ASTES

Advances in Science, Technology & Engineering Systems Journal

VOLUME 8-ISSUE 5|SEPT-OCT 2023

www.astesj.com
ISSN: 2415-6698

Editorial

In this issue, we present a collection of 09 accepted papers encompassing a diverse range of topics in the fields of technology, engineering, and education. These papers contribute to the advancement of knowledge and innovation within their respective domains. Each paper is a result of rigorous research, analysis, and experimentation, offering valuable insights and potential avenues for future exploration.

The first paper explores the integration of the Common Alerting Protocol (CAP) into the existing ISDB-T standard for digital terrestrial television systems. The researchers aim to enhance early warning capabilities for natural disasters by utilizing the ISDB-T standard alongside the Emergency Warning Broadcasting System (EWBS). The study involves the development of a CAP to EWBS translator and the design of transmitters compliant with both full-seg and one-seg ISDB-T standards [1].

Moving into the realm of robotics and education, the second paper discusses the implementation of augmented reality (AR) technology for training purposes in the context of the "Thailand 4.0" roadmap. The study introduces an AR-based visual programming interface for robot training, allowing users to understand the interconnected relationship between data and hardware functionality. The proposed AR-based training system proves effective in human resource development for the robotics and automation sector [2].

The third paper addresses the crucial issue of longitudinal stability in fixed-wing UAVs. Unmanned aerial vehicles (UAVs) are increasingly employed for various applications, and ensuring their stability is paramount. The research focuses on the design and modeling of a feedback controller, including PI, PID, and Fuzzy logic controllers. Results indicate the superiority of the fuzzy logic controller in managing the system's response [3].

Shifting towards sustainable energy solutions, the fourth paper investigates the impact of wind power integration on energy networks. The study highlights wind energy as a competitive and eco-friendly resource but acknowledges its unpredictable nature. The research uses simulation in MATLAB/SIMULINK to analyze the effects of wind power integration on the power quality of distribution networks. Additionally, the paper proposes the integration of Flexible Alternating Current Transmission Systems (FACTS) as a solution to maintain clean electrical quality [4].

The fifth paper provides a comprehensive review of the condition assessment of medium voltage assets in power distribution networks. Focusing on assets such as power transformers, photovoltaic systems, switchgear, lines and cables, and instrument transformers, the study identifies research gaps and emphasizes the need for deeper assessment in critical areas. The review covers traditional diagnostic methods and advanced AI-based approaches [5].

In the realm of education, the sixth paper introduces the Graded Multidisciplinary Model (GMM) as a means to facilitate STEM/STEAM education. The GMM is applied to the topic of logic gates, linking university content to high school content through practical activities in Minecraft. The study demonstrates the positive impact of the GMM on students' performance and skills in science, technology, engineering, art, and mathematics [6].

Shifting gears to social media networks, the seventh paper delves into the modeling of control agents using reinforcement learning. With a focus on opinion dynamics, the research proposes the use of intelligent reinforcement learning agents to shape opinions in social media networks. The paper presents a multi-agent system and Q-learning approach to control opinion dynamics effectively [7].

The eighth paper addresses the critical need for battery cell balancing in the context of renewable energy and electric vehicles. Introducing a novel near-field coupling method, the study focuses on a resonant coil design with an enhanced Q-factor for efficient balancing. Experimental validation demonstrates the proposed technique's superiority in speed and wireless energy transfer to cells, overcoming the limitations of traditional passive and active methods [8].

The ninth paper discusses the design and prototyping of a 3DOF worm-drive robot arm. The affordability and simplicity of the revolute arm make it suitable for executing pick-and-place tasks. The paper details the structural elements, kinematics, and operational metrics of the arm, highlighting its capabilities in endpoint position tracking and object manipulation [9].

In conclusion, this issue brings together a diverse array of research papers, showcasing advancements in early warning systems, robotics training, UAV stability, renewable energy integration, asset condition assessment, STEM/STEAM education, opinion dynamics, battery cell balancing, and robot arm design. The collective efforts of the researchers contribute significantly to the academic and practical understanding of these fields, paving the way for further exploration and innovation in the future.

References:

[1] N.B.B. Cifuentes, G.F.O. Cifuentes, “Transmission of the CAP Protocol through the ISDB-T Standard,” Advances in Science, Technology and Engineering Systems Journal, **8**(5), 1–7, 2023, doi:10.25046/aj080501.

[2] L. Yutthanakorn, S. Charoenseang, “Augmented Reality Based Visual Programming of Robot Training for Educational Demonstration Site,” Advances in Science, Technology and Engineering Systems Journal, **8**(5), 8–16, 2023, doi:10.25046/aj080502.

[3] V. Phunpeng, W. Wanna, S. Khaengkarn, T. Kerdphol, “Feedback Controller for Longitudinal Stability of Cessna182 Fixed-Wing UAVs,” Advances in Science, Technology and Engineering Systems Journal, **8**(5), 17–27, 2023, doi:10.25046/aj080503.

[4] A. Slimani, B. Merah, M.N. Tandjaoui, C. Benachaiba, “The Effect of Introduction of Wind Energy System on the Energy Networks,” Advances in Science, Technology and Engineering Systems Journal, **8**(5), 28–34, 2023, doi:10.25046/aj080504.

[5] E.G. Mesino, J. Caicedo, M. Mamaní, D.R. Quete, A.C. Piamba, D.G. Gómez, G.A. Mayor, J.C. Erazo, W.M. López, E. Jay, A.R. Quete, “Condition Assessment of Medium Voltage Assets: A Review,” Advances in Science, Technology and Engineering Systems Journal, **8**(5), 35–54, 2023, doi:10.25046/aj080505.

[6] M. Flores-Nicolás, M. Martínez-Reyes, F. de J. Matías-Torres, “The Graded Multidisciplinary Model: Fostering Instructional Design for Activity Development in STEM/STEAM Education,” Advances in Science, Technology and Engineering Systems Journal, **8**(5), 55–61, 2023, doi:10.25046/aj080506.

[7] M.N. Zareer, R. Selmic, “Modeling Control Agents in Social Media Networks Using Reinforcement Learning,” Advances in Science, Technology and Engineering Systems Journal, **8**(5), 62–69, 2023, doi:10.25046/aj080507.

[8] J. Jeon, D. Lee, “Resonance Coil Design for a Novel Battery Cell Balancing with using Near-Field Coupling,” Advances in Science, Technology and Engineering Systems Journal, **8**(5), 70–76, 2023, doi:10.25046/aj080508.

[9] I.S. Howard, “Design and Prototyping of a 3DOF Worm-drive Robot Arm,” Advances in Science, Technology and Engineering Systems Journal, **8**(5), 77–93, 2023, doi:10.25046/aj080509.

Editor-in-chief

Prof. Passerini Kazmersk

ADVANCES IN SCIENCE, TECHNOLOGY AND ENGINEERING SYSTEMS JOURNAL

Volume 8 Issue 5 **September-October 2023**

CONTENTS

Advances in Science, Technology and Engineering Systems Journal Vol. 8, No. 5, 1-7 (2023)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Transmission of the CAP Protocol through the ISDB-T Standard

Nelson Bolívar Benavides Cifuentes, Gonzalo Fernando Olmedo Cifuentes*

Universidad de las Fuerzas Armadas ESPE, Department of Electrical, Electroinics and Telecomunications, Sangolquí, 171103, Ecuador

ARTICLE INFO

Article history:
Received:11 April, 2023
Accepted:02 August, 2023
Online: 24 September, 2023

Keywords:
ISDB-T
CAP
Adalm Pluto
Transport Stream
DSM-CC
Emergency System

ABSTRACT

Early warning systems have had a significant impact on society by providing timely information to mitigate the effects of natural disasters. To enhance early warning capabilities, researchers are exploring the use of digital terrestrial television systems to broadcast alerts across large urban and rural areas. In this research project, the aim is to integrate the global early warning protocol CAP (Common Alerting Protocol) into the existing ISDB-T standard, alongside the standard's Emergency Warning Broadcasting System (EWBS). This integration will enable the creation of a hybrid system, allowing various global emergency alert devices that utilize the CAP protocol to be activated through the Digital Television signal. To achieve this, a CAP to EWBS translator was developed as part of the design proposal prior to transmission. Additionally, transmitters compliant with both the full-seg and one-seg ISDB-T standards were designed to support the CAP protocol. These transmitters utilize SDR (Software-Defined Radio) cards of the Adalm Pluto type. The CAP protocol, encoded in XML format, was transmitted through the ISDB-T transport stream using the DSM-CC data transmission protocol. By incorporating the CAP protocol into the ISDB-T standard and utilizing the DSM-CC data transmission protocol, this research project aims to enhance the early warning capabilities of digital terrestrial television systems.

1. Introduction

Traditionally, Early Warning Systems (EWS) have been implemented to alert the population about potential natural or man-made phenomena and mitigate their impacts [1-2]. To achieve this, various systems have adopted global protocols or developed their own protocols for emergency risk reduction [3-7].

In Ecuador, the ISDB-T International standard is being used for digital terrestrial television implementation. This standard, originating from Japan with Brazilian modifications, is widely adopted in South America [8-9]. The International ISDB-T standard includes an Early Warning Broadcasting System (EWBS) that utilizes area codes to deliver emergency information at the regional level, allowing for targeted alerts within specific areas [10-12].

However, relying solely on the EWBS system within the International ISDB-T standard poses limitations in terms of receiver compatibility and the replication of emergency alerts. As a result, there is a proposition to integrate global alert protocols for emergencies. One widely adopted international protocol is the Common Alerting Protocol (CAP), which interfaces with diverse technologies such as telephony, radio, fax, emails, and websites [13]. Hence, the objective of this article is to incorporate the CAP protocol into the International ISDB-T standard, thereby enabling hybrid digital terrestrial television standards to swiftly and effortlessly transmit emergency messages on a global scale.

The CAP protocol offers an open and non-proprietary digital message format for all types of alerts, utilizing XML for programming, which leads to reduced costs and operational complexity by providing a unified software interface for numerous existing warning systems [14].

The TV 3.0 project is a regional initiative led by Brazil that aims to develop and enhance the television transmission standard to provide viewers with a more advanced and enriched experience. It began in July 2020 and has been divided into several phases to achieve its objectives [15].

*Corresponding Author: Gonzalo Fernando Olmedo Cifuentes, Email: gfolmedo@espe.edu.ec

The first phase of the project involved inviting organizations from around the world to submit proposals for the six components of the TV 3.0 system: over-the-air physical layer, transport layer, video coding, audio coding, subtitles, and application coding [16]. A total of 36 responses were received from 21 different organizations. Similar proposals were merged to simplify testing and evaluations in the next phase [17-23].

The second phase took place during the years 2021 and 2022, focusing on comprehensive testing and evaluations of the proposed technologies and components for the TV 3.0 system. Currently, the third phase of the project is underway and is expected to last approximately two and a half years. A final decision on the over-the-air physical layer technology is anticipated in April 2024, with corresponding technical specifications being drafted from May to August 2024. Concurrently, research and development (R&D) activities are being conducted for the transport layer and application coding, including the development of adaptations, extensions, tools, and test suites. These activities will take place from March 2023 to August 2024. The official launch of the TV 3.0 system is expected to occur in 2025, offering viewers an enhanced television experience and a wide range of services and functionalities.

Within the TV 3.0 project, the implementation of an Advanced Emergency Alert System (EAS) has been identified as a crucial component. The EAS is designed to provide efficient and effective emergency alert notifications to the public through the TV 3.0 platform. It utilizes the Advanced Emergency Alerting (AEA) system based on ATSC 3.0, which is structured in an XML document called the Advanced Emergency Alerting Table (AEAT). This system enables broadcasters to transmit detailed emergency information, including text messages, images, and multimedia content, to inform and adequately prepare the public during emergency situations. Decision trees and protocols have been established to handle different types of alerts, including required alerts, optional alerts, and emergency action notifications, based on priority levels defined in the AEAT [22].

During the transition to digital television, there are still countries in the region that have not completed this process but have digital transmitters that can be utilized to send early emergency alerts through the broadcasting signal. This includes rural areas where television signals are available, as well as locations with Internet access. The research group has been working on projects related to the transmission and reception of emergency alerts using the EWBS system, which is part of the adopted ISDB-T standard in the country [24].

Additionally, the CAP protocol is a global protocol that is part of ATSC 3.0 and will be integrated into the TV 3.0 system. This protocol is used in other global communication mediums as well. Hence, there is a need to demonstrate that through the ISDB-T transport stream structure, it is possible to transmit any type of data, but this requires the implementation of specialized receivers. In this context, this article proposes a method to include the CAP protocol in the ISDB-T system, which is within the standard of the transport layer, and demonstrates its functionality in conjunction with the physical layer of the ISDB-T receiver according to the ISDB-T standard, which requires a special module. To achieve this, the design and implementation using software-defined radio are proposed.

This proposal and innovation aim to expand the capabilities of the ISDB-T system to enable the transmission of emergency alerts using the CAP protocol. By demonstrating the feasibility of this integration, it is expected to improve the efficiency and effectiveness of emergency alert systems in the context of digital television, contributing to the safety and well-being of the population.

This article is structured as follows: Section 2 outlines the methodology employed for information encryption and decryption. Section 3 presents the results obtained from the study, and Section 4 provides the concluding remarks.

2. Methodology

The aim of this study is to incorporate the CAP protocol into a Transport Stream (TS) flow and design a test scenario utilizing Software-Defined Radio (SDR) in both the transmission and reception stages.

To insert the CAP protocol in a TS stream, it is proposed to use the DSM-CC (Digital Storage Media Command and Control) standard. This standard allows the CAP Protocol to be transmitted cyclically using files, directories, services, and event streams through the data and object carousels [25-26].

For the test scenario in the transmission and reception stages, the cost-effective SDR Adalm Pluto will be utilized. This compact and easily transportable device serves various functions in the realms of education and research. The SDR Adalm Pluto operates in conjunction with the GNU Radio software, which is an open-source program compatible with multiple platforms. It allows for modular programming and simplifies the creation of pre-programmed Python blocks for signal transmission or reception [27].

2.1. TS Design

To use the DSM-CC in the International ISDB-T standard, the AIT table must be added to the PSI/SI tables; this table provides complete information about data transmission, the required activation status of the applications carried by it, etc. The data in the AIT table allows the transmitter to request that the receiver change the activation state of an application [28-32].

Within the PSI/SI Tables is the PMT table, which contains all the elements that make up the TS stream, such as audio, video, data, and the AIT table in the Elementary Stream field.

In each Elementary Stream field, some descriptors help to complement the description of each element of the TS stream, which is why the following descriptors are included in the AIT table: the Data Components descriptor and the Application signaling descriptor. These descriptors will indicate the type of application to be used within the TS flow. Its conformation is shown in Figure 1.

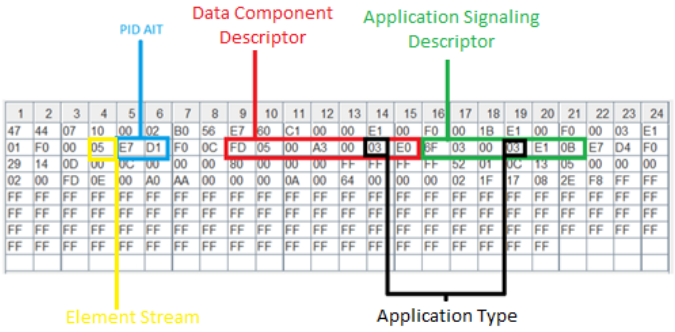


Figure 1: AIT Table Descriptors within the PMT Table

The descriptors for the DSM-CC section element are the TAG Association descriptor, the Identifier descriptor, and the Carousel Identifier descriptor. These descriptors mention the use of the data carousel within the TS stream. Its structure is presented in Figure 2.

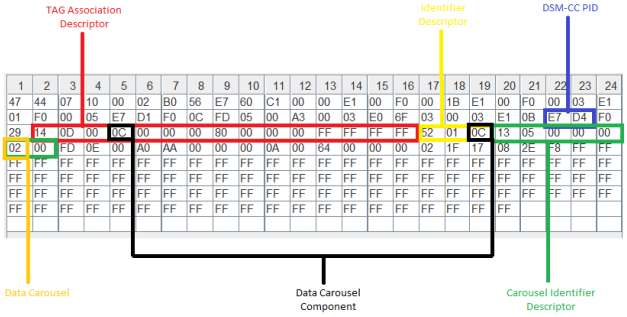


Figure 2: DSM-CC descriptors within the PMT table.

Next, the information that must go in the AIT table is presented, where the type of application is inserted, which must be the same as that shown in the PMT table; in turn, three descriptors are inserted, the first being the protocol descriptor transport indicating the use of data carousel. Second, the application descriptor suggests the priority of the data carousel in the TS flow. Finally, there is the application name descriptor, where the name of the CAP file is inserted in the ASCII code. These values can be seen in Figure 3.

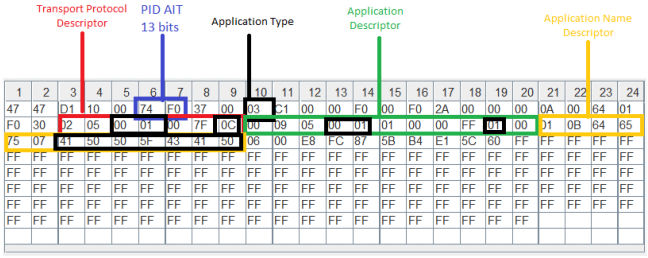


Figure 3: Components of the AIT table.

The TS flow and the DSM-CC section containing the multiplexed CAP protocol will be obtained with an audio and video signal.

2.2. Design of the Transmission and Reception System

With the generated TS stream, we build the transmission system in the GNU Radio software. The transmission system diagram with the GNU Radio software is presented in Figures 4, 5, and 6, and the description of each component is established in Table 1.

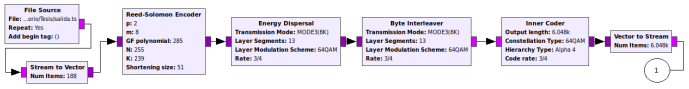


Figure 4: Transmission System in GNU Radio - Coding

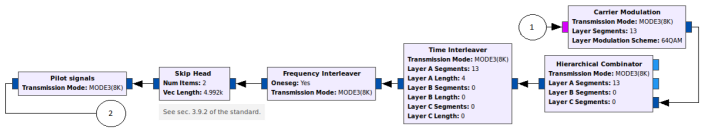


Figure 5: Transmission System in GNU Radio - Modulation

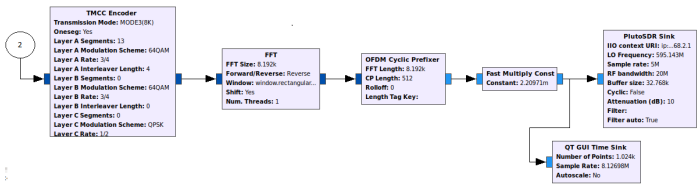


Figure 6: Transmission System in GNU Radio - Configuration

Table 1: Description of each Block of the Transmission System

Block	Description
FILE SOURCE	Read a stream from a file.
STREAM TO VECTOR	Converts an item stream to a GNU Radio block stream.
REED SOLOMON ENCODER	Create a Reed Solomon scrambler according to the standard.
ENERGY DISPERSAL	Implements an energy disperser compliant with the standard.
BYTE INTERLEAVER	Inserts a byte interleaver according to the standard.
INNER CODER	Create an internal encoder with puncturing.
VECTOR TO STREAM	Converts the GNURadio block stream to an item stream.
CARRIER MODULATION	Modulates the carrier with the parameters of Mode, number of segments, and modulation scheme.
HIERARCHICAL COMBINATOR	Divide the content to be transmitted into hierarchical layers.
TIME INTERLEAVER	Implementation of a Fornet interleaver as specified by the standard.
FREQUENCY INTERLEAVER	Performs frequency interleaving.

Block	Description
SKIP HEAD	Remove information considered junk from the header.
PILOT SIGNALS	Establishes pilot signals by the standard.
TMCC ENCODER	Block that encodes the TMCC carriers.
FFT	Implements the fast Fourier transform.
OFDM CYCLIC PREFIXER	Adds a cyclic prefix and performs pulse shaping on OFDM symbols.
FAST MULTIPLY CONST	Multiply the input by a constant to have the desired size.
PLUTO SDR SINK	Block that transmits the signal through the Adalm Pluto.
QT GUI TIME SINK	Graphically presents the transmission.

With these blocks, the SDR Adalm Pluto is connected, and for this software to recognize it, the IP address: 192.168.2.1 is put in the PLUTO SDR SINK block.

For the receiving system, a new program is created in GNU Radio. The reception system diagram is shown in Figures 7, 8, and 9, and the description of each component is established in Table 2.

