method. The density profiling and clustering approaches combined represents an promising low cost and safe measurement alternative in providing support to the liquid level control system.

This is crucial for the proposed DT architecture model being capable to establish the link between "what-if" scenarios and "forward-looking" analysis, by which it is possible to identify liquid level interfaces, predict safety factors the separator's and its fatigue life, and trace the major stress concentration point at operation time. Also, its predictions on safety factors and fatigue life under static and dynamic loads can be suitable for maintenance management and process calibration in order to enhance the equipment performance, stretch the equipment life span, improve reliability and also avoid potential human and economical losses.

5.2 Publications

During the writing of this dissertation, the following paper were published, being indirectly linked to the work here presented:

- **Santos, R. A. G. Dos**, et al. "Datacenter Thermal Monitoring Without Blind Spots: FBG-Based Quasi-Distributed Sensing." *IEEE Sensors Journal*, vol. 21, no. 8, 2021, pp. 9869–9876., doi:10.1109/jsen.2021.3058513.

Moreover, this work also led to the submission of the following paper, currently under review:

- **Santos, R. A. G. Dos**, Ribeiro, M. R. N., Leal-Junior, A. G., Pontes, M. J., & Neto, A. F. (2021). "Digital Twins and Augmented Observability: An Architecture for Oil & Gas Separator Vessels". IEEE Software Journal (*submitted*).

Lastly, the author presented the following short course, that is directly related to this dissertation:

- **Santos, R. A. G. Dos**, Ribeiro, M. R. N., Neto, A. F., Martinello, M., & Par-davila, M. (2019). "Remote Monitoring and Diagnostics (RM&D): Sensing and IoT, Communication, Modeling, Data Analytics and Artificial Intelligence". IMOC 2019.

5.3 Future Work

The proposed experimental separator DT has the potential to be evolved in future works, consisting in deploying its architecture model in edge and cloud environment, allowing separator's virtual models optimization and real time dynamic processing for the synchronization between physical and virtual elements. The edge computing infrastructure could offer the necessary resources for integrate the experimental separator vessel DT to the process control system, i.e., TO and TI integration establishment. Thus, the latency issues with feedback from the architecture model to the control system could be mitigated allowing its loop to be closed.

A metrological analysis is necessary to be made in order to statistically evaluate the proposed non-contact density profiling technique by computer vision. Other failure mechanisms, beyond fatigue in dynamic loads, will be considered, since separator vessels are subjected to diversified environment conditions in an industrial plant in terms of pressure, temperature, vibrations and chemical agents. Also, variable amplitude and stochastic loads will be considered, since the operational regime is non-deterministic for real separators.

Bibliography

- 2013, W. b. A. . *AISI 1020 Carbon Steel (UNS G10200)*. 2013. Disponível em: <<https://www.azom.com/article.aspx?ArticleID=9145>>. Citado na página 75.
- 2SOLVE. *Projeto Tanque de Testes*. [S.l.], 2019. Citado 4 vezes nas páginas 10, 66, 103, and 104.
- A., C. Y.; CIMBALA, J. M. *Fluid mechanics: Fundamentals and applications*. [S.l.]: McGraw-Hill Education, 2020. Citado 3 vezes nas páginas 9, 34, and 35.
- AHRENS, J.; GEVECI, B.; LAW, C. Paraview: An end-user tool for large-data visualization. In: *The Visualization Handbook*. [S.l.: s.n.], 2005. Citado na página 77.
- Al-Naamany, A. M.; Meribout, M.; Al Busaidi, K. Design and implementation of a new nonradioactive-based machine for detecting oil–water interfaces in oil tanks. *IEEE Transactions on Instrumentation and Measurement*, v. 56, n. 5, p. 1532–1536, 2007. Citado na página 56.
- ASLAM, M.; TANG, T. B. A high resolution capacitive sensing system for the measurement of water content in crude oil. *Sensors (Basel, Switzerland)*, v. 14, p. 11351–11361, 07 2014. Citado na página 60.
- ASLAM, M. Z.; BOON, T. T. Differential capacitive sensor based interface circuit design for accurate measurement of water content in crude oil. In: *2014 5th International Conference on Intelligent and Advanced Systems (ICIAS)*. [S.l.: s.n.], 2014. p. 1–6. Citado na página 60.
- ASM. *ASM handbook: Volume 11, Failure analysis and prevention*. [S.l.]: ASM International, 2002. 2686-2740 p. Citado 2 vezes nas páginas 23 and 79.
- AYALA, N. F. et al. Knowledge sharing dynamics in service suppliers’ involvement for servitization of manufacturing companies. *International Journal of Production Economics*, v. 193, p. 538–553, 2017. ISSN 0925-5273. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0925527317302712>>. Citado na página 17.
- Bangemann, T. et al. Integration of classical components into industrial cyber–physical systems. *Proceedings of the IEEE*, v. 104, n. 5, p. 947–959, 2016. Citado na página 18.
- BARKAT, A. et al. Basics of pharmaceutical emulsions: A review. *African journal of pharmacy and pharmacology*, v. 525, p. 2715–2725, 12 2011. Citado na página 47.
- BARTODZIEJ, C. J. The concept industry 4.0. In: _____. *The Concept Industry 4.0 : An Empirical Analysis of Technologies and Applications in Production Logistics*. Wiesbaden: Springer Fachmedien Wiesbaden, 2017. p. 27–50. ISBN 978-3-658-16502-4. Disponível em: <https://doi.org/10.1007/978-3-658-16502-4_3>. Citado na página 17.
- Batagelj, V. et al. Automation of reading of liquid-in-glass thermometers. In: *Proceedings of the 17th IEEE Instrumentation and Measurement Technology Conference [Cat. No. 00CH37066]*. [S.l.: s.n.], 2000. v. 1, p. 299–303 vol.1. Citado na página 60.

BEER, F. P. et al. *Mechanics of materials*. [S.l.]: McGraw-Hill Education, 2020. Citado 3 vezes nas páginas 9, 33, and 34.

BERTHOLD. *EmulsionSENS Multiphase Level Measurement*. 2021. Disponível em: <<https://www.berthold.com/en/process-control/products/level-measurement/multiphase-level-interface-measurement/emulsionsens/>>. Citado 3 vezes nas páginas 9, 23, and 46.