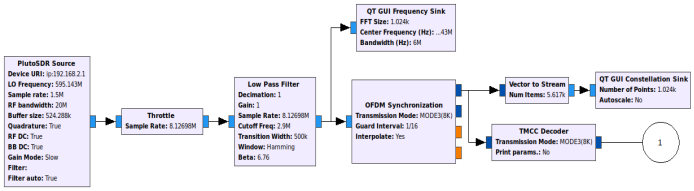


Figure 7: Reception System in GNU Radio – Configuration

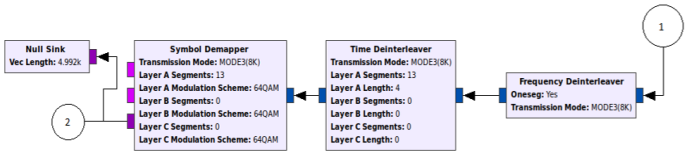


Figure 8: Reception System in GNU Radio - Demodulation

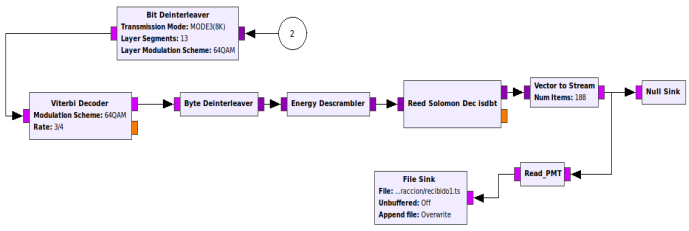


Figure 9: Reception System in GNU Radio - Decoding

Table 2: Description of each Block of the Reception System.

BLOCK	***DESCRIPTION***
PLUTO SDR SOURCE	Receive the signal from Adalm Pluto.
THROTTLE	Indicates the sample rate to use.
LOW PASS FILTER	Create a low pass filter to get the signal from the TS file.
QT GUI FREQUENCY SINK	Graphically indicates the shape of the signal.
OFDM SYNCHRONIZATION	Specifies OFDM information such as Mode, Guard Interval, and interpolation.
QT GUI CONSTEALLATION SINK	Graphically indicates the constellation of the system.
TMCC DECODER	Decodes TMCC carriers.
FREQUENCY DEINTERLEACER	Performs frequency interleaving.
TIME DEINTERLEAVER	Specifies a Forney interleaver as specified by the standard.
SYMBOL DEMAPPER	Removes the symbol map on each carrier and then generates a serial keyword.
BIT DEINTERLEAVER	Performs a bit of interleaving.
VITERBI DECODER	Implements a Viterbi decoder on the signal.
BYTE DEINTERLEAVER	Implements a Forney byte deinterleaver as specified by the standard.
ENERGY DESCRAMBLER	Specifies a power decoder as specified by the standard.
REED SOLOMON DEC ISDBT	Receives 204-byte blocks and performs a Reed Solomon decode.
VECTOR TO STREAM	Transforms the GNU Radio blocks to an elementary stream of 188 bytes for each packet.

BLOCK	*DESCRIPTION*
READ PMT	Block was created to verify the PMT table information mentioning its Elementary Stream types.
FILE SINK	Create a file with the information received.

With the received file, the information from the DSM-CC section is extracted as follows:

1. Look up the PID from the PMT Table and then identify the PID from the DSM-CC section.
2. Compare the PID frame by frame and verify that it matches the PID of the DSM-CC section. When a frame is found with the DSM-CC PID, its content is analyzed.
3. Identify the Table id to indicate if it is a data or control block.
4. Extract the content stored in a TS file as a data block.
5. Being a control block, the type of block is identified.
6. Finally, the number of data and control blocks found is obtained.

Subsequently, the data carousel is extracted, which allows removing the headers of the DBB sections that contain the data. To perform this function, the following steps are followed:

1. The number of modules that were generated in the extraction of the DSM-CC is verified.
2. Each module is analyzed to check if the file still contains information.
3. The message_length segment located in the byte following the table id is analyzed from each frame. This segment will indicate the size of the CAP file.
4. This process is repeated until the beginning and end of the message are stored.

Then, the object carousel is extracted; in this section, the type of file inside the TS is verified, this; case, it will only be a file type file where the CAP file will be obtained in XML format. The steps are the following:

1. The file size generated in the previous function is obtained.
2. The type of file is obtained.
3. Its header is parsed to extract its content.
4. At the end, you will have the extracted content.

3. Results

To verify the operation of the CAP protocol transmission, the TS flow is created with the information from the PAT, PMT, NIT, SDT, AIT, and DSM-CC tables.

Once all these elements have been verified, the signal is transmitted and received with the help of the GNU Radio software and SDR Adalm Pluto, the scenario mentioned in Figure 10.

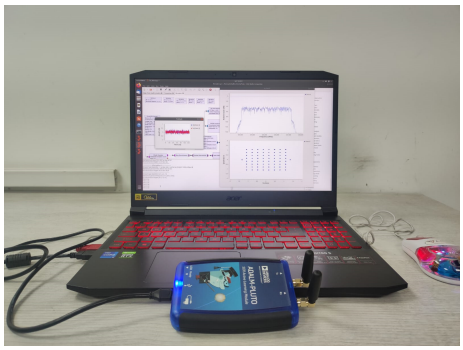


Figure 10: Scenario with GNU Radio and Adalm Pluto.

To check the operation of the communications system, a block called Read PMT was created at the reception; this block will read the information from the PMT and its four elementary Streams: the audio, the video, the AIT table, and the DSM-CC section. This content is indicated in Figure 11.

```
--------------------------------------------------
PMT
--------------------------------------------------
PID: 1031 y  in hex: 0x407
Section length: 86 y  in hex: 0x56
Section syntax error: 1 entonces True
CRC-32: [ 23  8  46 248]
Program number: 59232 y en hex: 0xe760
Version Number: 0 y en hex: 0x0
Current next indicator:  1 entonces True
PCR_PID: 256 y  in hex: 0x100
Program information length: 0 y  in hex: 0x0
--------------------------------------------------
---------- Information 1 ----------
Stream Type identifier: 27  in hex: 0x1b : Video ITU Recommendation H.264 e ISO/IEC 14496-10
Elementary Stream PID: 256  in hex 0x100
ES information length: 0  in hex: 0x0
--------------------------------------------------
---------- Information 2 ----------
Stream Type identifier: 3  in hex: 0x3 : Audio  ISO/EC 11172-3
Elementary Stream PID: 257  in hex 0x101
ES information length: 0  in hex: 0x0
--------------------------------------------------
---------- Information 3 ----------
Stream Type identifier: 5  in hex: 0x5 : Seccion
Elementary Stream PID: 2001  in hex 0x7d1
ES information length: 12  in hex: 0xc
--------------------------------------------------
---------- Information 4 ----------
Stream Type identifier: 11  in hex: 0xb :  ISO/IEC 13818-6 (tipo B)
Elementary Stream PID: 2004  in hex 0x7d4
ES information length: 41  in hex: 0x29
--------------------------------------------------
Stream type identifier: [27, 3, 5, 11]
Elementary Stream PID: [256, 257, 2001, 2004]
ES information length: [0, 0, 12, 41]
```

Figure 11: Content of the PMT Table at the reception

After some time, receiving and transmitting are paused, and the received.ts file is obtained. The steps for removing the DSM-CC section using a Python file are followed to extract the information from this received file, getting Figure 12.

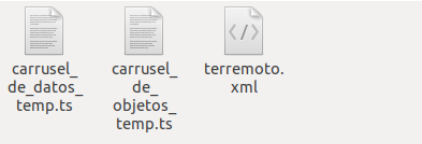


Figure 12: File Extracted from TS Stream.

In the end, the terremoto.xml file is successfully extracted. The cap-validator software checks the CAP file, resulting in a valid file, as shown in Figure 13 [33].

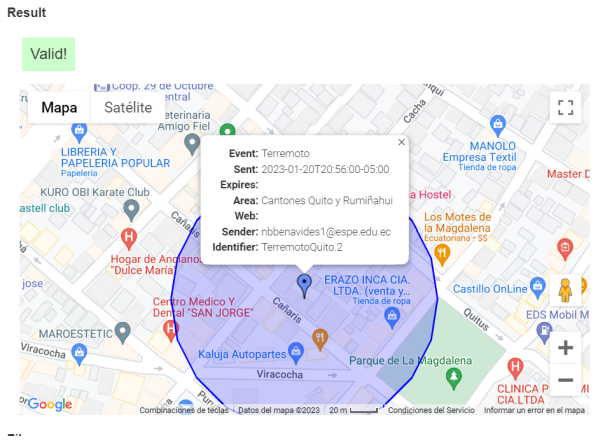


Figure 13. Result of cap-validator software.

The resulting process and the used diagram are shown in the QR code below in Figure 14.

Figure 14: QR code of CAP protocol transmission and extraction display.

4. Conclusions

The present study successfully demonstrated the transmission and reception of CAP protocol data in XML format through the DSM-CC data structure within an ISDB-T transport stream. The utilization of the CAP protocol resulted in the generation of minimal bytes in the DSM-CC, showcasing its efficiency in terms of bandwidth utilization when compared to transmitting an interactive application under the same transmission mode.

By employing GNU Radio and the SDR Adalm Pluto, a robust communication system for ISDB-T with full-seg content was developed. The system exhibited flawless reproduction of audio and video reception, while ensuring the complete download of the CAP file. Furthermore, the implementation of a one-seg reception system using the SDR Adalm Pluto and the DEKTEC DTU 215 modulating card yielded successful results, thanks to the minimal bandwidth requirements of the transmitted CAP file.

To ensure the quality of the CAP file, both during its construction and upon reception in both one-seg and full-seg scenarios, the cap-validator software was employed. This validation process confirmed the accuracy and integrity of the CAP information. Overall, the communications system demonstrated optimal performance and is well-suited to serve as a valuable asset in supporting the EWBS emergency system. By integrating the CAP system into the ISDB-T transport layer platform, emergency authorities and broadcasters can send timely and accurate emergency alerts, ensuring the safety and well-being of the population. The CAP structure, which is both part of ATSC 3.0 and will be part of TV 3.0, aims to provide a powerful tool for public authorities to disseminate news, weather updates, and critical information during emergencies, enhancing the overall emergency response capabilities within the broadcasting industry. Through the proposed approach in this article, the integration of CAP can even be achieved during the transition to digital television using the ISDB-T system, allowing for readiness in any migration to an advanced digital television system in the near future.

References

[1] F. Guzzetti, S. Luigi Gariano, S. Peruccacci, M. Brunetti, I. Marchesini, M. Rossi, M. Melillo, Geographical landslide early warning systems, Earth-Science Reviews, **200**, 2020. ISSN 0012-8252, https://doi.org/10.1016/j.earscirev.2019.102973.

[2] A. J. L. García and R. C. Rojas, "Abstract Extended of Novel Method for broadcasting Multiples Emergency Alert and Alarm Using ISDB-T Digital Terrestrial Television," 2019 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Jeju, Korea (South), 2019, 1-3, doi: 10.1109/BMSB47279.2019.8971946.

[3] G. Olmedo, F. Acosta, D. Villamarín, F. Santander, R. Achig, V. Morocho (2021). Prototype of a Centralized Alert and Emergency System for Digital Terrestrial Television in Ecuador, Advances in Intelligent Systems and Computing, 1326. Springer, Cham. https://doi.org/10.1007/978-3-030-68080-0_14.

[4] S. NURBADRI, K. ANWAR, D. ARSENO, Study On Early Warning Systems (EWS) for Indonesia Digital Terrestrial Television. JMECS (Journal of Measurements, Electronics, Communications, and Systems), [S.l.], **6**(1), 11-19, june 2020. ISSN 2477-7986

[5] Li, Z., Fang, L., Sun, X. et al. 5G IoT-based geohazard monitoring and early warning system and its application. J Wireless Com Network **2021**, 160, 2021. https://doi.org/10.1186/s13638-021-02033-y.

[6] G. Olmedo, F. Acosta, R. Haro, D. Villamarín, N. Benavides. (2019). Broadcast Testing of Emergency Alert System for Digital Terrestrial Television EWBS in Ecuador, Communications in Computer and Information Science, 1004. Springer, Cham. https://doi.org/10.1007/978-3-030-23862-9_13.

[7]. A. J. L. García and R. C. Rojas, "Abstract Extended of Novel Method for broadcasting Multiples Emergency Alert and Alarm Using ISDB-T Digital Terrestrial Television," 2019 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Jeju, Korea (South), 2019, 1-3, doi: 10.1109/BMSB47279.2019.8971946.

[8] M. Takada and M. Saito, "Transmission System for ISDB-T," in Proceedings of the IEEE, **94**(1), 251-256, Jan. 2006, doi: 10.1109/JPROC.2005.859692.

[9] L. C. d. P. Costa, C. S. Kurashima, M. G. De Biase, R. H. Alonso and M. K. Zuffo, "A technical analysis of digital television broadcasting in brazil," 2013 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), London, UK, 2013, 1-5, doi: 10.1109/BMSB.2013.6621795.

[10] ISDB-T HARMONIZATION DOCUMENT PART 3: Emergency Warning Broadcast System (EWBS) (November/2015) by ISDB-T International Forum.

[11] ARIB STD-B24:2008, Data coding and transmission specification for digital broadcasting.

[12] ARIB TR-B14:2006, Operational guidelines for digital terrestrial television broadcasting.

[13] P. K. Dalela, A. Saldhi, P. Bhave and V. Tyagi, "Common Alerting Protocol Compliant Emergency Warning And Alert System for Legacy Broadcasting Networks," 2020 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA), Victoria, BC, Canada, 2020, 1-5, doi: 10.1109/CogSIMA49017.2020.9216105.

[14] ITU-T, ITU-T Rec. X.1303 bis (03/2014) Common alerting protocol (CAP 1.2), 2014.

[15] FORUM ISDB-T, "TV3.0 Project", 2020, https://forumsbtvd.org.br/tv3_0/.

[16] FORUM ISDB-T, "Call for Porposals: TV3.0 Project", 17 July 2020, https://forumsbtvd.org.br/wp-content/uploads/2020/07/SBTVDTV-3-0-CfP.pdf.

[17] FORUM ISDB-T, "Testing and Evaluation Report: TV 3.0 Project - Over-the-air Physical Layer Laboratory Tests", 03 December 2021, https://forumsbtvd.org.br/wp-content/uploads/2021/12/SBTVD-TV_3_0-PL-Lab-Report.pdf.

[18] FORUM ISDB-T, "Testing and Evaluation Report: TV 3.0 Project - Over-the-air Physical Layer Field Tests", 03 December 2021, https://forumsbtvd.org.br/wp-content/uploads/2021/12/SBTVD-TV_3_0-PL-Field-Report.pdf.

[19] FORUM ISDB-T, "Testing and Evaluation Report: TV 3.0 Project - Transport Layer", 03 December 2021, https://forumsbtvd.org.br/wp-content/uploads/2021/12/SBTVD-TV_3_0-TL-Report.pdf.
[20] FORUM ISDB-T, "Testing and Evaluation Report: TV 3.0 Project - Video Coding", 03 December 2021, https://forumsbtvd.org.br/wp-content/uploads/2021/12/SBTVD-TV_3_0-VC-Report.pdf.
[21] FORUM ISDB-T, "Testing and Evaluation Report: TV 3.0 Project - Audio Coding", 03 December 2021, https://forumsbtvd.org.br/wp-content/uploads/2021/12/SBTVD-TV_3_0-AC-Report.pdf.
[22] FORUM ISDB-T, "Testing and Evaluation Report: TV 3.0 Project - Captions", 03 December 2021, https://forumsbtvd.org.br/wp-content/uploads/2021/12/SBTVD-TV_3_0-CC-Report.pdf.
[23] FORUM ISDB-T, "Testing and Evaluation Report: TV 3.0 Project - Application Coding", 03 December 2021, https://forumsbtvd.org.br/wp-content/uploads/2021/12/SBTVD-TV_3_0-AP-Report.pdf.
[24] G. Olmedo and Roman Lara-Cueva, "Evaluation of Emergency Warning Broadcast System (EWBS) for Digital Terrestrial Television in Ecuador", IEEE Global Communications Newsletter, December 2021, https://gcn.comsoc.org/evaluation-emergency-warning-broadcast-system-ewbs-digital-terrestrial-television-ecuador.
[25] Hongguang Zhang, Tianpu Jiang, Zhiqi Gu and Shibao Zheng, "Design and implementation of broadcast file system based on DSM-CC data carousel protocol," in IEEE Transactions on Consumer Electronics, **50**(3), 929-933, Aug. 2004, doi: 10.1109/TCE.2004.1341702.
[26] G. Olmedo, A. Nunez and D. Villamarín, "Design, implementation and evaluation of data carrousel extractor algorithm on MPEG-2 TS for digital terrestrial television," 2016 XLII Latin American Computing Conference (CLEI), Valparaiso, Chile, 2016, 1-8, doi: 10.1109/CLEI.2016.7833325.
[27] GNU RADIO, GNU Radio The Free Software & Open Software Radio Ecosystem, 2021.
[28] RTS/JTC-DVB-371, TS 102 809 - V1.3.1 - Digital Video Broadcasting (DVB); Signalling and carriage of interactive applications and services in Hybrid broadcast/broadband environments, 2017.
[29] Associação Brasileira de Normas Técnicas, ABNT NBR 15602-3: Codificação de vídeo, áudio e multiplexação Parte 3 Sistemas de multiplexação de sinais, 2007.
[30] Associação Brasileira de Normas Técnicas, ABNT NBR 15603-1: Multiplexación y servicios de información (SI) Parte 1 SI del sistema de radiodifusión, 2007.
[31] Associação Brasileira de Normas Técnicas, ABNT NBR 15603-2: Multiplexación y servicios de información (SI) Parte 2 Estructura de datos y definiciones de la información básica de SI, 2007.
[32] Associação Brasileira de Normas Técnicas, ABNT NBR 15603-3: Multiplexación y servicios de información (SI) Parte 3Sintaxis y Definiciones de Información extendida del SI, 2007.
[33] Google, CAP-VALIDATOR, 2014.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 8, No. 5, 8-16 (2023)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

xASTESJ

ISSN: 2415-6698

Augmented Reality Based Visual Programming of Robot Training for Educational Demonstration Site

Lucksawan Yutthanakorn, Siam Charoenseang*

Institute of Field Robotics, King Mongkut's University of Technology Thonburi, Bangkok, 10140, Thailand

ARTICLE INFO

Article history:
Received:27 July, 2023
Accepted:02 September, 2023
Online: 24 September, 2023

Keywords:
Augmented Reality
Educational Demonstration Site
Smart Factory
Robot Training System
Visual Block-based Programming

ABSTRACT

The human resource development of robotics and automation in the smart factory is an important factor in "Thailand 4.0" roadmap, which is following the industry 4.0 model. To pursue this goal of Thailand 4.0 roadmap of labor development, the effective and intuitive training system must be easy to understand. This study proposes the implementation of augmented reality (AR) technology for training purposes due to its ability to visualize real-time invisible data, such as device status. This involves the development of the AR-based visual programming interface with an educational demonstration site (demo site). The AR-based training system is started with animation content explaining the smart factory concept, followed by hands-on learning using Microsoft HoloLens 2 and IoT hardware devices in demo site. The demo site using an MQTT protocol, simulates an automated packing line in the smart factory. The hardware status is published in real-time to the MQTT broker. This approach enables users to comprehend the interconnected relationship between data and hardware functionality and allowing them to create their own programs to control the IoT hardware in the demo site. With this training, 22 targeted users have successfully grasped the smart factory concept and its correlation with hardware functionality. The block-based visual programming employed in the system enables easy comprehension of robot commands. Moreover, the AR application provides a smooth display at 48-60 frames per second. The proposed system's usability and value for specific tasks received high scores, ranging from 4 to 5 points, confirming its effectiveness for the targeted users. The proposed AR-based training system has proved that it has benefit for human resource development in the robotics and automation sector following the Thailand 4.0 roadmap. The proposed system empowered users to understand the smart factory concept and its practical implementation, leading to create new ideas for integrating AR and smart factory concepts into their manufacturing in the future.

1. Introduction

The industry 4.0 is the era of using the cyber-physical system [1]-[2] to manage data in manufacturing, which could be called as the smart factory. By transforming production into a digital process and connecting factories to the internet, industry 4.0 can enhance manufacturing performance. Germany is the first nation that mentions about the Industrie 4.0. The German government announces as the national strategic initiative [3] and establishes the platform for industry development to 4.0 and the criteria for evaluating the potential of the factory's deve lopment. The United State of America (USA) has an important strategy called the Smart Manufacturing. The strategy aims to manage the investments in manufacturing. The National Institute of Standards and Technology (NIST) of USA supports about implementing the robotics in manufacturing and creates the tests for evaluating potential of the automation in the production line. The quality control is the key to modern factory [4]. In Southeast Asia, Singapore prepares the Smart Industry Readiness Index (SIRI) for measuring industrial development readiness. The method emphasizes three dimensions are the process, the technology, and the organization. SIRI shares the framework for developing the industry to industry 4.0.

In Thailand, there is a model for developing the industry called Thailand 4.0.The model aims to help small and medium-sized enterprises (SMEs) become high-potential enterprises. Thailand 4.0 aims to implement the robotics and automation in the production line and labor skill improvement [5]. The Thailand 4.0 strategy identifies five new S-Curve industries: robotics, aviation and logistics, digital, biofuels and biochemicals, and medical hub.

*Corresponding Author: S. Charoenseang, KMUTT, Bangkok, 10140, Thailand, siam.cha@kmutt.ac.th

[6]. Labors need to improve their skills to support these industries. 70% of factories in Thailand are in 2.0-3.0 stage of industry [7]. The labors and entrepreneurs should prepare for the digital disruption and industry 4.0. These five new targeted industries need high-potential workers and the experts to drive them forward.

Although training in the industry is popular, there are few courses that offer hands-on workshops with hardware or demonstration sites. However, the industrial training system can benefit from the implementation of augmented reality (AR) technology. An AR-based application provides the visualization of data from real IoT equipment installed in an educational demonstration site (demo site). The demo site simulates the working of an automated packing line, demonstrating how the smart components work such as robot arm for pick-and-place jobs or sensors for detecting palettes. Furthermore, the AR-based application of the proposed system provides a visual block-based programming panel that supports beginner user in creating their own programs. The proposed solution, which combines the AR technology and the demo site, aims to increase worker's understands of the fundamentals of the smart factory concept and improve their robotics and automation skills.

2. Related Works

2.1. Comparison of AR, VR, and MR concepts

In the present, the augmented reality, virtual reality, and mixed reality are implemented in real life with various purposes, such as medical training application, assembly assistance task, augmented reality with data visualization in production line [8] and using these technologies for training system.

The experts explain the concept of augmented reality (AR) as a merging of 3D graphics with the real world [9]. There are three parts of AR-based system which are user, digital world, and real world. User can control the device in real world by interacting with augmented graphics appearing in the head-mount display device such as Microsoft HoloLens.

Virtual reality (VR) is created by all 2D or 3D graphic environment. User does not see any parts of real world. 2D or 3D graphic and virtual environments need a specific device to present the contents. User can see the virtual contents by using headset with head tracker and hand controller to interact with these graphics.

There is no specific clear definition for mixed reality (MR) that can explain the differences between MR and AR. However, the continuum of the popular source defines MR as the technology between the fully real world and fully virtual world [10]. In the other words, AR is a type of MR. Both AR and MR have augmented objects attached to the real environment, but MR may have more complex features such as attaching graphic textures with real characters from the real world, which is called Augmented Virtually – AV. The summary characteristics of AR, VR, and MR are shown in Table 1.

The proposed system aims to display status data in real time while hardware in the demo site is working. The user should be able to see the variation of the data while also seeing the real environment. Therefore, augmented reality technology would be more appropriate for this system than virtual reality.

Table 1: AR, VR, MR Characteristics Comparison

Topic	AR	VR	MR
Can see real environment through the device.	✓	×	Only specific area.
Environment [11]	Real world	Virtual world	Mixed
Can interact with virtual object	✓	✓	✓
Can interact with physical object [12]	✓	×	Only specific object.

2.2. Augmented Reality in Training and Industry

AR is frequently used in daily life, such as in the camera filters of social media applications, image tracking of 2D markers for modern-day advertising, and training systems in the medical field and industry.

Research on augmented reality technology for training has found that AR-based training systems are most commonly used in industry (35%), vocational training (26%), and medical (13%) applications [13]. Another research in training topics using AR or VR technology showed the successful implementation about AR-based training system in maintenance and assembly task [14]. The traditional way to learn the welding process is to practice with real equipment. However, virtual and augmented reality can now be used to create simulation training systems that are more flexible and cost-effective than the traditional methods. A study that compared two groups of students who used VR and AR for training and practical application found that the group that used the simulation training system had lower operating stress and higher levels of focus than the other group [15].

Furthermore, AR technology is not only for training system, but this technology can also be used for data visualization in CNC machines. The data, which can be inspected in real time by integrating AR technology with digital twins, is helpful for making the best decisions in the future [16].

2.3. Thailand 4.0, Thailand Professional Qualification Institute, and the survey about the smart factory training topics in Thailand.

Thailand is now a developing country. A roadmap to drive Thailand to be a developed country in 2032 is called Thailand 4.0 [17]-[18]. One of the agenda in the roadmap is about labor improvement.

There are 5 new s-curves industries including robotics for industry, aviation and logistics, biofuels and biochemicals, digital, and medical hub. To develop and re-skill labors in industry field requires a necessary skill about robotics and automation for digital transformation and industry 4.0 revolution.

The Thailand Professional Qualification Institute – TPQI creates career-standard certifications for qualified individuals. The

robotics and automation standard is under the robotics cluster in Mechatronics standard including 19 professional qualifications, 66 credits. This is a standard for 5 career levels, which are for operator, technician, engineer or staff, manager, system design engineer, and system integrator [19].

A review of training sessions showed that the curriculum emphasizes automation in industry 3.0, but there is not enough information about smart factories. Most training courses did not offer hands-on hardware practice. There are 3 of the 16 topics that contribute to the demo site for students, as shown in Table 2.

Table 2: Survey of Training Topics

Organization	Area of content	Providing the demo site for hands-on practice
Thai-German Dual Education and E-Learning Development Institute	Wireless connection, IoT, Programming	×
	Database	×
	Remote control, Sensor and actuator, Simulation, Programming	×
	Product management, Vision system, Robot and automation, Sensor and actuator, Programming	×
Technology Promotion Association (Thailand-Japan)	IoT, Sensor and actuator, Programming	×
Thailand Productivity Institute	AR/VR, IoT, Barcode scanner and actuator	×
	Data, Cyber security	×
	Cloud computing	×
	Robot and automation	×
	Robot and automation, Programming	×
Thai-German Institute (TGI)	Database	×
	Database, sensor and actuator	×
	Database, Wireless connection	✓
	AI	✓
	Wireless connection, IoT, Programming	✓
	Wireless connection, Sensor and actuator	×

Developing human resources for interpreters and engineers about the smart factory contents is beneficial for the company. This is the opportunity to develop an appropriate training system for this group using new technology and augmented reality to create a new training experience in industry. The benefits of AR technology, such as data visualization and real-world viewing, can help users easily understand the relationship between device operation and data. Additionally, user's understanding of the smart factory concept can be improved by watching the video animations provided in this proposed system.

3. Proposed System

The Thailand 4.0 roadmap requires human development to support the robotics and automation industry which is one of five new S-curves industries. To achieve this, reskilling and upskilling efforts in robotics, particularly in smart factory operations, are also required.

This proposed system is an AR-based training system for smart factory concept. It implements augmented reality (AR) technology to create an effective and intuitive training system. The system scenario while using the AR-based application of this proposed system is shown in Figure 1.

The contents of this system include three animation videos that cover the basics of smart factory operations and robot arm programming. The AR application includes content about the automation devices used in smart factories. The demonstration site (demo site) of smart factory simulates how the automation devices in a smart factory line work and are controllable. The AR contents will be displayed on the Microsoft HoloLens 2 headset.

3.1. System Scenario

The proposed system provides an AR-based application that displays the augmented UI panels, AR user interface, on the Microsoft HoloLens 2 headset. Users can view these panels above the demo site, and they are divided into two parts: the left panel and the right panel.

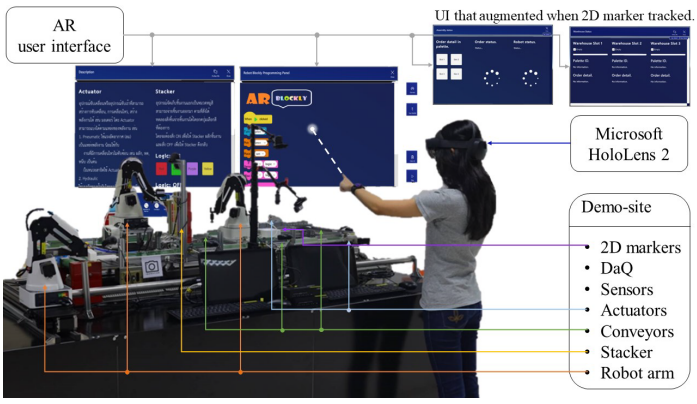


Figure 1: System Scenario

The left panel is the knowledge panel describing the device description, which includes four devices' contents explaining about the sensors used in the demo site and other popular sensors such as proximity sensors. Secondly, the actuator content provides details about various types of actuators used in the demo site. The third content is a content about conveyor which displays the type of conveyors used in the demonstration site. The final content is the robot arm content, which describes different types of robots and how they work with the demo site. User can switch the contents by clicking the button of each section and control the actuator, the conveyor, and the robot arm by interacting with the

buttons that display on each page. The real-time status of each sensor is displayed in the sensor content.

The right panel is for the block-based programming area. Each block is prepared for a specific command, such as controlling robots and conveyors, or if-else logic. To generate a command block, the user needs to click on the desired block displayed on the left side of the panel. The generated block will then appear on the right corner. User needs to drag each command block and snap them together to create an action command. The play button displayed next to this panel is for sending command to the demo site. After the commands are sent, the demo site will execute them step-by-step.

The module's status panel appeared after the 2D marker is tracked. The status includes the module's operational status. For the assembly module, the status panel provides information about operating states such as waiting, working, and so on. The status on the warehouse's panel consists of slot availability. The status of each module is displayed to users in real time.

3.2. System Diagram

Based on Figure 2: System Diagram, the status messages (msg) of each device at the physical level can be transferred to the local broker using the Message Queuing Telemetry Transport (MQTT) protocol. The Node-RED API receives the status and transfers it to the application. When the 2D marker is tracked, the user interface (UI) displays an augmented status panel over the 2D marker.

Once the user finishes programming, the command (cmd) is published from the application. The Node-RED API decodes the command at the API level and distributes it to the devices accordingly. The control of the device is based on the topic that was set earlier.

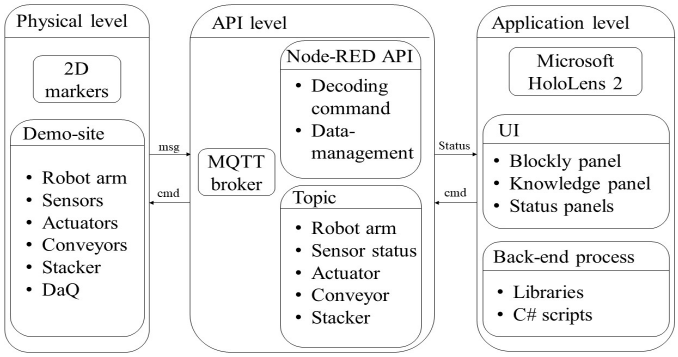


Figure 2: System Diagram

The Physical Level

The physical level consists of five components: MQTT broker, devices, IoT devices, Node-RED execution, and a 2D marker for tracking. The devices and IoT devices implemented in the demo site include robot arms, sensors, actuators, conveyors, stacker, and data acquisition system (DaQ). These physical devices in the real world can be controlled by the user through the application. 2D markers are attached in the demonstration site for providing module information to user when the markers are tracked.

The API Level

There are various libraries used for developing the API level (Application Program Interface), back-end management, and data publication through the broker. The application subscribes to the MQTT topics to receive and send the payload, and a broker publishes a command to that topic for the device to execute at the physical level.

The Application Level

This proposed system was developed using the Unity game engine to create the user interface for the application designed for the Microsoft HoloLens 2. Users can control the demo site by interacting with the user interfaces that appear in the Microsoft HoloLens 2. The user interfaces include the main programming panel (blockly), which provides a block-based programming area, and the knowledge panel, which displays information about the devices.

3.3. Application State Diagram

In Figure 3, the application starts in the initial state and then moves to the connecting state. At this state, the application, the local broker, and the demo site establish a connection. If the connection fails, the Wi-Fi connection in the setting window on the Microsoft HoloLens 2 needs to be checked. Once the connection is successful, the user can interact with the user interfaces displayed on the device to create commands that control the demo site. The application includes a decoding API that converts long commands into step-by-step tasks for the demo site to execute. Finally, the demo site completes the tasks according to the received command.

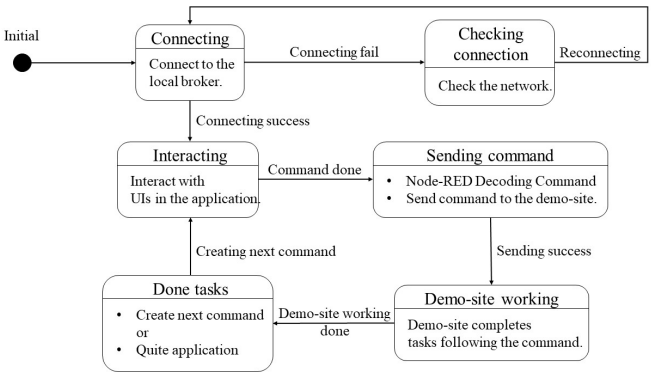


Figure 3: Application State Diagram

3.4. Data Flow

Data flow diagram of the proposed system is shown in Figure 4. There are two parts in the data flow. The first part starts with number 1 and is about the user interface and tracking system. The second part starts with number 2, and includes the MQTT connection and data from the demo site.

Starting at the time before using this system, the application and the demo site need to be connected to the local broker (process 2.0). The connection status shown in the graphic that is augmented on the Microsoft HoloLens 2 display.

The user interface shown on the device is managed by the UI control unit. This unit manages the position of the panels, the position of block command in blockly panel. The snapping function is a part of this unit and is the function for managing the blockly interface. The augmented status panel of the 2D markers developed by using the Vuforia Engine library to define the

markers. The UI control unit provides the position of these augmented graphics.

After user publishes their command, the command management system, which includes Node-RED decoding, receives the command and puts it in order. The first command is sent to the demo site and the next command is sent to the demo site after the first command is finished. While the demo site is working, the status data will be updated in real time and shown in the user interface.

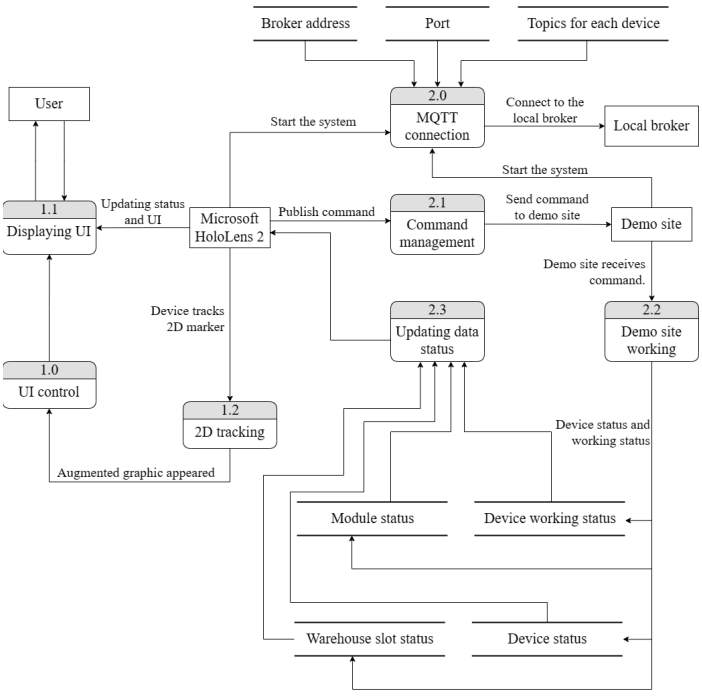


Figure 4: Data Flow Diagram

3.5. Data Synchronization

The device's status data is published to the MQTT local broker and updated in real time. The information such as sensor status (detected or no detected) or module's working status (working, waiting, done) is exchanged under the specific topic. The command order is sent after the previous order is completed. Therefore, the device's working status data and module's status data must be synchronized between the demo site, API, and the AR application.

3.6. Platforms (Hardware and Software)

Microsoft HoloLens 2

Figure 5 shows the Microsoft HoloLens 2, a head-mounted display for augmented reality applications. The see-through holographic display has a 3:2 aspect ratio and a 2K resolution. It is powered by a Qualcomm Snapdragon 850 computing platform with 4 GB of RAM and 64 GB of storage. It supports WIFI and Bluetooth 5 connections. A single battery charge can last for two to three hours. The device has various sensors including a gyroscope for head tracking, two infrared cameras for eye autofocusing, a 1-megapixel time-of-light depth sensor, light cameras, an accelerometer, a magnetometer, and an 8-megapixel camera for capturing 1080p30 video. It also has built-in spatial sound speakers [20].

Figure 5: Microsoft HoloLens 2 [21]

Hand gestures can be used to control the Microsoft HoloLens 2. The device can track the index finger and extend a pointer from it towards the augmented object displayed on the screen. To point, the user lifts the left or right hand up and open the palm as shown in Figure 6 (left) and Figure 7 (left). The dash line graphic will appear from the index finger to the target augmented object. The index finger can then be attached to the thumb to drag the graphic to a desired location, or released to click or choose, as shown in Figure 6 (right). Another method for clicking is to push the index finger through the interactive button. The device can detect which graphic is selected, as shown in Figure 7 (right).

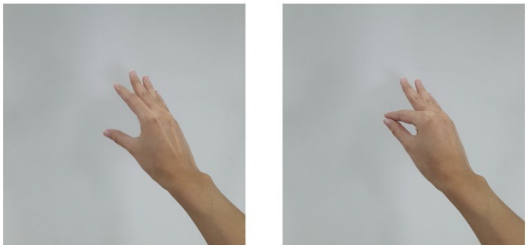


Figure 6: Pointing gesture (left), dragging gesture (right)

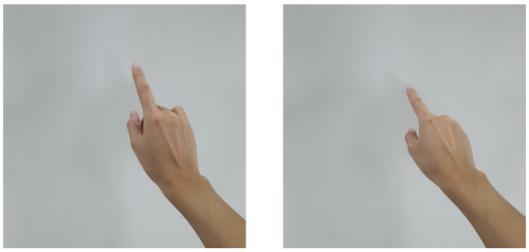


Figure 7: Pointing gesture (left), clicking gesture (right)

Microsoft has recommended size for the holographic buttons depends on the distance between user and the target holographic. The shorter distance (45 cm.) uses direct hand interaction which user can use the fingertip to interact the holographic directly. However, this proposed system chooses the hand-ray interaction instead since the ray that augmented on the fingertip is easy to point to the target holographic for the beginner. The target size should be 3.5 x 3.5 cm. for hand ray interaction. The designed distance is around 0.8-1 m. Then, the minimum designed target size of the button is 3.2 x 3.2 cm. Hence, user can interact with the button comfortably.

2D Tracking

The library for 2D tracking system is Vuforia Engine 10.8 which supports the Unity game engine to track the 2D markers using the computer vision based image recognition [22]. This library was developed for AR application [23]. Developers can upload their 2D markers to the Vuforia's database, with a minimum width of 320 pixels and a maximum size of 2.25 MB for

each file [24]. The Vuforia Engine library supports many devices, including the Microsoft HoloLens 2. Once the marker is tracked, the augmented object will appear either above the marker or in a specific position.

Demonstration Site

The demo site prepared for this proposed system is a collaborative project with the educational smart factory platform, which is a research project of the human-computer interface (HCI) laboratory of the Institute of Field Robotics (FIBO), King Mongkut's University of Technology Thonburi (KMUTT), Thailand.

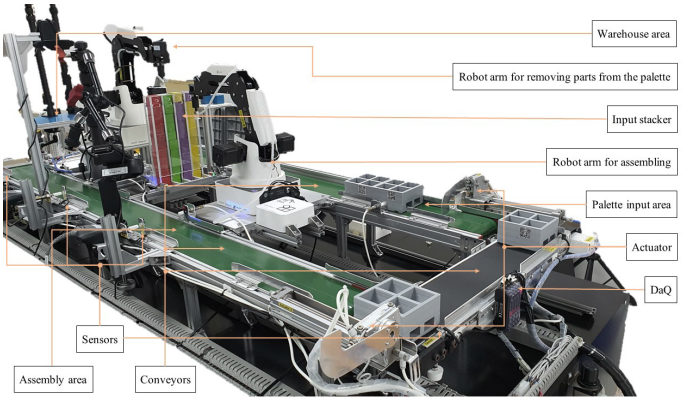


Figure 8: Devices in the Demonstration Site

This demo site simulates the assembly line and operates in separated modules, as shown in Figure 8. The proposed system focuses on the assembly module, the DaQ module consisting of sensors and actuators, and the conveyor module. The hardware includes robot arms, conveyors, sensors, actuators, and input stacker. The communication between the modules in the demo site utilizes the MQTT protocol.

3.7. AR-based Training System

The main contribution of this proposed system is to develop an AR-based application for smart factory concept training system. The system consists of two parts: an AR application including the API, and the contents of smart factory concept.

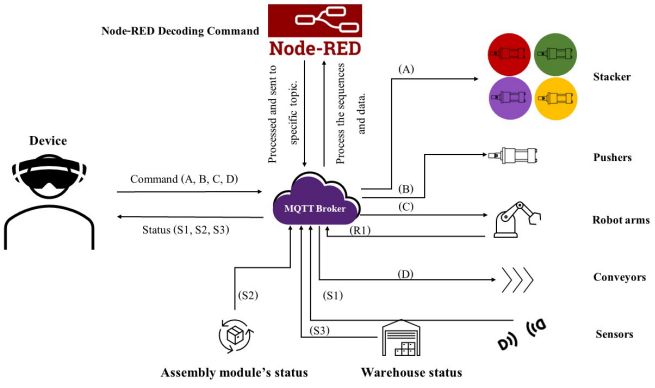


Figure 9: MQTT Data Flow Diagram

Communication between the AR application and devices.