BRADLEY, D.; ROTH, G. Adaptive thresholding using the integral image. *Journal of Graphics Tools*, Taylor & Francis, v. 12, n. 2, p. 13–21, 2007. Disponível em: <<https://doi.org/10.1080/2151237X.2007.10129236>>. Citado na página 63.

BRETTEL, M. et al. How virtualization, decentralization and network building change the manufacturing landscape: An industry 4.0 perspective. *World Academy of Science, Engineering and Technology, International Journal of Mechanical, Aerospace, Industrial, Mechatronic and Manufacturing Engineering*, v. 8, p. 37–44, 2014. Citado 3 vezes nas páginas 17, 18, and 19.

BUDYNAS, R. G.; NISBETT, J. K.; SHIGLEY, J. E. *Shigleys mechanical engineering design*. [S.l.]: McGraw-Hill, 2016. Citado 2 vezes nas páginas 9 and 39.

CENPES. *Nucleonic Density Profiler Measurement Test*. [S.l.], 2019. Citado 2 vezes nas páginas 46 and 57.

Chakravarthy, S.; Sharma, R.; Kasturi, R. Noncontact level sensing technique using computer vision. *IEEE Transactions on Instrumentation and Measurement*, v. 51, n. 2, p. 353–361, 2002. Citado na página 60.

The Future of Three Phase Separator Control, All Days de SPE Asia Pacific Oil and Gas Conference and Exhibition, (SPE Asia Pacific Oil and Gas Conference and Exhibition, All Days). SPE-77891-MS. Disponível em: <<https://doi.org/10.2118/77891-MS>>. Citado na página 24.

CHEN, J. et al. Foaming of crude oil: Effect of acidic components and saturation gas. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, v. 553, p. 432–438, 2018. ISSN 0927-7757. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0927775718304734>>. Citado na página 48.

CHENG, Y. et al. Cyber-physical integration for moving digital factories forward towards smart manufacturing: a survey. *The International Journal of Advanced Manufacturing Technology*, v. 97, p. 1209–1221, 2018. Citado na página 22.

CLANCY, A. S. *\$47M settlement reached in deadly south St. Louis boiler explosion*. 2019. Disponível em: <<https://www.ksdk.com/article/news/local/47m-settlement-reached-in-deadly-south-st-louis-boiler-explosion/63-3b84abc7-6efa-41da-a726-d42ebf8705a9#:~:text=Areportissuedamonth,whenthe-tank-was-installed.>> Citado na página 24.

COREYNEN, W.; MATTHYSSENS, P.; Van Bockhaven, W. Boosting servitization through digitization: Pathways and dynamic resource configurations for manufacturers. *Industrial Marketing Management*, v. 60, p. 42–53, 2017. ISSN 0019-8501. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0019850116300700>>. Citado na página 17.

CSB. *Goodyear Heat Exchanger Rupture*. 2008. Disponível em: <<https://www.csb.gov/goodyear-heat-exchanger-rupture/>>. Citado na página 24.

CSB. *Deadly Tesoro refinery blast caused by high temp "hydrogen attack"*. ISHN, 2010. Disponível em: <[https://www.ishn.com/articles/97860-deadly-tesoro-refinery-blast-caused-by-high-temp-hydrogen-attack#:~:text=TheApril2010fatalexplosion,aChemicalSafetyBoard\(CSB\)](https://www.ishn.com/articles/97860-deadly-tesoro-refinery-blast-caused-by-high-temp-hydrogen-attack#:~:text=TheApril2010fatalexplosion,aChemicalSafetyBoard(CSB))>. Citado na página 24.

CSB. *Tesoro Refinery Fatal Explosion and Fire*. 2010. Disponível em: <<https://www.csb.gov/tesoro-refinery-fatal-explosion-and-fire/>>. Citado na página 24.

CSB. *Loy Lange Box Company Pressure Vessel Explosion*. 2017. Disponível em: <<https://www.csb.gov/loy-lange-box-company-pressure-vessel-explosion-/>>. Citado na página 24.

DALENOGARE, L. S. et al. The expected contribution of industry 4.0 technologies for industrial performance. *International Journal of Production Economics*, v. 204, p. 383–394, 2018. ISSN 0925-5273. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0925527318303372>>. Citado 3 vezes nas páginas 17, 18, and 19.

DEBROY, T. et al. Building digital twins of 3d printing machines. *Scripta Materialia*, v. 135, p. 119–124, 2017. ISSN 1359-6462. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1359646216305991>>. Citado na página 20.

DELIGIANNIS, N. et al. Multi-modal dictionary learning for image separation with application in art investigation. *IEEE Transactions on Image Processing*, v. 26, n. 2, p. 751–764, 2017. Citado na página 60.

DHONDT, G.; WITTIG, K. *A Three-Dimensional Structural Finite Element Program*. 2021. Disponível em: <<http://www.calculix.de/>>. Citado na página 76.

DIXIT, R.; NASKAR, R. Review, analysis and parameterisation of techniques for copy–move forgery detection in digital images. *IET Image Processing*, v. 11, n. 9, p. 746–759, 2017. Disponível em: <<https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/iet-ipr.2016.0322>>. Citado na página 60.

DOWLING, N. E.; PRASAD, K. S.; NARAYANASAMY, R. *Mechanical behavior of materials: engineering methods for deformation, fracture, and fatigue*. [S.l.]: Pearson, 2013. 416-490 p. Citado 7 vezes nas páginas 9, 37, 38, 39, 40, 41, and 80.

Dregger, J. et al. The digitization of manufacturing and its societal challenges: a framework for the future of industrial labor. In: *2016 IEEE International Symposium on Ethics in Engineering, Science and Technology (ETHICS)*. [S.l.: s.n.], 2016. p. 1–3. Citado na página 17.

Díaz, C. A. R. et al. Liquid level measurement based on fbg-embedded diaphragms with temperature compensation. *IEEE Sensors Journal*, v. 18, n. 1, p. 193–200, 2018. Citado na página 55.

EPPEL, S. Tracing liquid level and material boundaries in transparent vessels using the graph cut computer vision approach. *ArXiv*, abs/1602.00177, 2016. Citado na página 60.

EPPEL, S.; KACHMAN, T. Computer vision-based recognition of liquid surfaces and phase boundaries in transparent vessels, with emphasis on chemistry applications. *ArXiv*, abs/1404.7174, 2014. Citado na página 67.

EVANS, F. et al. Future trends in oil and gas visualization. In: *Proceedings of the Conference on Visualization '02*. USA: IEEE Computer Society, 2002. (VIS '02), p. 567–570. ISBN 0780374983. Citado na página 61.

FATORACHIAN, H.; KAZEMI, H. A critical investigation of industry 4.0 in manufacturing: theoretical operationalisation framework. *Production Planning & Control*, Taylor & Francis, v. 29, n. 8, p. 633–644, 2018. Disponível em: <<https://doi.org/10.1080/09537287.2018.1424960>>. Citado na página 18.

FINGAS, M.; FIELDHOUSE, B. Studies of the formation process of water-in-oil emulsions. *Marine Pollution Bulletin*, v. 47, n. 9, p. 369–396, 2003. ISSN 0025-326X. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0025326X03002121>>. Citado na página 48.

GARCÍA-GOLDING, F. et al. Sensor for determining the water content of oil-in-water emulsion by specific admittance measurement. *Sensors and Actuators A: Physical*, v. 47, n. 1, p. 337–341, 1995. ISSN 0924-4247. Disponível em: <<https://www.sciencedirect.com/science/article/pii/0924424794009177>>. Citado na página 60.

GILCHRIST, A. *Industry 4.0: the industrial internet of things*. [S.l.]: Apress, 2016. Citado 3 vezes nas páginas 17, 18, and 19.

GLAESSGEN, E.; STARGEL, D. The digital twin paradigm for future nasa and u.s. air force vehicles. In: . [S.l.: s.n.], 2012. ISBN 978-1-60086-937-2. Citado 2 vezes nas páginas 20 and 21.

GONZALEZ, R. C.; WOODS, R. E. *Digital image processing*. [S.l.]: Pearson, 2018. Citado na página 62.

GRIEVES, M. Digital twin: manufacturing excellence through virtual factory replication. *White paper*, Florida Institute of Technology, v. 1, p. 1–7, 2014. Citado 2 vezes nas páginas 19 and 21.

Guirong Lu et al. A simple and valuable method for detecting levels of interface of oil-water and oil layer. In: *2009 ICCAS-SICE*. [S.l.: s.n.], 2009. p. 1836–1839. Citado na página 56.

HASSANIFARD, S.; HASHEMI, S. M. On the strain-life fatigue parameters of additive manufactured plastic materials through fused filament fabrication process. *Additive Manufacturing*, v. 32, p. 100973, 2020. ISSN 2214-8604. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2214860419303495>>. Citado na página 75.

HE, Y. et al. Reliability-oriented design of integrated model of preventive maintenance and quality control policy with time-between-events control chart. *Computers & Industrial Engineering*, v. 129, p. 228–238, 2019. ISSN 0360-8352. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0360835219300518>>. Citado na página 22.

- HENNIG, C. M. et al. *Handbook of cluster analysis*. [S.l.]: CRC Press, 2016. Citado 3 vezes nas páginas 41, 42, and 43.
- Hermann, M.; Pentek, T.; Otto, B. Design principles for industrie 4.0 scenarios. In: *2016 49th Hawaii International Conference on System Sciences (HICSS)*. [S.l.: s.n.], 2016. p. 3928–3937. Citado na página 18.
- HIBBELER, R. C. *Mechanics of materials*. [S.l.]: Pearson, 2018. Citado 8 vezes nas páginas 9, 27, 28, 29, 30, 31, 32, and 33.
- HJERTAKER, B. T.; JOHANSEN, G. A.; JACKSON, P. Level measurement and control strategies for subsea separators. *Journal of Electronic Imaging*, SPIE, v. 10, n. 3, p. 679 – 689, 2001. Disponível em: <<https://doi.org/10.1117/1.1379771>>. Citado 2 vezes nas páginas 55 and 56.
- HJERTAKER, B. T.; JOHANSEN, G. A.; JACKSON, P. Recent developments in hydrocarbon separator interface imaging. In: MCCANN, H.; SCOTT, D. M. (Ed.). *Process Imaging for Automatic Control*. SPIE, 2001. v. 4188, p. 81 – 92. Disponível em: <<https://doi.org/10.1117/12.417154>>. Citado na página 56.
- Hoffmann Souza, M. L. et al. A survey on decision-making based on system reliability in the context of industry 4.0. *Journal of Manufacturing Systems*, v. 56, p. 133–156, 2020. ISSN 0278-6125. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0278612520300807>>. Citado na página 21.
- Holstad, M. B. et al. Scattered gamma radiation utilized for level measurements in gravitational separators. *IEEE Sensors Journal*, v. 5, n. 2, p. 175–182, 2005. Citado 2 vezes nas páginas 24 and 56.
- HU, L. et al. Review of cyber-physical system architecture. *2012 IEEE 15th International Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing Workshops*, p. 25–30, 2012. Citado na página 19.
- ILIEV, A.; KYURKCHIEV, N.; MARKOV, S. On the approximation of the step function by some sigmoid functions. *Mathematics and Computers in Simulation*, v. 133, p. 223–234, 2017. ISSN 0378-4754. Biomath 2014 and Biomath 2015. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0378475415002554>>. Citado na página 49.
- Jazdi, N. Cyber physical systems in the context of industry 4.0. In: *2014 IEEE International Conference on Automation, Quality and Testing, Robotics*. [S.l.: s.n.], 2014. p. 1–4. Citado na página 18.
- JESCHKE, S. et al. Industrial internet of things and cyber manufacturing systems. In: _____. *Industrial Internet of Things: Cybermanufacturing Systems*. Cham: Springer International Publishing, 2017. p. 3–19. ISBN 978-3-319-42559-7. Disponível em: <https://doi.org/10.1007/978-3-319-42559-7_1>. Citado na página 17.
- KAGERMANN, H. Change through digitization—value creation in the age of industry 4.0. In: _____. *Management of Permanent Change*. Wiesbaden: Springer Fachmedien Wiesbaden, 2015. p. 23–45. ISBN 978-3-658-05014-6. Disponível em: <https://doi.org/10.1007/978-3-658-05014-6_2>. Citado 2 vezes nas páginas 18 and 19.