The MQTT protocol is utilized for communication between the AR application and the hardware of the demo site. MQTT broker is configured as the local broker. The scenario of the communication is shown in Figure 9. Each device has its own topics for receiving commands (S1, S2, S3) and updating their operational status (A, B, C, D).

The node-red decoding commands manage the received commands and send them in the correct order to devices, such as stackers, pushers, robot arms, conveyors, and sensors. All statuses from the ended devices will be published to the broker and displayed on the augmented status panels in the Microsoft HoloLens 2.

Node-RED Decoding Command

Node-RED is an online service block-based programming tool [25]. It can create an API for devices and online service using MQTT protocol.

The main API for managing and sending commands to devices in the correct order is the node-red decoding command. The flow chart diagram is shown in Figure 10. The purple block stands for the communication using MQTT protocol.

Starting with the initial state, the sensor and actuator in the palette module will be started. The command (cmd) from the application is sent to MQTT broker in form of string. After the node-red decoding command gets this payload, it will decode command string to command list. The first command set on the index 0 (cmd[0]) of the list will be sent to the device.

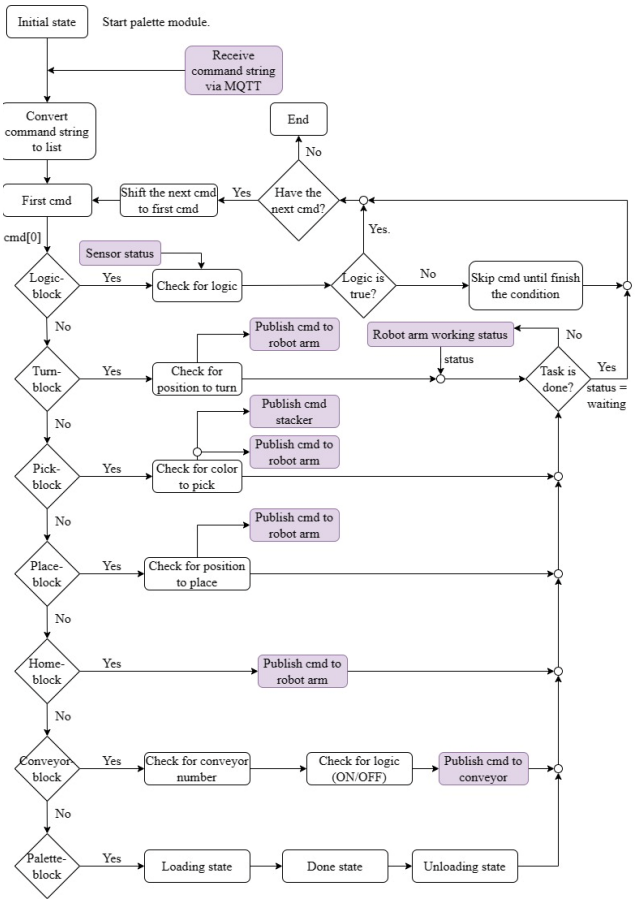


Figure 10: Node-red Decoding Command Flow Chart

The commands such as turning, picking, placing, and homing are related to the robot arm's operations. The working status of the

robot needs to be recognized in real time to prevent the mis-order tasks. The conveyor block controls whether the specific conveyor is on or off.

The logic block will check the sensor status and the logic that the user has previously identified such as the status of object detection. If the condition is true, the demo site will proceed to the next task under this condition scope. If the logic is false, the Node-RED decoding command will skip the tasks under this condition scope and proceed to the next tasks outside of this condition scope.

The user interface in this application consists of two panels as shown in Figure 11. The left panel displays the knowledge contents, while the right panel is the block-based programming area. Both panels are overlaid on the demo site, allowing the user to view them and the environment simultaneously.

Figure 11: User interface Displayed over the Demo Site (User's View).

Contents

Three video animations and four device contents are provided for the content about the smart factory concept. The videos cover key areas such as the relationship and differences between industry 4.0 and smart factories, the benefits of upgrading traditional manufacturing to a smart one, the cyber-physical system with the Internet of Things (IoT), and the Industrial Internet of Things (IIoT). They also introduce the basic phase of the programming structure for controlling the robot arm and provide the basic instructions for the robot arm's three kinds of movements, which are linear, joint, and circular. Additionally, the video animation also presents the professional standards for robotics and automation careers in Thailand, as defined by the Thailand Professional Qualification Institute (TPQI).

The AR application features four device contents that introduces the devices implemented in the demo site. Furthermore, the information about each device helps the users understand the other categories of devices that are frequently used in the smart manufacturing. The device contents are the sensor content, the actuator content, the conveyor content, and the robot arm content.

The real time sensor status is displayed in the sensor content panel. The actuator content panel provides examples of actuator types and easy control buttons that users can use hand gestures to interact with them. In general, the actuators are controlled by electricity or the pressure, such as pneumatic actuators. The buttons can control the input stacker and pusher actuators to push and pull the objects. The elements in the conveyor content panel and the robot arm content panel are similar to those in the actuator content panel. There are the knowledge contents displaying text and the buttons that can control the devices.

Examination and Feedback Forms

This application was evaluated using the pre-test, post-test, system satisfaction feedback forms, and content satisfaction feedback form.

The closed-ended questions in system satisfaction feedback form and the content satisfaction feedback form use the Likert's scale [26] to measure user's attitudes. The scale starts from 1 to 5 points where the minimum point stands for strongly disagree / very poor / definitely not, and the maximum point stands for strongly agree / excellent / definitely.

The ideas to develop the factory are collected from the open-ended questions. These questions include how to implement the knowledge gained from the training system in the user's factory, ideas of implementing the AR technology to their factory, and the suggestions for improving the system.

4. Experimental Results

4.1. Population and Requirement

The proposed system is developed for workers who need to reskill or upskill their performance about robotics and automation in smart factory. The requirements for volunteers include: 1. Working in the automotive or electrical industry. 2. Being a key person in developing their factory into a smart factory. 3. Being an engineer, manager, or staff who is interested in robotics in smart factories. There are 22 volunteers who evaluated the proposed system.

4.2. Tools

Pre-test and Post-test

The purpose of the pre-test is to assess the user's knowledge background. Questions in pre-test are about contents in the smart factory and fundamental of each device used in the demo site.

The post-test's questions are the same as the ones in the pre-test. User takes the post-test after watching all video contents and the AR-based training system. The expectation of these tests is that user should get more points after attending the training system. Points gained after using the AR-based training system demonstrate that user has more knowledge about the smart factory concept than they did previously.

System Satisfaction Feedback Form

This form includes general questions about user's background, usability of this training system, and suggestion for improving the system. The usability questions cover the assessment of values for specific task of the proposed system. The writing parts in this form is used to collect user's ideas about the factory improvement based on experience after attending this training system.

Content Satisfaction Feedback Form

This form is used to collect the feedback about the contents. The content satisfaction also includes questions for evaluating the values for specific task.

4.3. Tasks

Task 1: User Assessment

There are 2 examinations for user assessment. User needed to follow these steps.

1. Taking the pre-test to assess the user's knowledge background.
2. After users have completed all the contents and used the application, they must take the post-test to evaluate their knowledge progress.

Task 2: Learning

There are two types of smart factory contents that users need to learn. The first type is three videos about the smart factory concept. The second type is contents about the devices that are mostly used in smart manufacturing. The device contents are provided in the AR application.

Task 3: Using the Application

1. This task is about how to use the application in the Microsoft HoloLens 2 by watching a tutorial video of the application. The video provided the suggested hand gestures to interact with the Microsoft HoloLens 2, the block's name of the Blockly, and how to use the Blockly by dragging and snapping the blocks together.
2. User started by learning about the device contents (sensor, actuator, conveyor, robot arm) that appear in the application. They practiced their hand gestures for interacting by pushing interactive buttons in the knowledge panel.
3. User created commands to control the demo site using Blockly panel and completed the given task (Algorithm 1).
4. After user completed using the application, user needed to take the system satisfaction feedback form and content satisfaction feedback form.

During operation, the user's view in the Microsoft HoloLens 2 was live streamed to the developer in real time. If user had the problems with the system or the interaction, developer would suggest how to solve them.

User took an hour to hour and half per one test following the steps 1-9. The time they took is up to how fast they can adjust their hand gestures to interact with the user interface, and how to solve the given programming task by themselves.

The risk of using this training system may cause the motion sickness for some people. To relieve the symptom, user needs to take the Microsoft HoloLens 2 off, sit on the chair and get eyes rest for 10-15 minutes. No one in this study had a motion sickness symptom while using this training system.

Algorithm 1: Making the Condition.

```
Start;
Instruction 1;
 if condition then
   | Instruction 2;
   | Instruction 3;
 else
   | Instruction 4;
Instruction 5;
 end
```

4.4. Results

The frame rate of the proposed system is in the range of 48-60 fps, which can display the augmented graphics clearly and user can interact with those graphics smoothly.

The formula (1) is to calculate the learning rate of all users measured by the percent change of post-test scores and pre-test scores. While "Post" stands for post-test score, and "Pre" stands for pre-test score.

$$\%Change = \frac{(Post-Pre)}{Pre} \times 100 \quad (1)$$

Table 3: User's Changed Learning Rate

Group	Experience	% changes (%)	Amount of users
1	Had experience about smart factory and had to implement it in their factory before.	23.42	9
2	Had knowledge about the smart factory concept but did not implement it before.	36.37	4
3	Had no knowledge and experience about the smart factory concept before.	75.25	9

The Table 3 showed that the % changes of all users were increased as 23.42% for 9 users, 36.37% for 4 users, and 75.25% for 9 users. The result showed that this proposed system can deliver the knowledge of the smart factory concept to users. The differences of the % changes are because user had some experience about smart factory before using this proposed system. Users in group 1 had less % changes than the other groups but it did not mean that they obtained few knowledge. The users in group 1 have more ideas of implementing new technology such as AR in their manufacturing processes because they can apply their priori knowledge and experience with this proposed AR-based training system.

The evaluations of the usability and values for specific task are based on 5-point Likert's scale as shown in Table 4. The usability score of the proposed system obtained 4.44 from 5 points, while score of the values for specific task of the system is about 4.23 from 5 points, and the values for specific task related to the contents obtained 4.30 from 5 points. Moreover, all users had more ideas of implementation new technology such as AR and the smart factory concept to their works.

Table 4: 5-point Likert's Scale.

Point	Meaning
1	Strongly disagree
2	Disagree
3	Neither agree nor disagree
4	Agree
5	Strongly agree

5. Conclusions and Discussions

Thailand has the direction to increase the manufacturing performance using the implementation of the smart factory. In the smart industry, skills related to robotics and automation are required. Worker skill is the most important factor in the factory

development. The AR-based training system for industrial training with the industry 4.0 contents aimed to develop worker's knowledge and skills. This proposed system covered the foundation contents about the smart factory and devices used in the smart production line. Moreover, the block-based programming panel is provided in the AR application to support users to practice their programming skills. The Microsoft HoloLens 2 device was selected to work with this system. The graphic contents are displayed on the Microsoft HoloLens 2 screen over the real environment. Worker upskilling can be evaluated with the scores of the pre-test and post-test. The points gained after completing the training from the proposed system indicate that the user's knowledge of smart factories and robotics and automation has increased.

The smart factory contents in this proposed system enhanced the user's knowledges. The proposed system can display with 48-60 fps. along with 1440 x 936 pixels per eye. Furthermore, the evaluation scores of the values for specific task and usability were in the range of 4-5 points, indicating that users were satisfied with this proposed system.

The experimental results discovered that the user's learning rate also depends on their priori knowledge and experiences. However, this proposed AR-based training system can delivery knowledge and experience about the smart factory concept and robot visual programming through augmented reality technology. This helps users to be able to utilize the obtained experiential learning for improving their manufacturing process.

Currently, the contents in this proposed system follows the standards in robotics and automation in TPQI standards. In the future, the blockly panel could be improved to increase the UI performance in order to solve the display delay problem. The various contents about the smart factory and digital disruption can be added to the system. This proposed system can be developed for various contents for training in some specific production lines. Moreover, the display of the device's status in the proposed system can be implemented for the data visualization in the smart production line.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

We would like to acknowledge the funding from PMU-B's AI for All 1-2 funds, the KMUTT's fundamental fund for financial supports, and the Human-Computer Interface Lab at the Institute of Field Robotics for providing the infrastructure and hardware implemented in this research.

References

[1] The National Electronics and Computer Technology Center (NECTEC), 2019, Cyber-Physical System, https://www.nectec.or.th/research/research-project/nectec-cps.html.

[2] Z. Suleiman, S. Shaikholla, D. Dikhanbayeva, E. Shehab, A. Turkyilmaz, "Industry 4.0: Clustering of concepts and characteristics," Cogent Engineering, **9**(1), 2022, doi:10.1080/23311916.2022.2034264.

[3] Germany Industrie 4.0, https://ati.ec.europa.eu/sites/default/files/2020-06/DTM_Industrie%204.0_DE.pdf.

[4] NIST, Smart Manufacturing | NIST, https://bit.ly/2IJXL4o.

[5] The Excise Department, Thailand 4.0, https://www.excise.go.th/cs/groups/public/documents/document/dwnt/mjgy/~edisp/uatucm282681.pdf.

[6] Business Opportunities, https://www.eeco.or.th/en/business-opportunities.

[7] The Federation of Thai Industries, 2020, Thailand Competetiveness 2016: Tha IMD Perspective, 6, http://www.nfcrbr.or.th/site/attachments/article/81/White%20paper.pdf .

[8] D. Mourtzis, V. Siatras, V. Zogopoulos, "Augmented reality visualization of production scheduling and monitoring," Procedia CIRP, **88**, 151–156, 2020, doi:10.1016/j.procir.2020.05.027.

[9] M. Speicher, B.D. Hall, M. Nebeling, "What is Mixed Reality?," in Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, ACM, New York, NY, USA: 1–15, 2019, doi:10.1145/3290605.3300767.

[10] Paul Milgram, Fumio Kishino, "A Taxonomy of Mixed Reality Virtual Displays", in IEICE Transaction on Information and Systems, Figure 1, 3.

[11] Reid Manning, 2022, Augmented, Virtual, and Mixed Reality Training Landscape [Online], Available: https://www.spheregen.com/augmented-virtual-and-mixed-reality-employee-training.

[12] S. Ke, F. Xiang, Z. Zhang, Y. Zuo, "A enhanced interaction framework based on VR, AR and MR in digital twin," Procedia CIRP, **83**, 753–758, 2019, doi:10.1016/j.procir.2019.04.103.

[13] F.-K. Chiang, X. Shang, L. Qiao, "Augmented reality in vocational training: A systematic review of research and applications," Computers in Human Behavior, **129**, 107125, 2022, doi:10.1016/j.chb.2021.107125.

[14] N. Gavish, T. Gutiérrez, S. Webel, J. Rodríguez, M. Peveri, U. Bockholt, F. Tecchia, "Evaluating virtual reality and augmented reality training for industrial maintenance and assembly tasks," Interactive Learning Environments, **23**(6), 778–798, 2015, doi:10.1080/10494820.2013.815221.

[15] F. Torres-Guerrero, L. Neira-Tovar, L. Torres-Treviño, An Introductive Training for Welding Workshop: A Biometric Evaluation Using Virtual Reality Scenes to Improve Practice, 319–331, 2019, doi:10.1007/978-3-030-18715-6_27.

[16] Z. Zhu, C. Liu, X. Xu, "Visualisation of the Digital Twin data in manufacturing by using Augmented Reality," Procedia CIRP, 81, 898–903, 2019, doi:10.1016/j.procir.2019.03.223.

[17] Educational Management, 2559, Thailand 4.0 Blueprint, https://waa.inter.nstda.or.th/stks/pub/2017/20171114-draeqa-blueprint.pdf.

[18] Office of The Official Information Commission, Thailand 4.0, http://www.oic.go.th/FILEWEB/CABINFOCENTER3/DRAWER049/GENERAL/DATA0000/00000702.PDF.

[19] Thailand Professional Qualification Database System, Robotics and Automation, https://tpqi-net.tpqi.go.th/home/occ/industrialInfo/RAC.

[20] Microsoft, HoloLens 2, https://www.microsoft.com/en-us/HoloLens/hardware#OneGDCWeb-Banner-ffqoo6k.

[21] Mikko Ruotsalainen, Lari Komulainen, Alex Tejada, "Virtual Lenses to Support Study Work at Customer Site", School of Engineering Science, 2021.

[22] Vuforia, Getting Started with Vuforia Engine in Unity, https://library.vuforia.com/getting-started/getting-started-vuforia-engine-unity.

[23] Vuforia, Image Targets, https://library.vuforia.com/objects/image-targets.

[24] D. Amin, S. Govilkar, "Comparative Study of Augmented Reality Sdk's," International Journal on Computational Science & Applications, **5**(1), 11–26, 2015, doi:10.5121/ijcsa.2015.5102.

[25] Node-RED, https://nodered.org.

[26] I. Elaine Allen and Christopher A. Seaman. (2022). Likert Scales and Data Analyses, https://www.bayviewanalytics.com/reports/asq/likert-scales-and-data-analyses.pdf.

Advances in Science, Technology and Engineering Systems Journal Vol. 8, No. 5, 17-27 (2023)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Feedback Controller for Longitudinal Stability of Cessna182 Fixed-Wing UAVs

Veena Phunpeng[1,*], Wilailak Wanna [1], Sorada Khaengkarn[1], Thongchart Kerdphol[2]

[1]*School of Mechanical Engineering, Institute of Engineering, Suranaree University of Technology, Nakhon Ratchasima, 30000, Thailand*

[2]*Department of Electrical Engineering, Faculty of Engineering, Kasetsart University, Bangkok, 10900, Thailand*

ARTICLE INFO

Article history:
Received:05 July, 2023
Accepted:08 August, 2023
Online: 24 September, 2023

Keywords:
Unmanned Aerial Vehicles
Fixed Wing
Fuzzy Logic Controller
PID Controller
Feedback loop controller

ABSTRACT

Unmanned aerial vehicles (UAVs) are becoming increasingly popular for both civil and military applications. Unmanned aerial vehicles can be categorized into two categories: rotary-wing and fixed-wing. Due to its capacity to fly long distances and carry substantial payloads, fixed-wing UAVs are gaining popularity and are currently utilized for various tasks. However, when confronted with disturbances such as weather or wind gusts, fixed-wing UAVs can rapidly lose stability, leading to a loss of lift and stalling. Consequently, it is vital to ensure the stability of fixed-wing UAVs. For the longitudinal stability management of a fixed wing unmanned aerial vehicle, the design and modeling of a feedback controller, including a PI controller, PID controller, and Fuzzy logic controller, are discussed in this article. MATLAB/SIMULINK©2021 will be used to design the control system and compare the response of each controller during the simulation. The controller's response to various input formats will indicate its capacity to regulate the system's behavior. Our results indicate that the fuzzy logic controller was superior to the PI and PID controllers at controlling the system's response according to the desired or input behavior.

1. Introduction

Unmanned aerial vehicles (UAVs) are becoming increasingly popular in a variety of civil and military applications, including aerial photography, shipping and delivery, geo-graphic mapping, disaster management, precision agriculture, search and rescue, weather forecasting, wildlife monitoring, law enforcement, and entertainment [1–3]. Unmanned aerial vehicles are small aircraft without a pilot on board. They can be remotely commanded and operated to accomplish operations in inaccessible and extremely dangerous regions. Unmanned aerial vehicles can be categorized into two basic categories: rotary-wing and fixed-wing [4]. Due to their capacity to fly long distances and carry heavy payloads, fixed-wing UAVs are currently gaining popularity and are employed for a variety of tasks [5]. Yet, fixed-wing UAVs quickly lose their stability when disturbances arise (weather or wind gusts). The instability of unmanned aircraft will cause the wings to lose lift, resulting in the stalling of fixed-wing UAVs [2,6–8]. Thus, the fixed-wing UAV's stability is vital.

Accidentally, Lion Air Indonesia's Boeing 737 MAX passenger aircraft crashed into the ocean in October 2018, resulting in 189 fatalities. Additionally, Ethiopian Airlines' Boeing 737 MAX caused 157 deaths in March 2019 [9]. Clearly, the breakdown of the flight control system caused the airplane to crash. The error-generating automatic flight control system is known as the Maneuvering Characteristics Augmentation System (MCAS) [10]. MCAS is a Boeing Company-developed system fitted in Boeing 737 MAX aircraft. MCAS is a mechanism that aids in the maintenance of stopped flights caused by an excessive angle of attack. MCAS is software coupled to the sensor that measures the angle of attack of an aircraft. The resultant MCAS mistake is the result of a faulty measurement of the rise of impact, in which the system calculates the impact angle to be greater than it actually is. The MCAS is activated when the angle of attack is greater than the angle of attack set by the system. This causes the elevator to adjust the aircraft to a progressiveMCASly decreasing angle when the pilot detects a change in the aircraft's pitch angle. The MCAS problem causes the plane to crash. Based on actual events aboard the Boeing 737 MAX, the significance of flight control systems in actual aircraft during flight is emphasized. Therefore, flight control systems are essential for unmanned aircraft. It relates directly to flight control and safety.

*Corresponding Author: Veena Phunpeng, veenap@g.sut.ac.th

The stability of fixed-wing UAVs can be broken down into three categories: directional stability, lateral stability, and longitudinal stability, with longitudinal stability being an essential type. It relates to the nose-up or nose-down position, takeoff or landing, and direct flying altitude of the UAV.

Fixed-wing UAVs employ a standard industrial feedback controller to reduce UAV expenses, such as a Proportional, Integral, Derivative, Proportional-Integral, and proportion-al-derivative controller, Proportional-Integral-Derivative controller, etc. [7,8,11]. Commercial controllers such as the Pixhawk and Paparazzi are commonly utilized in the flight control system of fixed-wing UAVs [4,12]. This present work describes control systems for fixed-wing UAVs.

In this study, the longitudinal stability management of a fixed wing unmanned aerial vehicle is constructed and simulated using a feedback controller comprising PI controller, PID controller, and fuzzy logic controller. MATLAB/SIMULINK©2021 will be used to simulate the control system and compare the responses of each controller. Modeling of aviation systems is explained in this study's summary. P Controller, I Controller, D Controller, PI Controller, PID Controller, and Fuzzy Logic Controller define the feedback controller. In the last section, the design of the control system is implemented, and MATLAB/SIMULINK©2021 simulation results are presented.

2. Aircraft System Modeling

The movement of a fixed wing unmanned aerial vehicle can be divided into three actions according to the imaginary rotation around the X, Y, and Z axes: roll, pitch, and yaw [8], as shown in Figure 1.

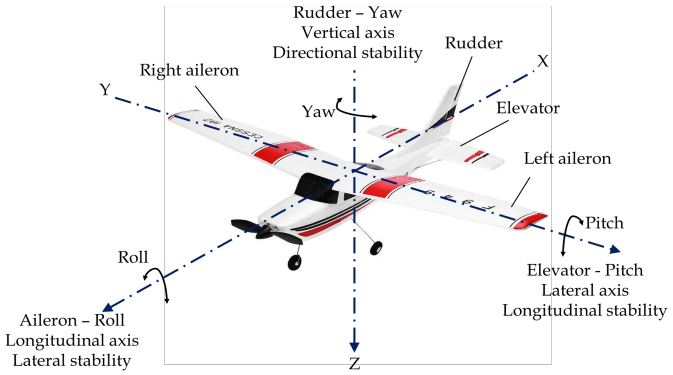


Figure 1: The movement of fixed-wing UAV.

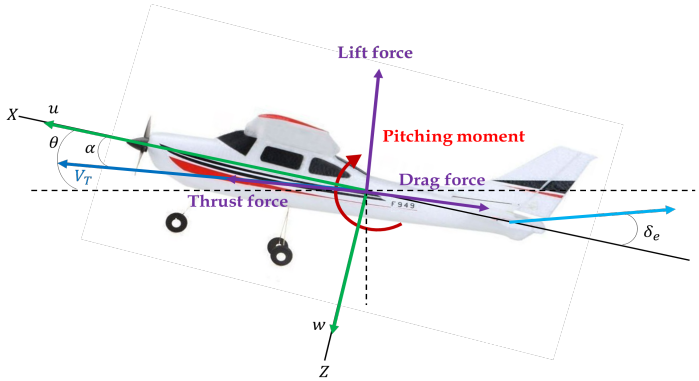


Figure 2: Pitch movement of fixed-wing UAV.

Pitch is the nose up or down of fixed-wing UAVs. It revolves around the Y-axis, an imaginary axis that goes from one wing tip to the other. It is controlled by a control surface known as an elevator, which is mounted on an empennage (horizontal stabilizer); when the elevator turns up and down (i.e., angle of the elevator: δ_E), which causes a pitch angle: θ as shown in Figure 2.

The pitch of the fixed-wing UAV can be written as a transfer function. In this research, the pitch transfer function of the fixed-wing UAV can be written as a transfer function, as shown in equation 1 [6,7,13].

$$\frac{\theta_s(s)}{\delta_{E_s}(s)} = \frac{-19893s^2 - 105510s - 15567}{86.1189s^4 + 1985.9478s^3 + 16150s^2 + 2082.5s + 945.7337} \quad (1)$$

where $\theta_s(s)$: Pitch angle
$\delta_{E_s}(s)$: Elevator angle

Equation 1 is a model of a fixed-wing Cessna182 UAV, with a ratio of 1:6.65 compared to a real scale, of which the fundamental dynamics of a mathematical model are similar to the real-sized aircraft. The wingspan of the model is 5.4135 ft. The dynamic of the pitching control system includes an actuator that can force the control surface to move as needed. The mathematical model used to control the elevator up and down is expressed in equation 2 [6,7,13].

$$\frac{\delta_E}{v_E} = -\frac{10}{s+10} \quad (2)$$

Where δ_E : Angle of the elevator
v_E : Input voltage

3. Feedback controller

Many systems in modern industries and other systems need a controller, such as an automobile steering control system, speed control systems, temperature control systems, tracking systems, flight control systems, etc. A controller controls the process or plant that responds to the method according to the desired output response. The control system can be divided into two systems 1) an Open-loop control system and 2) a Closed-loop control system.

The open-loop control system [12] is used in processes that do not require much control precision. Therefore, it is a process that does not require a back-measurement or feedback signal to be used to calculate constants for use in control. An open control process has system components, as shown in Figure 3. The application of the open-loop control system, such as fan control systems, washing machine control, etc.

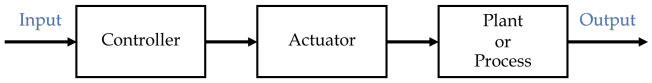


Figure 3: Open-loop control system.

A closed-loop control system [14,15] or feedback control (see Figure 4) is a controller in which the actual signal received from the system is measured and compared with the desired output signal to calculate errors. The error can be used to determine the system control constants to adjust the system to the desired output.

The block diagram of the closed-loop control system is shown in Figure 5.

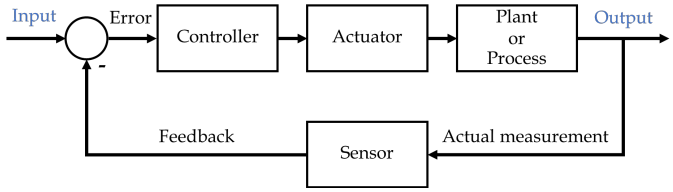


Figure 4: Closed-loop control system or feedback control system.

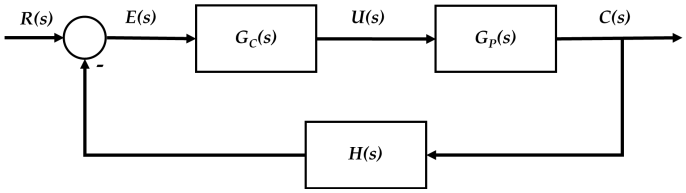


Figure 5: Block diagram of a closed-loop control system.

where R(s) : Input signal
E(s) : Error signal
GC(s) : Transfer function of the controller
U(s) : Control signal
GP(s) : Transfer function of plant or process
C(s) : Actual response signal
H(s) : Transfer function of the measurement sensor

General controllers for closed-loop control systems can be divided into Proportional-Integral (PI) controllers, Proportional-Derivative (PD) controllers, and Proportional-Integral-Derivative (PID) controllers. These controllers have used various industrial systems such as temperature, pressure, chemical, and vehicle manufacturing. The type of controller will have a different effect on the system response. Selecting a control type in the control design will depend on the response requirements of that system. The controllers can be divided into the following types.

3.1. P-Controller

A proportional controller [16] (P controller) is a control that determines the controller's gain or the proportion of the output signal to the input signal of the controller. The proportional controller has proportional gain to be used as an extension of the controller's input signal. If the proportional gain is high, the system will have a fast response. The result is that the response has an oscillation. On the other hand, when the proportional gain is low, the system will slowly respond. The proportional equation of the proportional control system can be found in equation 3.

$$G_C(s) = \frac{U(s)}{E(s)} = K_P \tag{3}$$

where K_P: Proportional Gain.

3.2. I-Controller

An integral controller is a working principle of controlling the output signal by signal integration. The integral controller can be described in equation 4.

$$G_C(s) = \frac{U(s)}{E(s)} = \frac{K_I}{s} \tag{4}$$

where K_I: Integral Gain

3.3. D-Controller

A derivative controller is a control that represents the derivation of the input signal of the controller. The derivative controller can be described in equation 5.

$$G_C(s) = \frac{U(s)}{E(s)} = K_D s \tag{5}$$

where K_D: Derivative Gain.

3.4. PI-Controller

A proportional-Integral [16] (PI) controller consists of a proportional or "P" controller and an integral or "I" controller. The P controller can be found in this equation.

$$G_C(s) = \frac{U(s)}{E(s)} = K_P + \frac{K_I}{s}$$

3.5. PID-Controller

The proportional integral derivative controller combines the proportional, integral, and derivative controllers shown in this equation.

$$G_C(s) = \frac{U(s)}{E(s)} = K_P + \frac{K_I}{s} + K_D s$$

Each controller has a different behavior, which can be summarized in Table 1.

In the control process, the gain must be selected according to the instability of the process. Select the controller gain that can be obtained from the Ziegler–Nichols method [17–19]. The procedure for finding the controller gain is as follows: The first step is to set K_I and K_D to the value 0, then increment K_P from zero until the maximum value is reached, resulting in K_U (Ultimate Gain). The system output will have a fixed vibration with the oscillation period or P_U. As shown in Table 2, the Ultimate gain (K_U) and Oscillation period (P_U) can be used to calculate the controller gain (K_P, K_I, and K_D).

Table 1: Characteristics of PID Gains K_P, K_I, and K_D.

Characteristic	K_P	K_I	K_D
Rise time	Decreases	Minimal Decrease	Minimal Decrease
Settling time	Minimal Decrease	Increase	Decrease
Overshoot	Increase	Increase	Decrease
Steady-state error	Decrease	Decrease	No impact
Stability	Decrease	Decrease	Increase

Table 2: Ziegler–Nichols method using Ultimate gain (K_U) and Oscillation period (P_U).

Controller	K_P	K_I	K_D
Proportional (P)	$0.5K_U$	-	-
Proportional-integral (PI)	$0.45K_U$	$\frac{0.54K_U}{P_U}$	-
Proportional- integral- derivative (PID)	$0.6K_U$	$\frac{1.2K_U}{P_U}$	$\frac{0.6K_UP_U}{8}$

3.6. Fuzzy Logic Controller

A fuzzy logic controller is a type of feedback controller that can be applied to a wide variety of systems [13,20–26]. Due to the ability to control nonlinear systems, complex systems Include systems that require immediate response (Real-time system) by a fuzzy logic controller is a controller that uses reasoning and decision-making principles similar to human decision-making with a work process structure as shown in Figure 6, which consists of four components as follows:

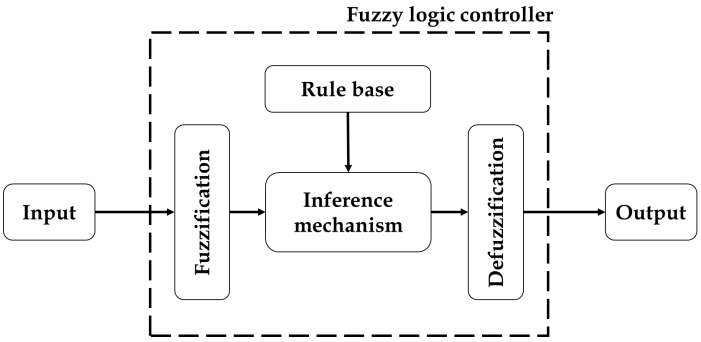


Figure 6: Fuzzy logic controller architecture.

- Fuzzification It is the part to convert the crisp input signal (Input) that is in the form of the original set with data values 0 and 1 (which can be interpreted as "false" and "true" as shown in Figure 7) into the form of a fuzzy set whose value is between 0 and 1 (the truth value is between false (0) and true (1), as shown in Figure 7).
- Rule base or rule base that was created as a tool used to evaluate the results of the input signal. The nature of the rule is in the form of "If-Then".
- The inference mechanism is used to make decisions or interpret problems related to the rule base.
- Defuzzification This is the process of converting the fuzzy set data back to the original data set.

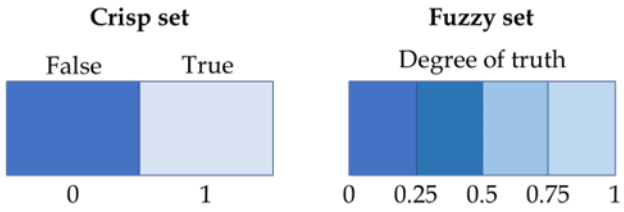


Figure 7: Crisp set and fuzzy set.

4. Simulation

This section shows the design of the control system through the program MATLAB/SIMULINK©2021. For the PI and PID controllers, the controller gains of PI and PID controllers are tuned by using Ziegler–Nichols method as shown in Figures 8 and 9. Where K_U = 3.162 and P_U = 0.384 s. The gain controller (K_P, K_I, and K_D) is shown in Table 3.

To design a fuzzy logic controller for longitudinal stability control of fixed-wing UAVs, able to design a control system through the program MATLAB/SIMULINK©2021, as shown in Figure 10.

From the structure of the fuzzy logic control system, as shown in Figure 10, it can be seen that there are two input signals for the fuzzy logic controller, the error signal (e) is the difference between the actual measurement pitch angle (output) is compared to the desired pitch angle. The error rate (Δe) is the error change rate over time. The fuzzy logic controller provides an output signal that can be used to control the transfer function of the pitch angle.

To design a fuzzy logic controller, the data value must be changed to a fuzzy set when receiving an incoming signal (crisp input signal) as a number or in the original set format. It uses the semantic characteristics of linguistic variables, which are linguistic variables for the input signal (error and error rate). The output signal is {NB, NM, NS, ZO, PS, PM, PB}, which means a fuzzy set. They have the following meanings: Negative Big, Negative Medium, Negative Small, Zero, Positive Small, Positive Medium, and Positive Big, respectively.

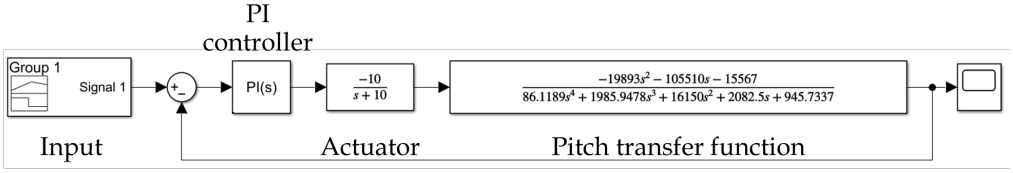


Figure 8: The PI control system.

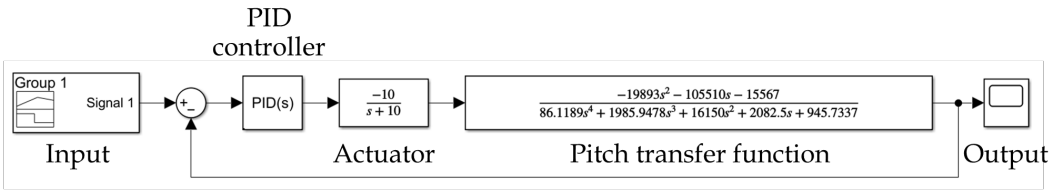


Figure 9: The PID control system.

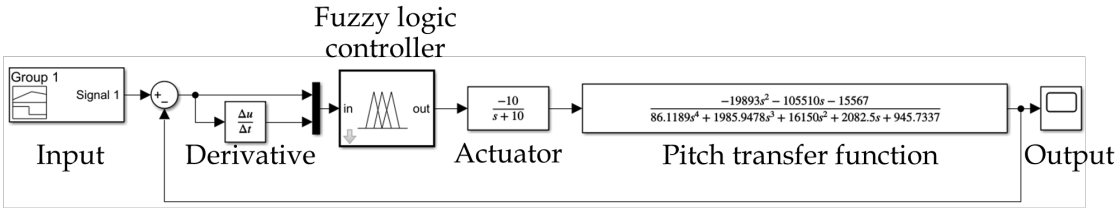


Figure 10: Fuzzy logic control system.

Table 3: The Ziegler–Nichols tuning PI and PID controller method uses Ultimate gain (K_U) and Oscillation period (P_U).

Controller	K_P	K_I	K_D
Proportional-integral (PI)	1.423	4.447	-
Proportional- integral- derivative (PID)	1.892	9.881	0.091

This research uses two types of membership functions for fuzzy sets: The triangular membership function for {NM, NS, ZO, PS, PM} and the trapezoidal membership function for {NB, PB} are shown in Figure 11 – 13.

The fuzzy rule base provides an output signal for the system's control. The fuzzy rule base is set according to the "If-Then" rule. The fuzzy rule base has 49 rules, as shown in Table 4.

Figure 14 shows the architecture of a Fuzzy-PID controller. This sort of controller combines a fuzzy logic controller with a PID controller employing an adaptive controller to select the best appropriate controller gains. Figure 15-16 illustrates the membership function employed by the Fuzzy-PID controller, whilst Table 5-7 displays the fuzzy rule base.

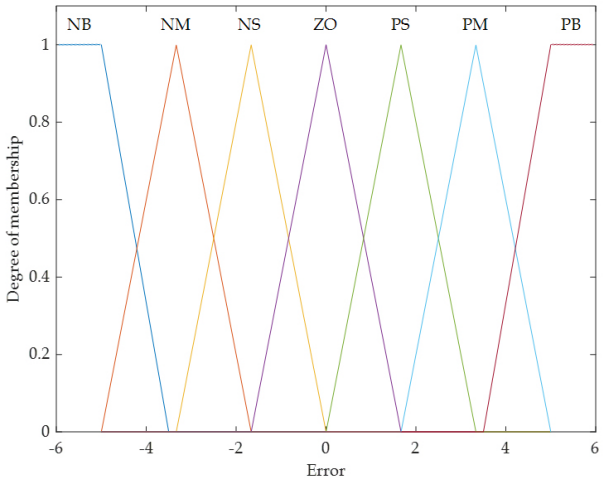


Figure 11: Membership function of error.

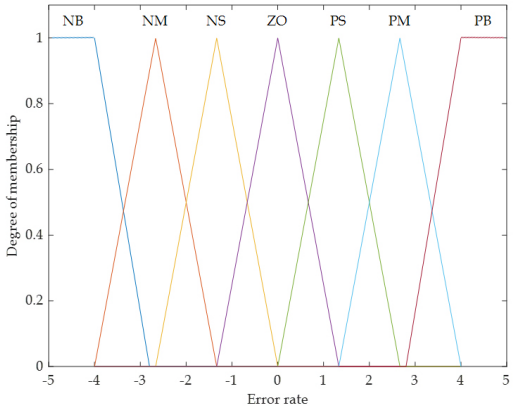


Figure 12: Membership function of error rate.

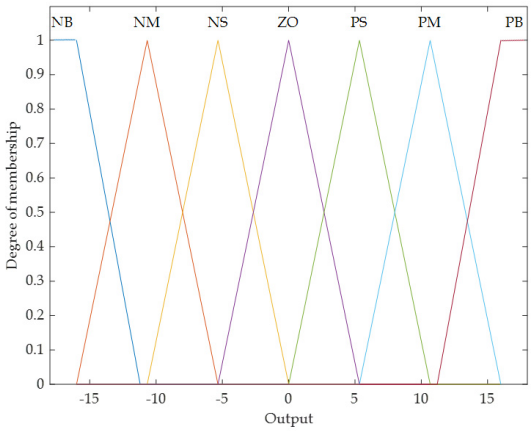
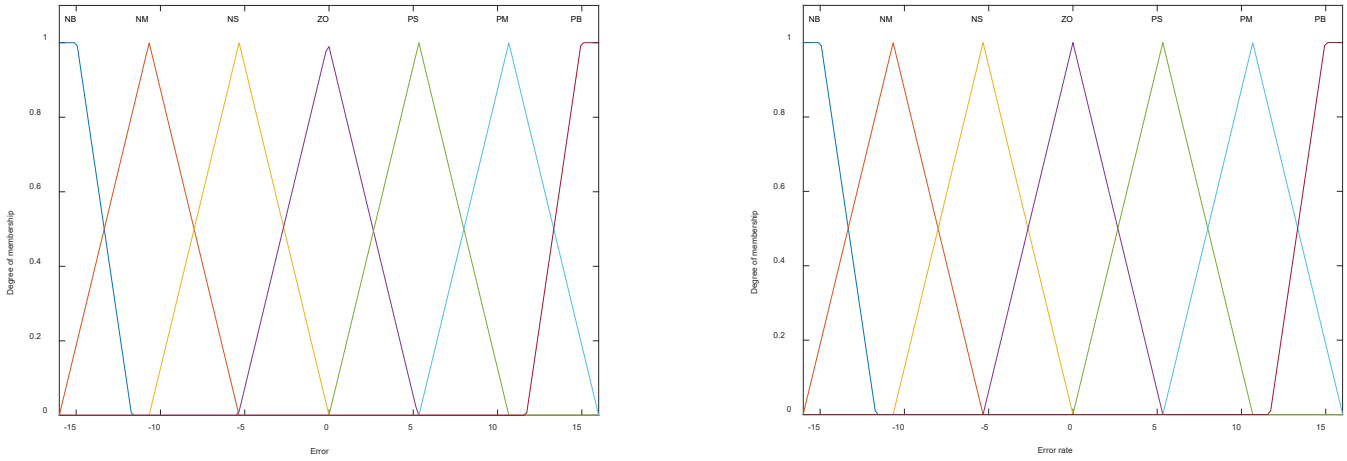


Figure 13: Membership function of output.