- KHAN, W. et al. Industrial internet of things: Recent advances, enabling technologies and open challenges. *Computers & Electrical Engineering*, v. 81, p. 106522, 2020. ISSN 0045-7906. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0045790618329550>>. Citado na página 56.
- KIANGALA, K.; WANG, Z. Initiating predictive maintenance for a conveyor motor in a bottling plant using industry 4.0 concepts. *The International Journal of Advanced Manufacturing Technology*, v. 97, p. 1–21, 08 2018. Citado na página 21.
- KRIEGEL, H.-P. et al. Density-based clustering. *WIREs Data Mining and Knowledge Discovery*, v. 1, n. 3, p. 231–240, 2011. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/widm.30>>. Citado 2 vezes nas páginas 43 and 44.
- KUHN, G.; OLIVEIRA, M.; FERNANDES, L. An improved contrast enhancing approach for color-to-grayscale mappings. *The Visual Computer*, v. 24, p. 505–514, 07 2008. Citado na página 62.
- La, H. J.; Kim, S. D. A service-based approach to designing cyber physical systems. In: *2010 IEEE/ACIS 9th International Conference on Computer and Information Science*. [S.l.: s.n.], 2010. p. 895–900. Citado na página 19.
- LANGEVIN, D. et al. Crude oil emulsion properties and their application to heavy oil transportation. *Oil & Gas Science and Technology – Rev. IFP Oil & Gas Science and Technology – Rev. IFP*, v. 59, p. 511–521, 09 2004. Citado 2 vezes nas páginas 56 and 72.
- LASI, H. et al. Industry 4.0. *Business & Information Systems Engineering*, v. 6, p. 239–242, 08 2014. Citado 2 vezes nas páginas 17 and 18.
- LEAL-JUNIOR, A. G. et al. Multi-interface level in oil tanks and applications of optical fiber sensors. *Optical Fiber Technology*, v. 40, p. 82–92, 2018. ISSN 1068-5200. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1068520017305266>>. Citado 2 vezes nas páginas 56 and 72.
- LEE, E. A. Cyber-physical systems - are computing foundations adequate? In: *Position Paper for NSF Workshop On Cyber-Physical Systems: Research Motivation, Techniques and Roadmap*. [s.n.], 2006. Disponível em: <<http://chess.eecs.berkeley.edu/pubs/329.html>>. Citado na página 20.
- LI, B.-r. et al. Framework and case study of cognitive maintenance in industry 4.0. *Frontiers of Information Technology & Electronic Engineering*, v. 20, p. 1493–1504, 11 2019. Citado na página 22.
- LI, C.; GUOYI, X. Research on ultra short wave method to measure water content of crude oil. In: *2008 9th International Conference on Signal Processing*. [S.l.: s.n.], 2008. p. 2956–2959. Citado na página 60.
- LIM, J. et al. A review on the effects of emulsions on flow behaviours and common factors affecting the stability of emulsions. *Journal of Applied Sciences*, v. 15, p. 167–172, 02 2015. Citado na página 47.

- LIU, Q. et al. A visual measurement of water content of crude oil based on image grayscale accumulated value difference. *Sensors (Basel, Switzerland)*, MDPI, v. 19, n. 13, p. 2963, Jul 2019. ISSN 1424-8220. 31284392[pmid]. Disponível em: <<https://doi.org/10.3390/s19132963>>. Citado 3 vezes nas páginas 60, 61, and 67.
- Liu, Y. et al. Review on cyber-physical systems. *IEEE/CAA Journal of Automatica Sinica*, v. 4, n. 1, p. 27–40, 2017. Citado 2 vezes nas páginas 19 and 61.
- LOGAN, D. L.; IMRAN, Q. *A first course in the finite element method*. [S.l.]: Cengage Learning, 2017. 1-30 p. Citado na página 74.
- MADAN, G. *Modern Control System Theory*. [S.l.]: New Delhi, India, Wiley Eastern Limited, 1984. Citado na página 25.
- MAKEEV, Y. V.; LIFANOV, A. P.; SOVLUKOV, A. S. Microwave measurement of water content in flowing crude oil with improved accuracy. In: *2014 24th International Crimean Conference Microwave Telecommunication Technology*. [S.l.: s.n.], 2014. p. 956–957. Citado na página 60.
- MAKEYEV, Y. V.; LIFANOV, A. P.; SOVLOUKOV, A. S. On-line microwave measurement of crude oil water content. In: *2009 19th International Crimean Conference Microwave Telecommunication Technology*. [S.l.: s.n.], 2009. p. 839–840. Citado na página 60.
- MATERIAL, W. *Home*. 2021. Disponível em: <<https://www.theworldmaterial.com/astm-sae-aisi-1020-carbon-steel/>>. Citado na página 75.
- Drilling Automated Realtime Monitoring Using Digital Twin*, Day 2 Tue, November 13, 2018 de *Abu Dhabi International Petroleum Exhibition and Conference*, (Abu Dhabi International Petroleum Exhibition and Conference, Day 2 Tue, November 13, 2018). D021S030R004. Disponível em: <<https://doi.org/10.2118/192807-MS>>. Citado na página 66.
- MODARRES, M.; KAMINSKIY, M.; KRIVTSOV, V. *Reliability engineering and risk analysis: a practical guide*. [S.l.]: CRC Press., 2017. 1-13 p. Citado na página 21.
- MOHRI, M.; ROSTAMIZADEH, A.; TALWALKAR, A. *Foundations of machine learning*. [S.l.]: The MIT Press., 2018. Citado na página 56.
- NEGRI, E.; FUMAGALLI, L.; MACCHI, M. A review of the roles of digital twin in cps-based production systems. *Procedia Manufacturing*, v. 11, p. 939–948, 2017. ISSN 2351-9789. 27th International Conference on Flexible Automation and Intelligent Manufacturing, FAIM2017, 27-30 June 2017, Modena, Italy. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2351978917304067>>. Citado na página 21.
- NORTON, R. L. *Machine design: an integrated approach*. [S.l.]: Prentice Hall, 2011. 303-418 p. Citado 6 vezes nas páginas 30, 35, 36, 37, 79, and 80.
- O'DONOVAN, P. et al. An industrial big data pipeline for data-driven analytics maintenance applications in large-scale smart manufacturing facilities. *Journal of Big Data*, v. 2, p. 25, 11 2015. Citado na página 19.