Table 4: Fuzzy logic rules base for fuzzy logic controller.

Error \ Error rate	NB	NM	NS	ZO	PS	PM	PB
NB	NB	NB	NB	NB	NB	NB	NM
NM	NB	NB	NB	NB	NB	NM	NS
NS	NB	NB	NB	NB	NM	NS	NS
ZO	NM	NS	NS	ZO	ZO	PS	PM
PS	PS	PS	PM	PB	PB	PB	PB
PM	PS	PM	PB	PB	PB	PB	PB
PB	PM	PB	PB	PB	PB	PB	PB



(a) (b)

Figure 15: Membership function of input for Fuzzy-PID controller (a) Error (b) Error rate.

Table 5: Fuzzy logic rules base for K_P.

Error \ Error rate	NB	NM	NS	ZO	PS	PM	PB
NB	PB	PB	PM	PM	PS	ZO	ZO
NM	PB	PB	PM	PS	PS	ZO	NS
NS	PM	PM	PM	PS	ZO	NS	NS
ZO	PM	PM	PS	ZO	NS	NM	NM
PS	PS	PS	ZO	NS	NS	NM	NM
PM	PS	ZO	NS	NM	NM	NM	NB
PB	ZO	ZO	NM	NM	NM	NB	NB

Table 6: Fuzzy logic rules base for K_I.

Error \ Error rate	NB	NM	NS	ZO	PS	PM	PB
NB	PB	PB	PM	PM	PS	ZO	ZO
NM	PB	PB	PM	PS	PS	ZO	NS
NS	PM	PM	PM	PS	ZO	NS	NS
ZO	PM	PM	PS	ZO	NS	NM	NM

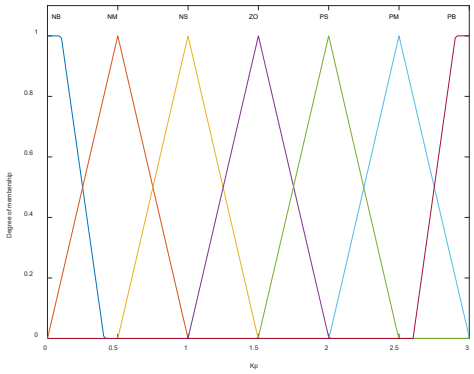
PS	PS	PS	ZO	NS	NS	NM	NM
PM	PS	ZO	NS	NM	NM	NM	NB
PB	ZO	ZO	NM	NM	NM	NB	NB

Table 7: Fuzzy logic rules base for K_D.

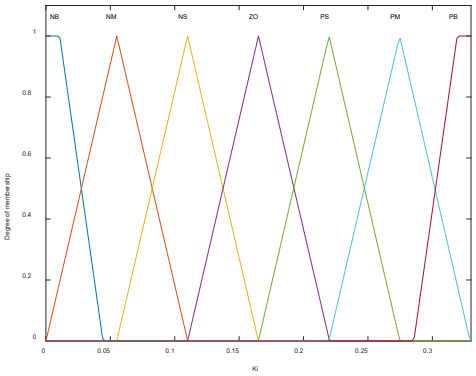
Error / Error rate	NB	NM	NS	ZO	PS	PM	PB
NB	PB	PB	PM	PM	PS	ZO	ZO
NM	PB	PB	PM	PS	PS	ZO	NS
NS	PM	PM	PM	PS	ZO	NS	NS
ZO	PM	PM	PS	ZO	NS	NM	NM
PS	PS	PS	ZO	NS	NS	NM	NM
PM	PS	ZO	NS	NM	NM	NM	NB
PB	ZO	ZO	NM	NM	NM	NB	NB

5. Results and Discussion

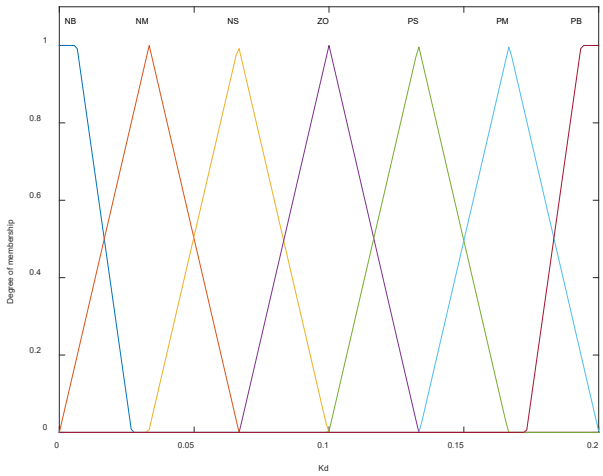
This section shows the simulation results of the designed control system for the pitch control system of the Cessna182 fixed-wing UAV through MATLAB/SIMULINK©2021. The results show the comparison of the response of the original system (no controller), PI controller, PID controller, and fuzzy logic controller.



(a)



(b)



(c)

Figure 16: Membership function of output for Fuzzy-PID controller (a) K_P (b) K_I (c) K_D.

For the original system (no controller), it was found that the system had zero position for open-loop transfer function at -0.1519 and – 5.152 and pole position at -0.0618 ± 0.2359i, – 10 and -11.4685 ± 7.2878i, as shown in Figure 17. For the pole at position -11.4685 ± 7.2878i, the system has a high overshoot and takes a long time to reach equilibrium (settling time).

Therefore, to improve the system's behaviour, reduce the oscillation (overshoot), and increase the stability of the pitch system. The control system (PI controller, PID controller, and fuzzy logic controller) was designed and simulated through MATLAB/SIMULINK©2021, with the following input format that can be divided into 3 cases.

Figure 18 shows the response of step-up input. The noncontrolled system is slow to respond and significantly overshoots with high errors. When the PI controller controls the system, the system has a faster response but a high overshoot. The PID controller provides a quick response and lower overshoot

than the PI controller. Comparing these results with the fuzzy logic controller shows that the reaction does not have an overshoot.

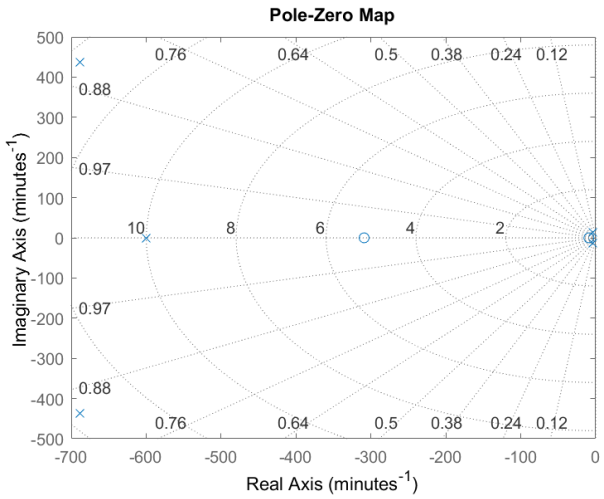


Figure 17: Pole and zero of system.

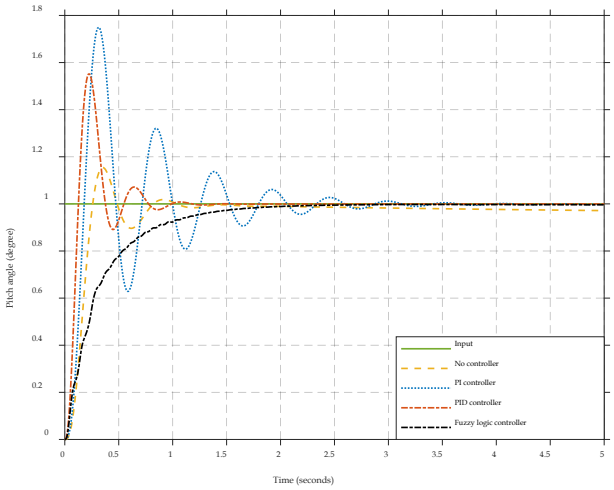


Figure 18: Input as a step up.

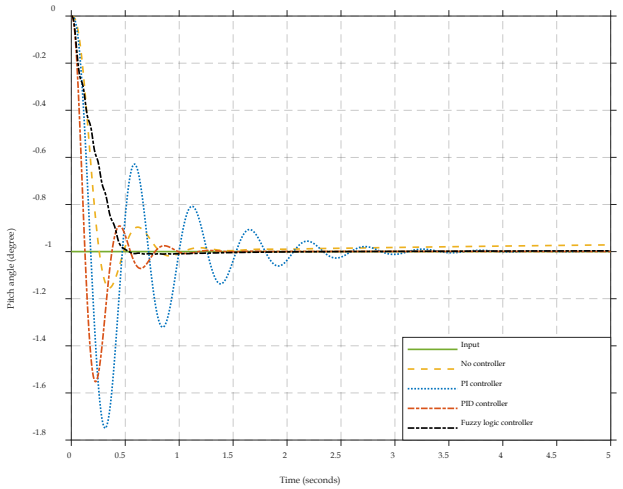


Figure 19: Input as a step-down.

Figure 19 shows that the system response with step-down input has the same response as a system with step-up input. Systems with no controller have slow response and overshoot. The PI and PID controller systems have a faster response but also a high overshoot. The fuzzy logic controller provides the best response without overshoot and fast system response.

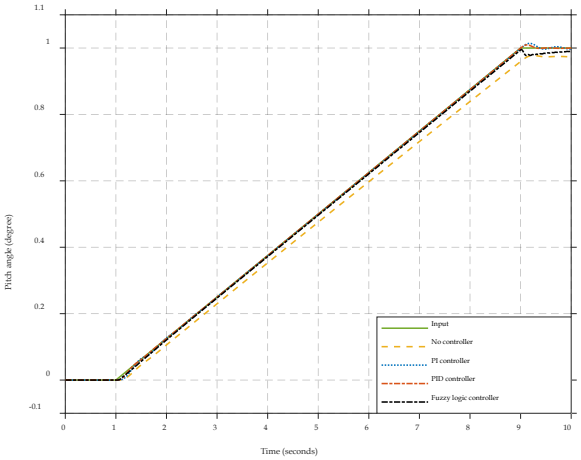


Figure 20: Input as ramp up.

Figure 20 illustrates the response of the system to a ramp-up input. Without a controller, the system malfunctions. The system's response demonstrates that both PI and PID controllers result in overshoot. Nonetheless, the fuzzy logic controller responds adequately to the desired input without overshooting.

For the ramp-down input, the response is shown in Figure 21 The system response result clearly indicates that the system has an error when there is no controller system. The PI and PID controllers decrease the system's error but have high overshoot. The fuzzy logic controller provides a satisfactory response with reduced error and overshoot.

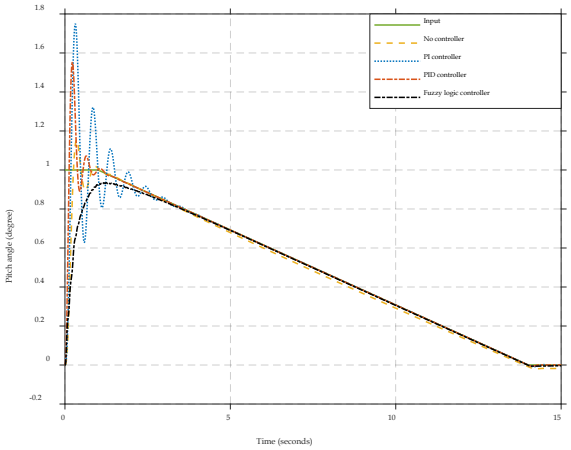


Figure 21: Input as ramp down.

In the second case, the input is a combined disturbance divided into three modes. The response in the second case is shown in Figure 22 – 26.

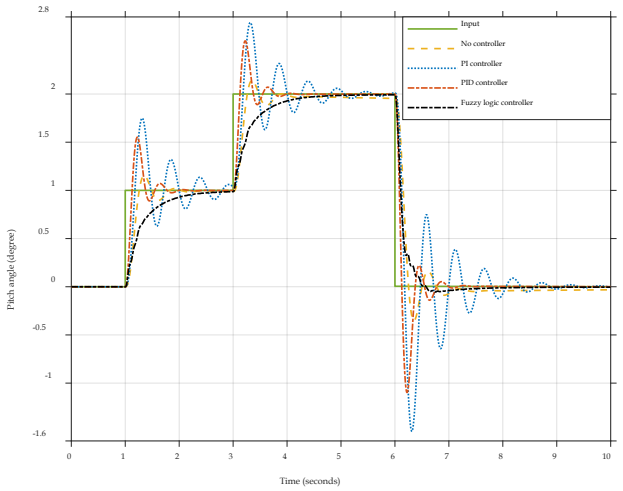


Figure 22: Input mode 1.

The first mode's response is shown in Figure 22. The system response shows indicates system without a controller has errors and overshoots. PI controllers provide high system response with high overshoot compared to PID controllers and fuzzy logic controllers. The fuzzy logic controller provides a system response that is fault-free and overshoot-free, meaning the system has no oscillation.

For the second mode (Figure 23), this input combines the step input and the ramp input. The system response is as close as possible to the input. From the system's response, it was found that no controller system has the most valuable error. Fuzzy logic controllers provide the best system response compared to PI and PID controllers.

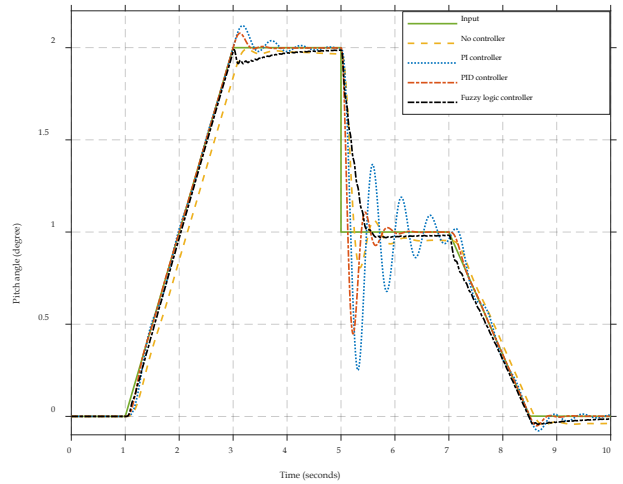


Figure 23: Input mode 2.

The input is a sine wave in the third mode (Figure 24). From the system's response, it was found that no controller system has the most valuable error. The PID and Fuzzy logic controllers provide the best system response and are as close as possible to the input.

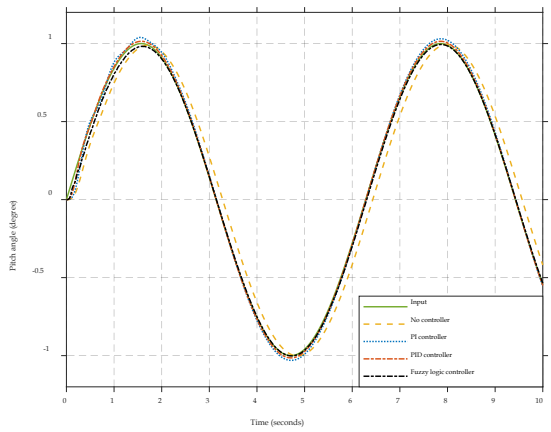


Figure 24: Input mode 3.

For the third case, set the system input as a continuous disturbance, as shown in Figure 25, and zoom-in of input as a constant disturbance shown in Figure 26. From the system's response, the no-controller system has a high error. PI and PID controllers, a high overshoot is provided. The fuzzy logic controller provides the most input-compliant system response. The system response was no error and overshoot.

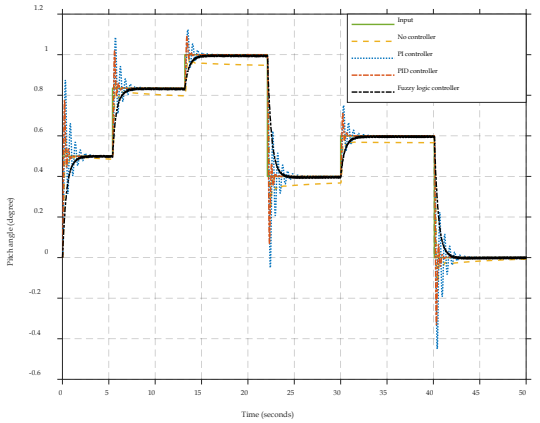


Figure 25: Input as a continuous disturbance.

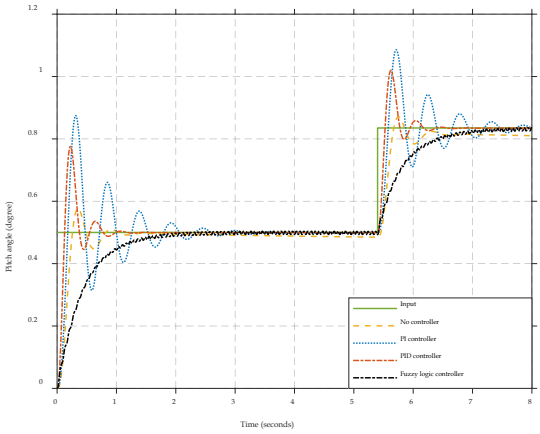


Figure 26: Zoom-in of input as a continuous disturbance.

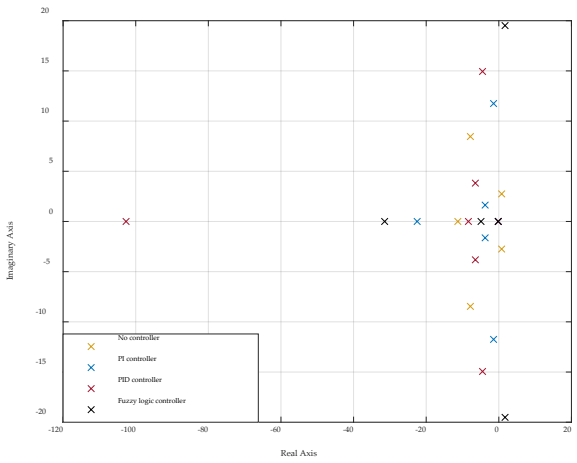


Figure 27: Pole displacement of input as a continuous disturbance.

Figures 27 show the pole displacement of each input mode. The pole displacement of each input mode is the same position because each input mode uses the same controller gains and rules (no controller, PI controller, PID controller, and fuzzy logic controller). Therefore, pole displacements are the same. From the pole position of no control system, PI control systems and PID control systems have pole displacement on imaginary axes that show the system has overshoot or oscillation in the system. For fuzzy logic control systems, make the system not have a pole on an imaginary axis. The system has a fuzzy logic control system without overshoot or oscillation.

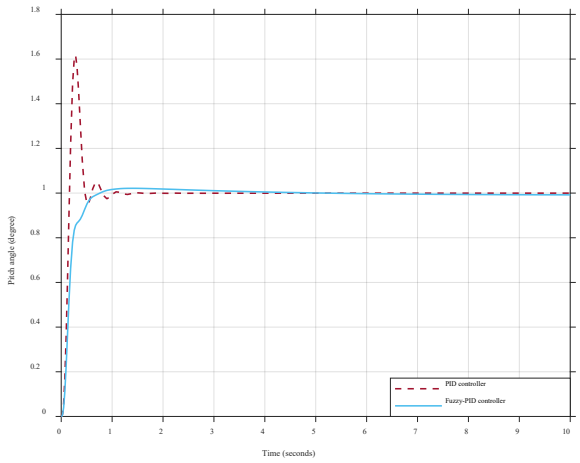


Figure 28: Response of PID and Fuzzy-PID controller.

Figure 28 compares the PID controller's response to that of the fuzzy PID controller. The control results show that a controller with the application of a fuzzy logic controller has a better system response than a system using a PID controller. As observed from the control effect, when using a controller with the application of a fuzzy logic controller, the overshoot and rise time can be reduced. The steady-state errors are removed.

6. Conclusion

UAVs tend to lose stability easily when they are disturbed by disturbances (Weather or wind gusts). This can cause the UAV to lose control and possibly stall easily. Therefore, UAV systems need a controller that is used to control the stability of the UAV. The longitudinal stability is the most important in many types of stability of UAVs due to the longitudinal stability directly influencing the pitch motion of the UAVs. The pitch motion directly involves UAVs' take-off, landing, and holding altitude. In this research, the design and simulation of longitudinal stability control of a fixed-wing unmanned aerial vehicle with a feedback controller include PI controller, PID controller, and fuzzy logic controller. The simulation uses MATLAB/SIMULINK©2021 to design the control system with each controller and compare the response of each controller. The controller's response according to the input format shows the ability to control system behavior. The comparison of the PI, PID, and fuzzy logic controllers show the fuzzy logic controller was able to control the system response as desired or according to the behavior of the input to the system. In the future, in addition to longitudinal stability control, fuzzy logic controllers can also be applied to control other systems in UAVs.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This research was made possible by the Suranaree University of Technology (Grant Number FF7-707-66-12-82), Thailand Science Research and Innovation (TSRI), and National Science Research and Innovation Fund (NSRF) (NRIIS number 179358).

References

[1] H. Shakhatreh, A.H. Sawalmeh, A. Al-Fuqaha, Z. Dou, E. Almaita, I. Khalil, N.S. Othman, A. Khreishah, M. Guizani, Unmanned Aerial Vehicles (UAVs): A Survey on Civil Applications and Key Research Challenges, IEEE Access, **7**, 48572–48634, 2019, doi:10.1109/ACCESS.2019.2909530.

[2] Y. Cheng, S. Xu, Z. Huang, B. Xiong, C. Wang, "Stability control of small fixed-wing UAV," in Proceeding of the 11th World Congress on Intelligent Control and Automation, IEEE: 1056–1060, 2014.

[3] V. Phunpeng, T. Kerdphol, "Integrated Piezoelectric and Flexible Solar Systems for Enhancing Inflight Operation Time of Small UAV," Suranaree Journal of Science and Technology, In press, 1–25, 2022.

[4] M.-C. Wen, C.-H. Yang, E.-X. Sung, T.-H.R. Wu, Optimized integration of UVAs surveys, and image-based modeling strategies for digital terrain model reconstruction Program of Smart Living View project, 2014.

[5] J.L. Hernandez, I. González-Hernández, R. Lozano, "Attitude and altitude control for a fixed wing UAV applied to photogrammetry," in 2019 International Conference on Unmanned Aircraft Systems (ICUAS), IEEE: 498–502, 2019.

[6] J.H. Yang, H.K. Xu, "Robust Controller Design for Non-Minimum Phase UAV System and System Analysis," IEEE Access, **6**, 70734–70769, 2018, doi:10.1109/ACCESS.2018.2879649.

[7] B.M. Albaker, M. Majidi, B.M. Albaker, N.A. Rahim, "Flight path PID controller for propeller-driven fixed-wing unmanned aerial vehicles Related papers Flight Cont rol Syst em for Small High-Performance UAVs EXPERIMENTAL FRAMEWORK FOR EVALUAT ION OF GUIDANCE AND CONT ROL ALGORIT HMS FOR UAVs Flight path PID controller for propeller-driven fixed-wing unmanned aerial vehicles," International Journal of the Physical Sciences, **6**(8), 1947–1964, 2011, doi:10.5897/IJPS11.162.

[8] J. Yang, X. Wang, S. Baldi, S. Singh, S. Fari, "A software-in-the-loop implementation of adaptive formation control for fixed-wing UAVs,"

IEEE/CAA Journal of Automatica Sinica, **6**(5), 1230–1239, 2019, doi:10.1109/JAS.2019.1911702.

[9] J. Herkert, J. Borenstein, K. Miller, "The Boeing 737 MAX: Lessons for Engineering Ethics," Science and Engineering Ethics, **26**(6), 2957–2974, 2020, doi:10.1007/s11948-020-00252-y.

[10] S. Makó, M. Pilát, P. Šváb, J. Kozuba, M. Čičváková, "Evaluation of MCAS System," Acta Avionica Journal, 2020, doi:10.35116/aa.2020.0003.

[11] Z. Zheng, Z. Jin, L. Sun, M. Zhu, "Adaptive Sliding Mode Relative Motion Control for Autonomous Carrier Landing of Fixed-Wing Unmanned Aerial Vehicles," IEEE Access, **5**, 5556–5565, 2017, doi:10.1109/ACCESS.2017.2671440.

[12] E. Ebeid, M. Skriver, J. Jin, "A Survey on Open-Source Flight Control Platforms of Unmanned Aerial Vehicle," in Proceedings - 20th Euromicro Conference on Digital System Design, DSD 2017, Institute of Electrical and Electronics Engineers Inc.: 396–402, 2017, doi:10.1109/DSD.2017.30.

[13] M.J. Patyra, J.L. Grantner, K. Koster, Digital Fuzzy Logic Controller: Design and, 1996.

[14] Electrical4U, "What is a Control System? (Open Loop & Closed Loop Control Systems Explained)," Https://Www.Electrical4u.Com/, 2020.

[15] M. Saoudi, A. El-Sayed, H. Metwally, "Design and implementation of closed-loop control system for buck converter using different techniques," IEEE Aerospace and Electronic Systems Magazine, **32**(3), 2017, doi:10.1109/MAES.2017.150261.

[16] V. Jain, "Types of Controllers | Proportional Integral and Derivative Controllers | Electrical4U," Https://Www.Electrical4u.Com/, 2021.

[17] V.V. Patel, "Ziegler-Nichols Tuning Method," Resonance, **25**(10), 2020, doi:10.1007/s12045-020-1058-z.

[18] V.V. Patel, "Ziegler-Nichols Tuning Method: Understanding the PID Controller," Resonance, **25**(10), 2020, doi:10.1007/s12045-020-1058-z.

[19] V. Kumar, A. Patra, "Application of Ziegler-Nichols Method for tuning of PID controller," International Journal of Electrical and Electronics Engineers, **8**(2), 2016.

[20] M.J. Patyra, J.L. Grantner, K. Koster, "Digital fuzzy logic controller: Design and implementation," IEEE Transactions on Fuzzy Systems, 4(4), 1996, doi:10.1109/91.544304.

[21] C. Elmas, O. Deperlioglu, H.H. Sayan, "Adaptive fuzzy logic controller for DC-DC converters," Expert Systems with Applications, **36**(2 PART 1), 2009, doi:10.1016/j.eswa.2007.11.029.

[22] J. Carvajal, G. Chen, H. Ogmen, "Fuzzy PID controller: Design, performance evaluation, and stability analysis," Information Sciences, **123**(3), 2000, doi:10.1016/S0020-0255(99)00127-9.

[23] A.A. Thorat, S. Yadav, S.S. Patil, "Implementation of Fuzzy Logic System for DC Motor Speed Control using Microcontroller," Journal of Engineering Research and Applications (IJERA), **3**(2), 2013.

[24] A.C. Soh, E.A. Alwi, R.Z.A. Rahman, L.H. Fey, "Effect of Fuzzy Logic Controller Implementation on a Digitally Controlled Robot Movement," Kathmandu University Journal of Science, Engineering and Technology, **4**(1), 1970, doi:10.3126/kuset.v4i1.2881.

[25] B. Hamed, M. Almobaied, "Fuzzy PID Controllers Using FPGA Technique for Real Time DC Motor Speed Control," Intelligent Control and Automation, **02**(03), 2011, doi:10.4236/ica.2011.23028.

[26] V. Phunpeng, S. Khaengkarn, T. Hengmeechai, T. Kerdphol, "Longitudinal Stability Assessment of Fixed-Wing UAV based on Fuzzy Closed-Loop Control," in Proceeding - 2023 International Electrical Engineering Congress, iEECON 2023, 2023, doi:10.1109/iEECON56657.2023.10126642.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 8, No. 5, 28-34 (2023)

www.astesj.com

ASTESJ

ISSN: 2415-6698

The Effect of Introduction of Wind Energy System on the Energy Networks

Ahmed Slimani[1], Benyoucef Merah[1], Mohammed Nasser Tandjaoui[*2], Chellali Benachaiba[2]

[1]*Department of Computer Science, Faculty of Science Exact, Tahri Mohamed University, Bechar, Algeria*

[2]*Department of Electrical Engineering, Faculty of Technology, Tahri Mohamed University, Bechar, Algeria*

ARTICLE INFO

Article history:
Received:21 August, 2022
Accepted:28 February, 2023
Online: 24 September, 2023

Keywords:
Wind energy
Distributed network
Decentralized production
Sustainable energy
Hybrid system

ABSTRACT

The increase electricity production based on renewable energies sources responds to the requirements of sustainable development. In recent years, the investment in the wind energy sector has increased to satisfy green electricity generation and environment protection. Wind energy is currently the most competitive energy resource that is part of the world's countries views on the future of energy production and exploitation and its impact on the environment, it plays a very important role to reduce the CO2 emissions injected by the energy sectors, namely transport and industry. Wind energy is a good alternative to fossil fuels as its ecological impact is low. This technology is developing rapidly of which the wind turbine industry is the most dynamic of the industries producing large equipment for power generation. Therefore, wind energy has a major drawback, which is the inability to predict accurately because of its uncontrolled fluctuations, which enter as a major component of energy flow problems within the electrical system. This paper presents the effects of wind power integration on the power quality of distribution network when the weak nature of the network in remote areas and the uncertainty of the wind are taken into consideration. The proposed model design has been tested in simulation using MATLAB/SIMULINK, where the results obtained confirm the electrical degradation in terms of quality in electrical networks integrating wind turbines, to ensure a clean electrical quality, the optimal solution is the integration of FACTS systems which will be the subject of the next paper.

1. Introduction

Today more than 85% of the energy produced is obtained from fossil fuels (petroleum, coal, and natural gas) and nuclear energy, which generates strong environmental pollution by emissions of greenhouse gases affecting climate change or, in the case of nuclear, long-lasting radiation pollution, which is not yet resolved, of the storage of radioactive waste. The constant growth of energy consumption in all its forms and the associated polluting effects, mainly caused by the combustion of fossil fuels, are at the heart of the issue of sustainable development and the care of the environment in a discussion for the future of the planet [1, 2]. The massive use of fossil and fissile energies, even if it has invaded the entire field of human activity today, remains an epiphenomenon on the scale of human history; it appears through two small peaks, one during the 19th century with coal and the discovery of steam engines, the other in the 20th century with oil, gas and nuclear power [2-5]. Two of the major challenges for our century are the fight against climate change and the diversification of the sources of energy that we currently use, to serve the needs of the poor nations which represent more than two-thirds of the population, the world needs to forge a new energy strategy, which, in order to respect the global environment, must first be based on sobriety and energy efficiency and inevitably use renewable sources which undoubtedly have an important role to play

The electricity industry thus seems to be backtracking, when most of the electrical energy was generated locally by small isolated systems for direct use. The old steam generators, used to provide heat and electricity, have found their modern equivalents in the form of micro-turbines, fuel cells, internal combustion engines and small gas turbines. In addition to economic interest, other arguments have argued in favor of a transition to small-scale decentralized energy systems; these include environmental impacts, the vulnerability of centralized energy systems in the event of an attack, and the reliability of electricity [1, 4]. With regard to technologies of decentralized energy production systems, there are currently two types of systems with a strong footprint: renewable energy systems and cogeneration systems. These systems, favored by the public authorities, have become an economically viable reality and are therefore imposed. With regard to renewable energies, the conversion devices range from a few tens of watts to a few megawatts. Wind power is the only one that is both inexpensive and easily exploitable. In addition to the production of wind energy on the grid, the market for small and

*Corresponding Author: Mohammed Nasser Tandjaoui, tandjaoui.naceur@univ-bechar.dz

medium-sized wind power plants, intended to supply isolated sites, is very promising and booming in Electricity Market's energy production. Wind energy is the fastest growing renewable energy in the world. It is almost universally recognized as the most promising energy source for producing clean electricity in the short to medium term and it contributes to the preservation of the environment. The wind is an element whose study is very complex, its characterization depends on several parameters such as the measurement of the wind speed, its direction, the effect of the roughness of the ground, the effect of obstacles., the effect of the stability of the atmosphere, etc. In this study, we will present the technical anomalies produced by the integration of wind turbines in the distribution network.

3. Description of energy system used in the production of wind energy

Wind turbines always use a synchronous or asynchronous squirrel cage generator as a generator, where two main processes are applied: fixed speed and variable speed. A wind turbine is made up of a tower at its top is fixed the nacelle. The nacelle consists of a system for transforming wind energy into electrical energy with its control. Figure 1 shows the components of the electro-mechanical chain. Cables for transporting electrical energy, control and command elements, equipment for connecting to the low-voltage distribution network, are laid out inside the tower [3, 6-12].

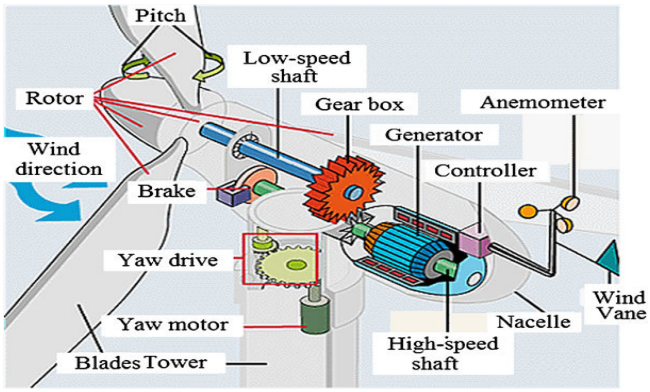


Figure 1: Components of the electromechanical coupling of wind turbine drive system

4. Conversion of wind energy into electrical energy

A sustainable development of the means of energy production based on wind energy will only be achievable with a decrease in implementation cost, so that they can be integrated into the energy production market [6]. The production of electricity from wind turbines or wind turbines in an isolated site requires the use of the electric machine - static converter assembly [3.13]. The wind turbine gets its energy from the wind, therefore there is a relationship between wind speed V and rotational speed ω, torque and power on the rotor of a wind turbine. The mechanical power that can be extracted from the wind is determined by the following expression:

$$P_v = \frac{1}{2}\rho S V^3 \quad (1)$$

ρ : Air density (kg/m3),

S : Surface crossed by air (m2), S = π. R2,

R : Wind turbine rotor radius,

V : Wind speed (m/s).

The power supplied by the rotor of the wind turbine is given by the following relation:

$$P = C_P . \frac{\rho . S . V^3}{2} \quad (2)$$

Cp : Coefficient depending on the shape of the rotor and the wind speed. Thus there is a linear relationship between the wind speed and the angular speed of the rotor [6]

$$k = \omega \frac{R}{V} \quad (3)$$

ω : Angular speed of the wind turbine rotor.

The coefficient Cp in the wind turbine is a function of k, and it is given by the manufacturer according to the type of sensor as indicated in table 1.

Table 1. values of coefficients k and Cp for the wind turbine

k	0	1,3	2,1	2,7	3,4	5	5,8	6,3
Cp	0	0,1	0,2	0,3	0,4	0,3	0,2	0,1

We define the torque developed on the rotor of the wind turbine as follows:

$$T = \frac{P}{\omega} \quad (4)$$

With: $$T = C_P . \frac{\rho . S . V^3}{2\omega} = C_P . \frac{R . \rho . S . V^2}{2k} \quad (5)$$

In the context of the insertion of decentralized production units in the electricity networks, it will be necessary to study the means to implement these units for reactive power compensation and active power control, as well as the control of the rms value of the voltage. Medium power production units (usually < 600 kVA) are generally equipped with asynchronous cage machines.

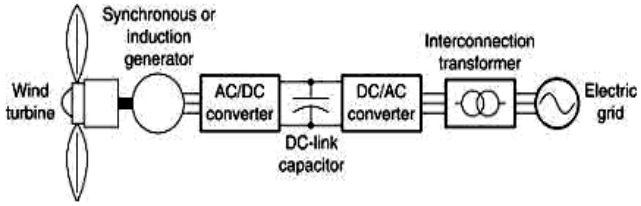


Figure 2: Wind power conversion chain

In fact, the asynchronous cage generator is currently the most widely used electrical machine in the production of fixed speed wind power. This generator can operate at variable speed, through the use of power converters, and can generate electrical power output over a wide range of wind speeds. Figure 2 shows the complete structure of the conversion chain fitted with its control members. The scarcity of non-renewable energies, atmospheric pollution, global warming and nuclear risks have raised awareness that economic development that respects the environment, in which we live, is necessary [8, 9].

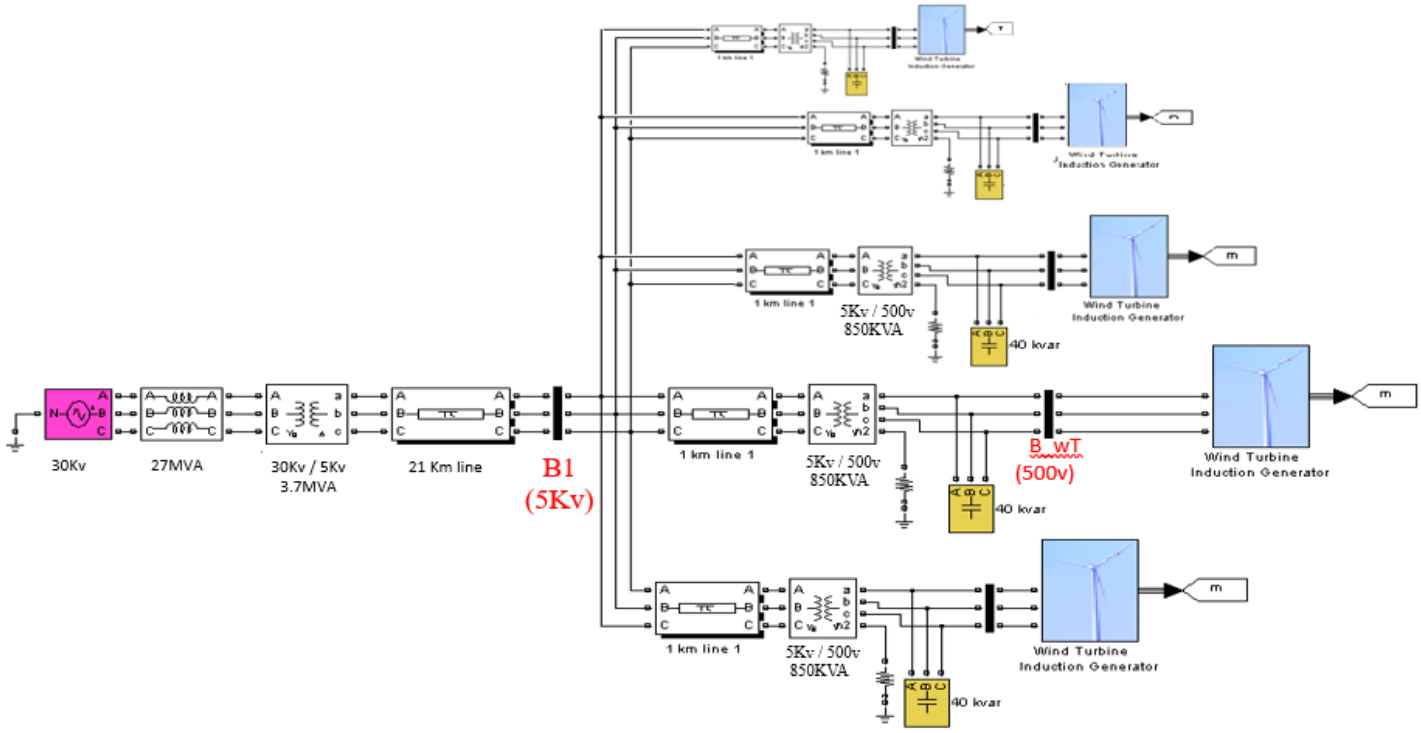


Figure 4: Typical configuration of a hybrid model (wind-diesel)

5. Mechanical regulation of wind turbine speed

From equation (2), we can see that the power produced increases indefinitely with the speed of the wind, which is not the case in reality. The safety devices of the wind machine ensure that the expression of power is distributed over different ranges of wind speeds (Figure 3).

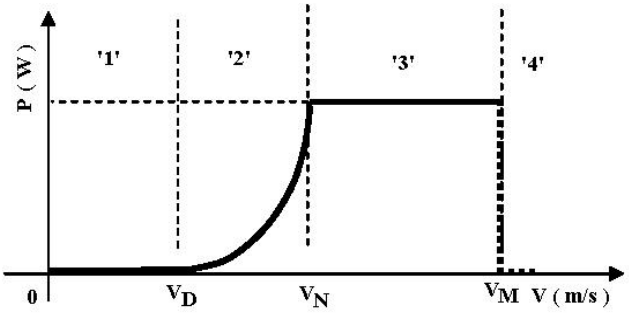


Figure 3: Power response of a stationary wind generator

Figure3 shows that the wind turbine starts to generate power at a cut-in-speed of VD and reaches its rated capacity at VN and continues to produce rated power up to a wind speed of VM. At higher wind speeds, the power yield decreases, as can be seen from Figure 3[10].