PARROTT, A.; WARSHAW, L. *Industry 4.0 and the digital twin: Manufacturing meets its match*. [S.l.]: Deloitte University Press, 2017. Citado na página 19.

PETROWIKI. *Oil and gas separators*. PetroWiki, 2021. Disponível em: <https://petrowiki.spe.org/Oil_and_gas_separators>. Citado 2 vezes nas páginas 22 and 77.

POINDEXTER, M. et al. Factors contributing to petroleum foaming. 1. crude oil systems. *Energy & Fuels - ENERG FUEL*, v. 16, 04 2002. Citado na página 48.

PORTER, K.; SIMONS, R.; HARRIS, J. Comparison of three techniques for scour depth measurement: Photogrammetry, echosounder profiling and a calibrated pile. *Coastal Engineering Proceedings*, v. 1, n. 34, p. sediment.64, Oct. 2014. Disponível em: <<https://journals.tdl.org/icce/index.php/icce/article/view/7675>>. Citado na página 18.

QIN, J.; LIU, Y.; GROSVENOR, R. A categorical framework of manufacturing for industry 4.0 and beyond. *Procedia CIRP*, v. 52, p. 173–178, 2016. ISSN 2212-8271. The Sixth International Conference on Changeable, Agile, Reconfigurable and Virtual Production (CARV2016). Disponível em: <<https://www.sciencedirect.com/science/article/pii/S221282711630854X>>. Citado na página 17.

Qiu, T. et al. Edge computing in industrial internet of things: Architecture, advances and challenges. *IEEE Communications Surveys Tutorials*, v. 22, n. 4, p. 2462–2488, 2020. Citado na página 56.

RAO, S. S. *The finite element method in engineering*. [S.l.]: Butterworth-heinemann, an imprint of Elsevier., 2018. 3-52 p. Citado na página 74.

REISCHAUER, G. Industry 4.0 as policy-driven discourse to institutionalize innovation systems in manufacturing. *Technological Forecasting and Social Change*, v. 132, p. 26–33, 2018. ISSN 0040-1625. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0040162517316657>>. Citado na página 18.

RIEGEL, J.; MAYER, W.; HAVRE, Y. van. *FreeCAD Documentation*. 2021. Disponível em: <<https://wiki.freecadweb.org/>>. Citado na página 76.

ROCSOLE. *Solutions For Separators - SEE BEYOND*. Rocsole, 2021. Disponível em: <<https://www.rocsole.com/solutions-for-separators>>. Citado 2 vezes nas páginas 23 and 24.

ROSEN, R. et al. About the importance of autonomy and digital twins for the future of manufacturing. *IFAC-PapersOnLine*, v. 48, n. 3, p. 567–572, 2015. ISSN 2405-8963. 15th IFAC Symposium on Information Control Problems in Manufacturing. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2405896315003808>>. Citado na página 22.

Saldivar, A. A. F. et al. Industry 4.0 with cyber-physical integration: A design and manufacture perspective. In: *2015 21st International Conference on Automation and Computing (ICAC)*. [S.l.: s.n.], 2015. p. 1–6. Citado na página 18.

SCHLEICH, B. et al. Shaping the digital twin for design and production engineering. *CIRP Annals*, v. 66, n. 1, p. 141–144, 2017. ISSN 0007-8506. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0007850617300409>>. Citado 2 vezes nas páginas 20 and 21.

SCHRAMM, L. *Emulsions : fundamentals and applications in the petroleum industry / Laurier L. Schramm, editor*. [S.l.: s.n.], 1992. ISBN 0-8412-2006-9. Citado na página 48.

SCHUBERT, E. et al. Dbscan revisited, revisited: Why and how you should (still) use dbscan. *ACM Trans. Database Syst.*, Association for Computing Machinery, New York, NY, USA, v. 42, n. 3, jul. 2017. ISSN 0362-5915. Disponível em: <<https://doi.org/10.1145/3068335>>. Citado na página 44.

SCHUMACHER, A.; EROL, S.; SIHN, W. A maturity model for assessing industry 4.0 readiness and maturity of manufacturing enterprises. *Procedia CIRP*, v. 52, p. 161–166, 2016. ISSN 2212-8271. The Sixth International Conference on Changeable, Agile, Reconfigurable and Virtual Production (CARV2016). Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2212827116307909>>. Citado na página 18.

SEBER, G. A. F.; WILD, C. J. *Nonlinear regression*. [S.l.]: Wiley-Interscience, 2003. Citado na página 45.

SENGUPTA, D.; KISHORE, P. Continuous liquid level monitoring sensor system using fiber Bragg grating. *Optical Engineering*, SPIE, v. 53, n. 1, p. 1 – 8, 2014. Disponível em: <<https://doi.org/10.1117/1.OE.53.1.017102>>. Citado na página 55.

SHALEV-SHWARTZ, S. *Online learning and online convex optimization*. [S.l.]: Now Publishers, 2012. Citado na página 42.

SHARMA, P.; LAO, L.; FALCONE, G. A microwave cavity resonator sensor for water-in-oil measurements. *Sensors and Actuators B: Chemical*, v. 262, p. 200–210, 2018. ISSN 0925-4005. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0925400518302351>>. Citado na página 60.

SHI, T. et al. Capacitance-based instrumentation for multi-interface level measurement. *Measurement Science and Technology*, IOP Publishing, v. 2, n. 10, p. 923–933, 1991. ISSN 0957-0233. Citado na página 56.

SMYTH, G. K. Nonlinear regression. In: _____. *Encyclopedia of Environmetrics*. American Cancer Society, 2006. ISBN 9780470057339. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470057339.van017>>. Citado na página 45.

SOLUTIONS, F. T. A. et al. *Properties: Polycarbonate (PC) - Conductive Polycarbonate - Properties - Supplier Data by Goodfellow*. 2021. Disponível em: <<https://www.azom.com/properties.aspx?ArticleID=2008>>. Citado na página 75.

SONY, M.; NAIK, S. Industry 4.0 integration with socio-technical systems theory: A systematic review and proposed theoretical model. *Technology in Society*, v. 61, p. 101248, 2020. ISSN 0160-791X. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0160791X19304051>>. Citado na página 22.

STEWART, M.; ARNOLD, K. E. *Surface Production Operations, Volume 1: Design of Oil Handling Systems and Facilities*. [S.l.]: Gulf Professional Publishing, 2011. 1-18, 150-243, 244-310 p. Citado 3 vezes nas páginas 9, 22, and 48.

- SULZER. *The iPhase™ Level Measurement and Water Profiler*. 2021. Disponível em: <<https://www.sulzer.com/en/shared/products/sulzer-iphase>>. Citado na página 23.
- SURESH, S. *Fatigue of material*. [S.l.]: Cambridge University Press, 1991. 221-255 p. Citado 2 vezes nas páginas 39 and 80.
- SÖDERBERG, R. et al. Toward a digital twin for real-time geometry assurance in individualized production. *CIRP Annals*, v. 66, n. 1, p. 137–140, 2017. ISSN 0007-8506. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0007850617300380>>. Citado na página 20.
- TAKAGI, Y. et al. Development of a noncontact liquid level measuring system using image processing. *Water Science and Technology*, v. 37, n. 12, p. 381–387, 1998. ISSN 0273-1223. Instrumentation, Control and Automation of Water and Wastewater Treatment and Transport Systems 1997. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0273122398003564>>. Citado na página 60.
- TAO, F. et al. Digital twin-driven product design, manufacturing and service with big data. *The International Journal of Advanced Manufacturing Technology*, v. 94, 02 2018. Citado na página 19.
- TAO, F. et al. Data-driven smart manufacturing. *Journal of Manufacturing Systems*, v. 48, p. 157–169, 2018. ISSN 0278-6125. Special Issue on Smart Manufacturing. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0278612518300062>>. Citado na página 19.
- TAO, F. et al. Digital twins and cyber–physical systems toward smart manufacturing and industry 4.0: Correlation and comparison. *Engineering*, v. 5, n. 4, p. 653–661, 2019. ISSN 2095-8099. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S209580991830612X>>. Citado na página 19.
- TAO, F. et al. Digital twin-driven product design framework. *International Journal of Production Research*, Taylor & Francis, v. 57, n. 12, p. 3935–3953, 2019. Citado na página 21.
- Tao, F.; Zhang, M. Digital twin shop-floor: A new shop-floor paradigm towards smart manufacturing. *IEEE Access*, v. 5, p. 20418–20427, 2017. Citado 2 vezes nas páginas 20 and 21.
- TAO, F. et al. Digital twin driven prognostics and health management for complex equipment. *CIRP Annals*, v. 67, n. 1, p. 169–172, 2018. ISSN 0007-8506. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0007850618300799>>. Citado na página 21.
- Terrissa, L. S. et al. A new approach of phm as a service in cloud computing. In: *2016 4th IEEE International Colloquium on Information Science and Technology (CiSt)*. [S.l.: s.n.], 2016. p. 610–614. Citado na página 21.
- THE Online Materials Information Resource. 2021. Disponível em: <<http://www.matweb.com/search/DataSheet.aspx?MatGUID=10b74ebc27344380ab16b1b69f1cffbb>>. Citado na página 75.

THE Online Materials Information Resource. 2021. Disponível em: <<http://www.matweb.com/search/DataSheet.aspx?MatGUID=501acbb63cbc4f748faa7490884cdbca&ckck=1>>. Citado na página 75.

TRACERCO. *Profiler™ GPS*. Tracerco, 2019. Disponível em: <<https://www.tracerco.com/products/nucleonic-instrumentation/profiler-gps/>>. Citado 3 vezes nas páginas 9, 23, and 24.

UHLEMANN, T. H.-J.; LEHMANN, C.; STEINHILPER, R. The digital twin: Realizing the cyber-physical production system for industry 4.0. *Procedia CIRP*, v. 61, p. 335–340, 2017. ISSN 2212-8271. The 24th CIRP Conference on Life Cycle Engineering. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2212827116313129>>. Citado na página 21.

Vachálek, J. et al. The digital twin of an industrial production line within the industry 4.0 concept. In: *2017 21st International Conference on Process Control (PC)*. [S.l.: s.n.], 2017. p. 258–262. Citado na página 20.

VIRTANEN, P. et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, v. 17, p. 261–272, 2020. Citado na página 50.

WANG, L.; TÖRNGREN, M.; ONORI, M. Current status and advancement of cyber-physical systems in manufacturing. *Journal of Manufacturing Systems*, v. 37, p. 517–527, 2015. ISSN 0278-6125. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0278612515000400>>. Citado 3 vezes nas páginas 17, 18, and 19.

WANG, S. et al. Implementing smart factory of industrie 4.0: An outlook. *International Journal of Distributed Sensor Networks*, v. 12, n. 1, p. 3159805, 2016. Disponível em: <<https://doi.org/10.1155/2016/3159805>>. Citado 2 vezes nas páginas 18 and 19.

WANG, T.-H. et al. Liquid-level measurement using a single digital camera. *Measurement*, v. 42, n. 4, p. 604–610, 2009. ISSN 0263-2241. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0263224108001711>>. Citado na página 60.

WONG, S.; LIM, J.; DOL, S. Crude oil emulsion: A review on formation, classification and stability of water-in-oil emulsions. *Journal of Petroleum Science and Engineering*, v. 135, p. 498–504, 2015. ISSN 0920-4105. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0920410515301315>>. Citado na página 48.

YANG, W. Q.; BRANT, M. R.; BECK, M. S. A multi-interface level measurement system using a segmented capacitance sensor for oil separators. *Measurement Science and Technology*, IOP Publishing, v. 5, n. 9, p. 1177–1180, sep 1994. Disponível em: <<https://doi.org/10.1088/0957-0233/5/9/021>>. Citado na página 56.

YU, C. Liquid level measurement by using an image method. *International Conference on Signal Processing Proceedings, ICSP*, v. 2015, p. 2320–2323, 01 2015. Citado na página 60.

YU, C.; XU, X.; LU, Y. Computer-integrated manufacturing, cyber-physical systems and cloud manufacturing – concepts and relationships. *Manufacturing Letters*, v. 6, p. 5–9,

2015. ISSN 2213-8463. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2213846315000176>>. Citado na página 19.

ZHANG, H. et al. A digital twin-based approach for designing and multi-objective optimization of hollow glass production line. *IEEE Access*, v. 5, p. 26901–26911, 2017. Citado na página 21.

ZHAO, Y. et al. High-accuracy low-water-content measurement of crude oil based on a near-infrared spectral absorption method. *Optical Engineering*, SPIE, v. 43, n. 10, p. 2216 – 2217, 2004. Disponível em: <<https://doi.org/10.1117/1.1783291>>. Citado na página 60.

ZHONG, R. Y. et al. Intelligent manufacturing in the context of industry 4.0: A review. *Engineering*, v. 3, n. 5, p. 616–630, 2017. ISSN 2095-8099. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2095809917307130>>. Citado na página 19.

Appendix

APPENDIX A – Clustering Pseudocodes

Algorithm 1: K-means clustering (Lloyd's algorithm)

1. **Seeding:** Choose k seed points c_1, c_2, \dots, c_k . Set $J_{old} = \infty$. Compute the current k-means cost J using seed points as centers, i.e.

$$J_{curr} = \sum_{i=1}^n \min_j d^2(x_i, c_j)$$

2. **While** $J_{curr} < J_{old}$

- a) **Voronoi partitioning:** Computing the Voronoi partitioning of the data based on the centers c_1, c_2, \dots, c_k . In other words, create k clusters C_1, C_2, \dots, C_k such that $C_i = x : d(x, c_i) \leq \min_{j \neq i} d(x, c_j)$. Break ties arbitrarily.

- b) **Reseeding:** Compute new centers c'_1, c'_2, \dots, c'_k , where

$$c'_j = \text{mean}(C_j) = \frac{1}{|C_j|} \sum_{x \in C_j} x$$

Set $J_{curr} = J_{old}$. Update the current k-means cost J_{curr} , using the new centers.

- c) **Output:** The current set of centers c_1, c_2, \dots, c_k .
-

Algorithm 2: Hierarchical clustering (single linkage algorithm)

Input: An $n(n-1)$ set of dissimilarities

1. Determine the smallest dissimilarity, d_{ik} .
2. Agglomerate objects i and k : that is, replace them with a new object, $i \cup k$; update dissimilarities such that, for all objects $j \neq i, k$:

$$d_{i \cup k, j} = \min\{d_{ij}, d_{kj}\}$$

Delete dissimilarities d_{ij} and d_{kj} , for all j , as these are no longer used.

3. While at least two objects remain, return to step 1.
-

Algorithm 3: Spectral clustering

Input: Similarity matrix \mathbf{S} , number of clusters K

1. Transform \mathbf{S} Calculate $d_i \leftarrow \sum_{j=2}^n S_{i,j}$ the node degrees, for $i = 1 : n$. Form the transition matrix \mathbf{P} with $P_{i,j} \leftarrow S_{i,j}/d_i$, for $i, j = 1 : n$
2. Eigendecomposition Compute the largest K eigenvalues $\lambda_1 \geq \dots \geq \lambda_K$ and eigenvectors $\mathbf{v}^1, \dots, \mathbf{v}^K$ of \mathbf{P} .
3. Embed the data in K -th principal subspace Let $\mathbf{x}_i = [\mathbf{v}_i^2 \ \mathbf{v}_i^3 \ \dots \ \mathbf{v}_i^K] \in \mathbb{R}^{K-1}$, for $i = 1, \dots, n$.
4. Run the K -means algorithm on the "data" $\mathbf{x}_{1:n}$.

Result: Output the clustering C in step 4.

Algorithm 4: DBSCAN clustering

Input: DB : Database**Input:** ϵ : Radius**Input:** $minPts$: Density threshold**Input:** $dist$: Distance function**Data:** $label$: Point labels, initially undefined**foreach** $point\ p$ **in** database DB **do do** **if** $label(p) \neq undefined$ **then** | **continue** **end** Neighbors $N \leftarrow RangeQuery(DB, dist, p, \epsilon)$ **if** $|N| < minPts$ **then** | $label(p) \leftarrow Noise$ **continue** **end** $c \leftarrow$ next cluster label $label(p) \leftarrow c$ Seed set $S \leftarrow N \setminus p$ **foreach** q **in** S **do do** **if** $label(q) = Noise$ **then** | $label(q) \leftarrow c$ **end** **if** $label(q) \neq undefined$ **then** | **continue** **end** Neighbors $N \leftarrow RangeQuery(DB, dist, p, \epsilon)$ $label(p) \leftarrow c$ **if** $|N| < minPts$ **then** | **continue** **end** $S \leftarrow S \cup N$ **end****end**

Annex

ANNEX A – Experimental Separator Vessel's CAD Design Extractions

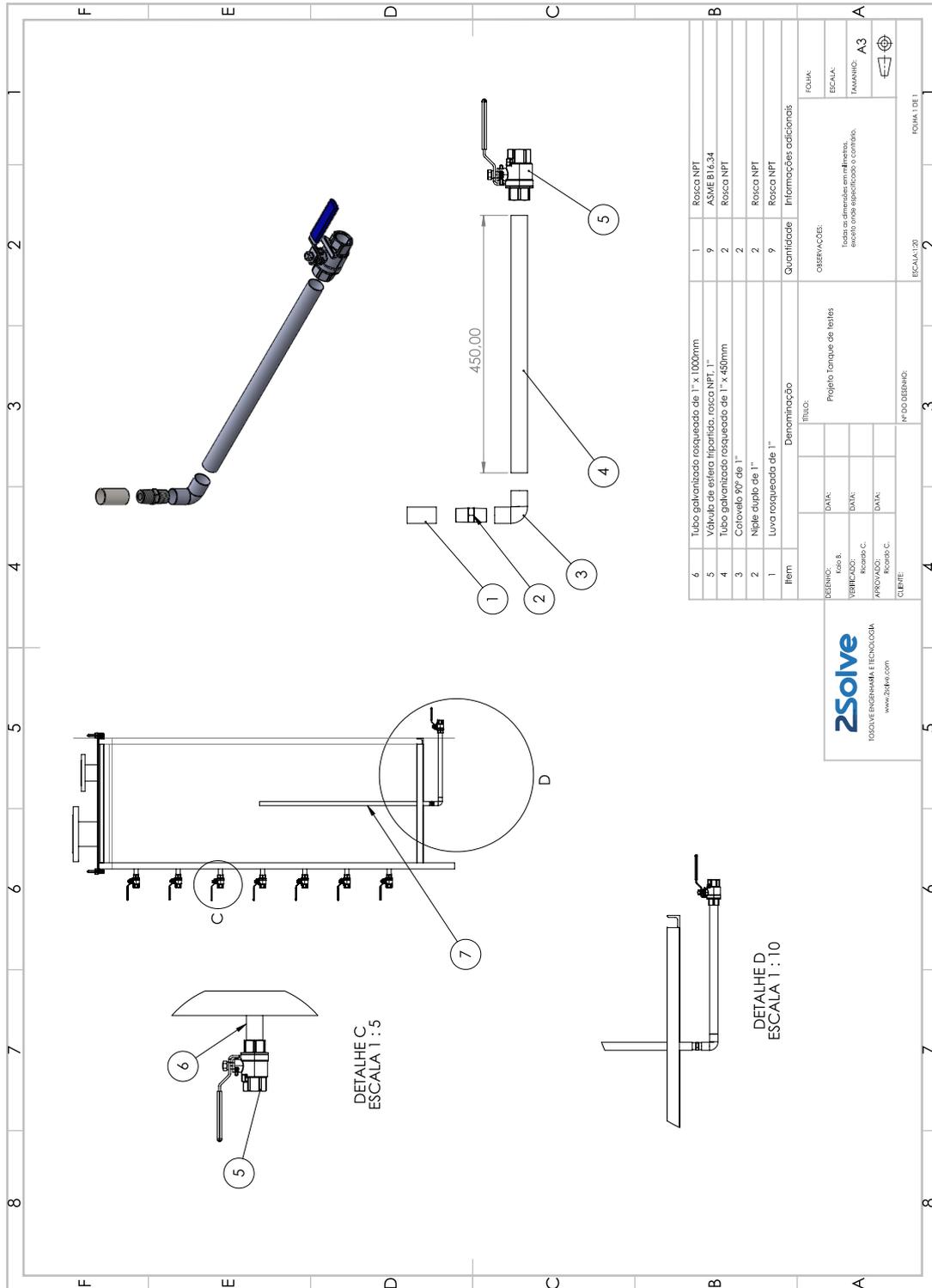


Figure A.0.1 – Extraction from CAD design of the experimental separator vessel (page 2) (2SOLVE, 2019).

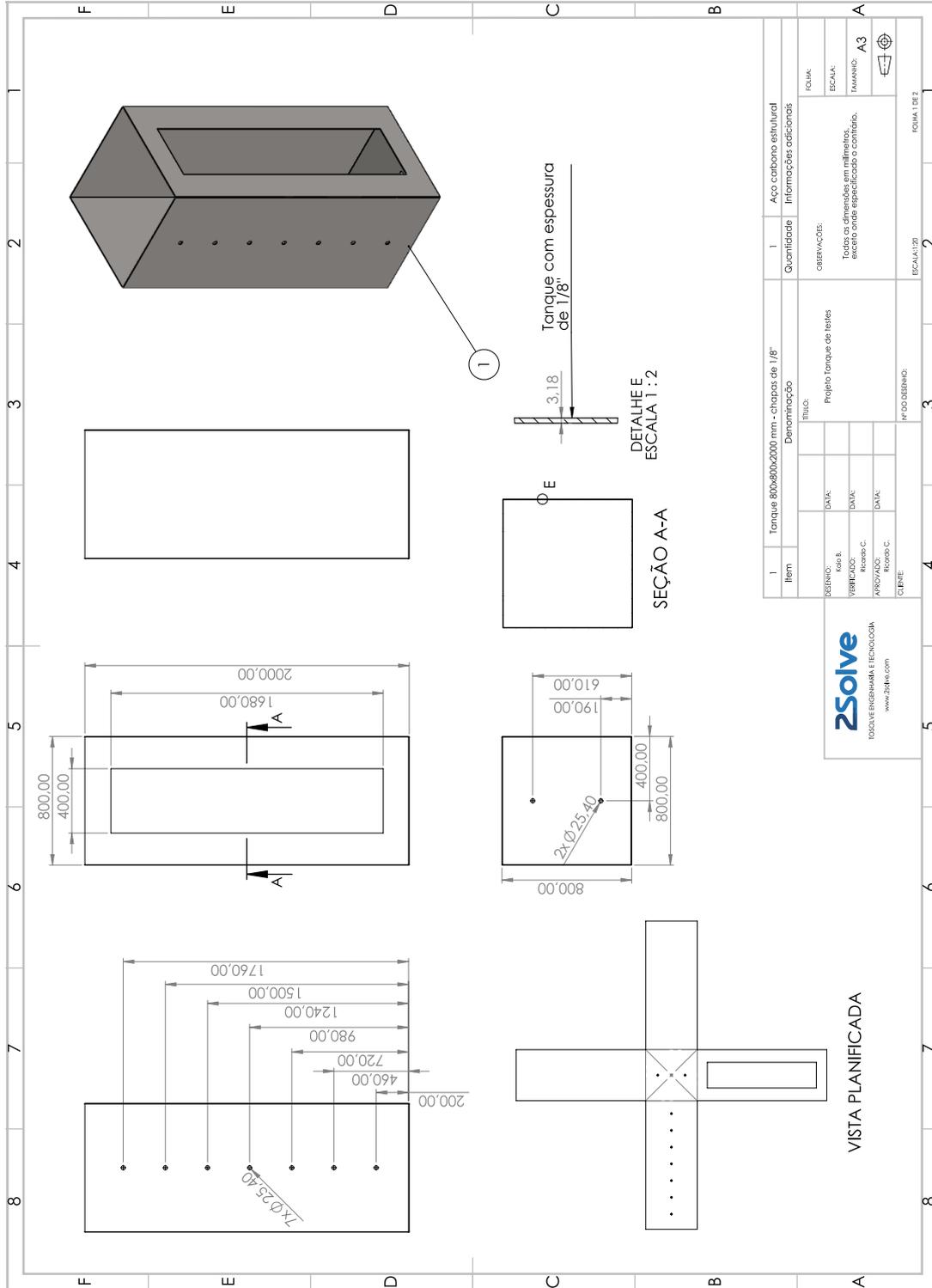


Figure A.0.2 – Extraction from CAD design of the experimental separator vessel (page 5) (2SOLVE, 2019).