6. The impact of wind turbine on electrical network

Traditional production is based on conventional thermal or nuclear power stations as well as on hydraulic power stations. This type of electricity production is flexible in use (due to the availability of primary energy) and meets the technical criteria for safeguarding the network, while renewable energies are highly dependent on the availability of the primary source and do not or hardly meet the criteria for safeguarding the network or for electrical pollution. In this section, the problems induced by the integration of wind generators into the networks, the various constraints for connecting centralized production to the network and finally the prospects for improving the integration of these wind generators into the network will be developed. The incorporation of wind turbines into the networks poses several problems. In this section, the various effects of their integration into networks will be introduced. The flow of current in this conductor will create a voltage drop ΔV, this voltage drop can be expressed as a function of the active and reactive powers passing through the conductor, either:

$$\Delta V \approx \frac{rP + xQ}{V_2} = V_1 + V_2 \qquad (6)$$

The voltage is therefore a local quantity because it differs at every point of the network. This quantity is therefore regulated locally.

It can therefore be seen that the voltage is dependent on fluctuations in power and on the impedance values involved depending on where it's on the network. Often by zone on the network, by producing or absorbing the reactive energy using the production groups, capacitors, or FACTS systems (flexible alternative current transmission system's)

On the transport network x >> r, then:

$$\Delta V \approx \frac{xQ}{V_2} = V_1 - V_2 \qquad (7)$$

Rapid fluctuations in small voltage amplitudes are called flicker. In the case of wind turbines, these variations are due to fluctuations in wind speed, the mechanical limits of the wind turbine (e.g. pitch control) and the shadow effect caused by the

passage of the blades in front of the mast. Fixed speed wind turbines are the most unfavorable from the point of view of these phenomena. The wind turbine technology that is best suited to limiting the impact on the grid of these variations is that completely interfaced with the grid via power electronics, thus allowing a certain decoupling between the turbine and the grid. Wind turbines, interface to the network via fully controlled converters equipped with IGBT transistors, currently the most widely used, generate high frequency harmonics (kHz), but these harmonics can be fairly easily limited unlike partially controlled converters, equipped with thyristors, generating low frequency harmonics requiring substantial filters to eliminate them. Wind generators, like most decentralized generators, are very sensitive to network disturbances and tend to disconnect quickly during a voltage drop or during a frequency variation [7, 8].

7. Behavior of an autonomous WDHS at different wind turbine speeds

Wind-Diesel Hybrid System (WDHS) can be a solution for remote areas of a large interconnected grid. Now the two power sources tested before are synchronized, in order to keep the voltage and frequency of our autonomous WDHS constant, with load variation and wind disturbance. Wind turbines convert the kinetic energy of the wind into mechanical energy and then into electricity. The blades of the wind rotor capture some of the energy contained in the wind and transfer it to the hub that is attached to the shaft of the wind turbine.

This then transmits mechanical energy to the electrical generator, which converts mechanical energy into electrical energy. The generator can then be linked directly or indirectly to the grid.

The combination directly with the network, which is the case in our study, requires the generator to run at "fixed" or very weakly variable speed by playing on the slip of the asynchronous machine. If the generator is indirectly linked to the grid, the introduction of power converters between the generator and the grid results in a decoupling between the frequency of the electrical network and the rotational speed of the electrical machine. This allows variable speed operation using of a synchronous, asynchronous or even special machine.

In this work, the operation of the complete system presented in figure 4, has been simulated in the MATLAB \ SIMULIMK environment where the power output by the wind turbine depends nonlinearly on the wind speed (Figure 5). The performance of the system has been considered with the parameters of deferent principal component of this model presented in the following table.

Table 2: Principal parameters of WDHS

Wind turbine	V_{out} = 575V / f = 50Hz
	P = 3Mw
	Rs (p.u) Ls (p.u)
	0.004843/0.1278 , 0.004377/0.1791
	/ F_F = 0.01 in
	Γ_L = 6.77, / C_i = 5.04 (p.u)
	Np = 3
Diesel generator	**V_{out} = 5kV, / f = 50Hz**
	P = 2MW and Q = 500VAR
	Vdc = 45V, / Rst = 0.0036 (p.u)
	Np = 2

The WDHS system chosen in this work uses a 30kV with a 27MVA the same a parameter of the diesel generator where it is connected in a B1 bus bar, and the wind turbine is connected in the B_WT side, where it maintains a voltage of 500 V and a power of 850kVA delivered by generator of the wind turbine, the proposed load varies between 50kW and 17MW.

The characteristic of the wind turbine presented in figure 5 gives the wind speeds ranging from 4m/s to 12m/s, this shows that the position of the maximum of the power depending on the rotational speed changes with the wind speed. If the wind speed is for example fixed in 10 m/s, the power supplied by the turbine will describe the curve drawn by red color, and the nominal wind speed giving the nominal mechanical power (1pu = 3MW).

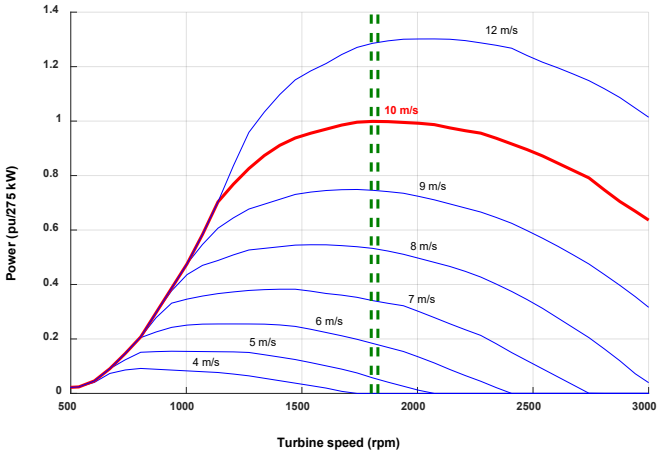


Figure 5: Characteristics of the wind turbine

The aim of variable speed is to adapt the speed of the turbine to the speed of the wind to continuously work at maximum power. In this context, our study is based on a fixed speed of around 10 m/s that show in figure 6.

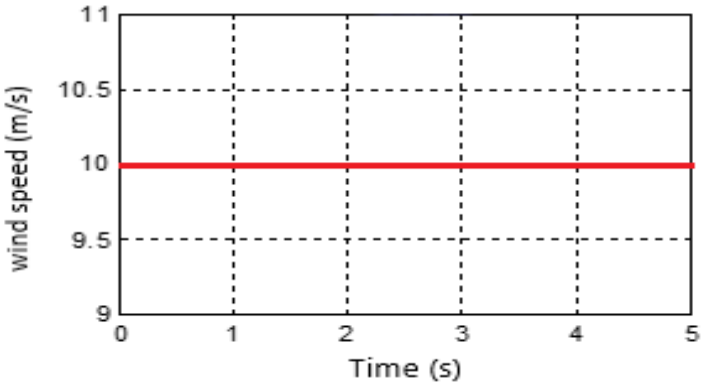


Figure 6: Wind speed fixed at 10 m/s

The simulation results show the proper functioning of our WDHS. At the start of the simulation, the wind turbine begins to provide active power but remains insufficient. To compensate for the lack of power, the diesel generator is primed. The regulation of voltage and frequency in the case of constant wind is ensured by the diesel generator.

Hence, the speed and mechanical torque of the turbine shown in figures 7 and 8 respectively, indicates that the mechanical torque generates variations which are due to the mechanical part of the wind turbine, depending on the construction, location or other external phenomena, and the electromagnetic couple is oscillated at the beginning until the instant 3s, where it stabilizes.

Figures 9 and 10 present the active and reactive powers measured and references by blue and red colors respectively, the results obtained show that the oscillations presented before the instant 0.5s and after the instant 1.5s are caused by different wind variations, which make the energy system unstable

The voltage and power delivered by the Asynchronous machine of the wind turbine are displayed in figures 11 and 12, when the power is beginning to stabilize at time 1,5s and the voltage remains sinusoidal throughout the simulation period.

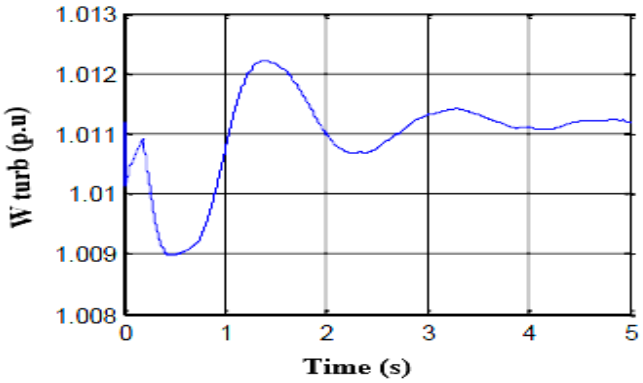


Figure 7: Rotor speed of turbine

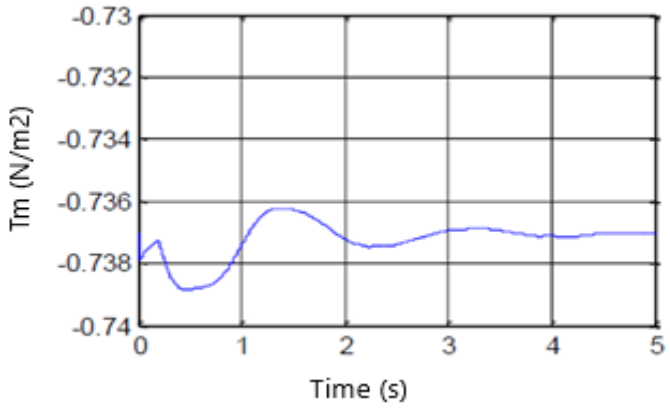


Figure 8: Torque of turbine

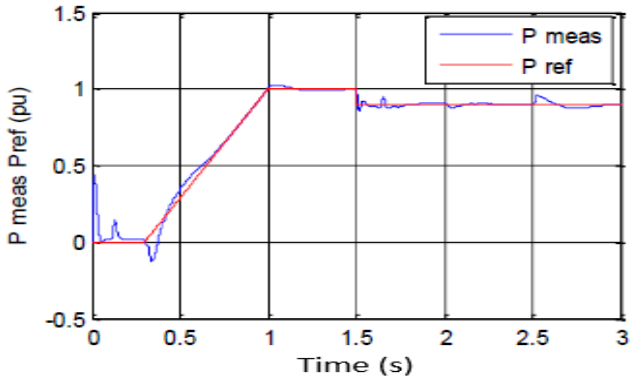


Figure 9: Active energy of the system

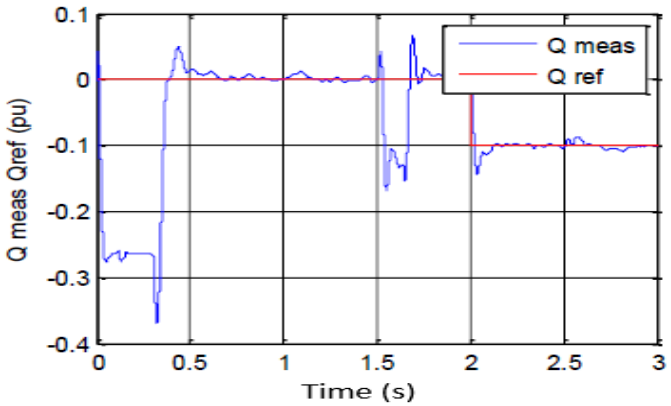


Figure 10: The system energy reactive power

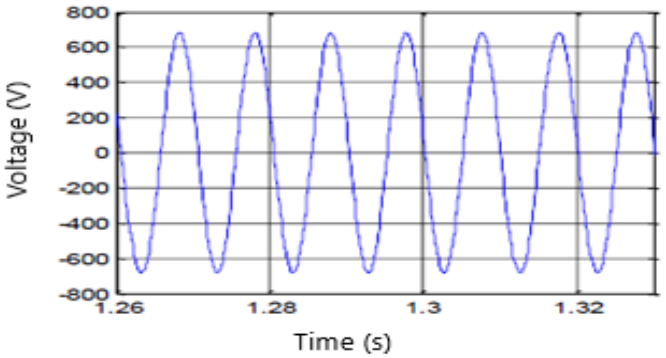


Figure 11: Asynchronous machine voltage of turbine

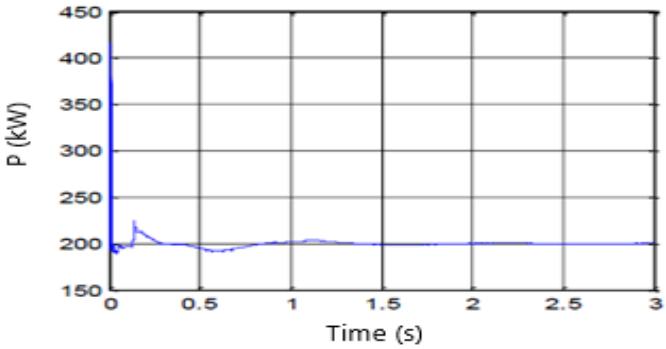


Figure 12: Asynchronous machine power

When the WDHS underwent a variation of the active load and the wind speed and was still weak. In this case, the diesel generator supports the power demand. On the hand, at low wind speeds both the diesel generator is required to feed the load, on the other hand When the wind power exceeds the load demand, the diesel generator must be put into the under-powered state. In this case, the diesel generator excitation system is used as a grid voltage controller to maintain it at its nominal value. The load is considered constant throughout the simulation period. The simulation results presented by different figures like (figure 13, 14, 15, 16, 17 and 18) showed that the intermittent nature of wind turbines imposes an issues that affect the power quality of the main electricity grid, to ensure an optimal electrical power, it's important the use the FACTS systems (Flexible Alternative Current Transmission Systems) to maintain controllable reactive power flow between the generating units and the utility network. Dynamic compensation of reactive power is an effective means of safeguarding power quality as well as voltage stability.

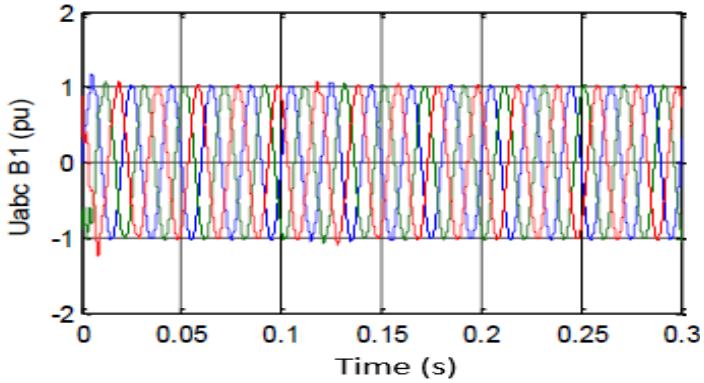


Figure 13: Busbar connection voltage in B1

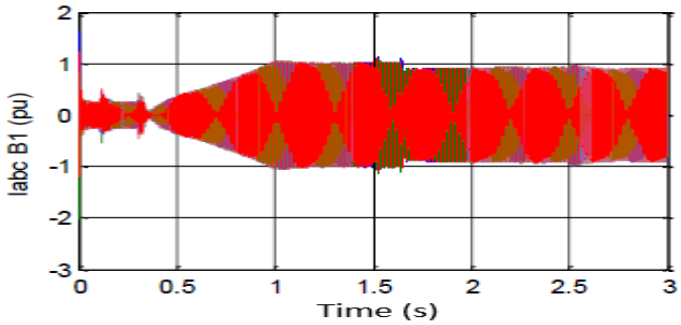


Figure 14: Busbar connection current in B1

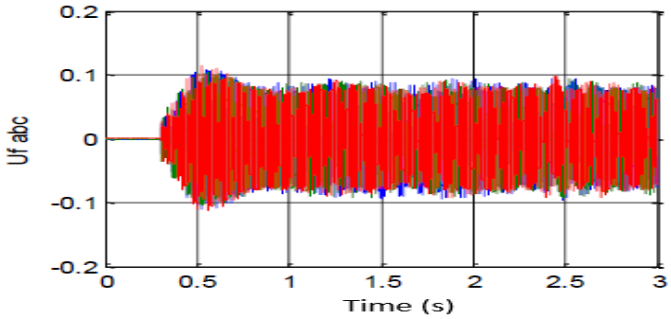


Figure 15: Busbar connection voltage in B_wt

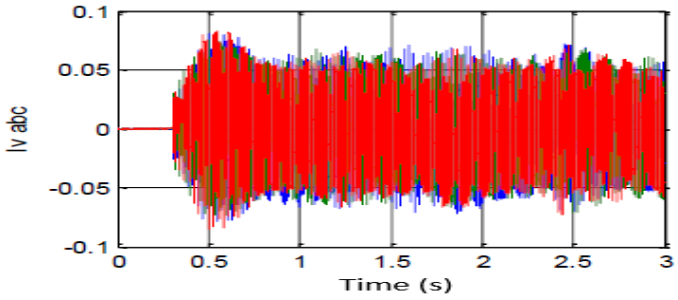


Figure 16. Busbar connection current in B_wt

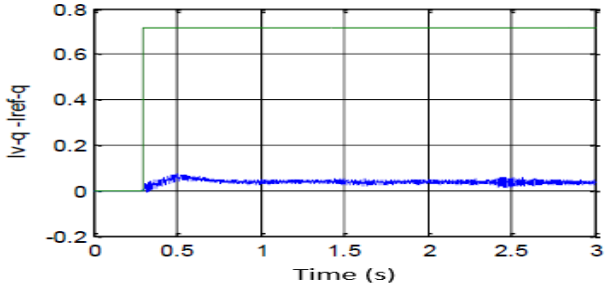


Figure 17: The current sequence Iq in B_wt with refference current

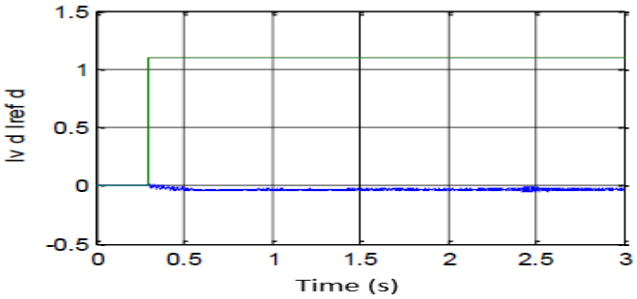


Figure 18: The current sequence Id in B_wt with refference current

8. Conclusion

The significant expansion of wind energy requires solving a series of technical economic questions. A current energy source, wind energy is the most advanced technology due to its installed power and recent improvements in power electronics and control. In addition, the applicable regulations favor the increasing number of wind farms due to the attractive economical reliability.

The system studied uses a set of very different equipment intended to coexist in a cogeneration structure supplying an isolated load. To better understand the nature of the interactions between the different subsets of this system, an approach was adopted to study the dynamics of each subsystem. The connection of a diesel generator (DG) to distribution networks greatly affects the performance and reliability of the networks.

This paper focuses on the impact of DGs on electrical network control and protection arrangements. Thus, the study of hybridization was made by following the key elements of the system: the wind turbine, and the connection of the assembly. At first glance, we placed this subject in its socioeconomic context. The latter is marked by the findings of environmental deterioration caused by the ever-increasing use of fossil fuels.

In addition, the difficulties of transporting the fuel used to supply the DGs are a classic means of producing energy in isolated areas, and clearly show the need for a renewable energy source. Thus two complementary issues were addressed such as the continuity of electricity production accompanied by good quality to be supplied to sensitive customers. However, our studies on this vast problem which is the integration of renewable sources, namely wind turbines in the conventional network, remains in a dynamic of improvement.

The technological development in wind power generation have solved several problems presented in large wind farms connected to low-power transmission grids

A dynamic study over longer periods, taking into account energy storage and switching sources, could also be added to the present models. The integration of the wind power can be encouraged by the development of hybrid system, combining wind energy with diesel generator, with an optimized management of energy system.

The impact of the wind on the quality of the electrical energy is decisive because it disturbs the electrical parameters of the network during integration. For this, the strengthening of control systems is mandatory or the use of smart control methods is strongly recommended. Finally, to ensure the optimal quality of

electrical energy and given the randomness of the wind, we have integrated FACTS which will be the subject of the next paper

References

[1] M. N. Tandjaoui, C. Benachaiba, O. Abdelkhalek, "Role of power electronics in grid integration of renewable energy systems", Journal of Electrical Engineering, **16**(1), 2016.

[2] S. A. Zegnoun, M. N. Tandjaoui, M. Djebbar, C. Benachaiba, B. Mazari, "Power quality enhancement by using D-FACTS systems applied to distributed generation", International Journal of Power Electronics and Drive System, **10**(1), 330~341, 2019.

[3] M. N. Tandjaoui, C. Benachaiba, O. Abdelkhalek et B. Denai, "Renewable energy and its impact in power quality of distribution systems", SSC5'2012, College of Applied Sciences. Umm Al-Qura University. Makkah, April 16-18, 2012.

[4] M.N. Tandjaoui, M. Haidas, C. Benachaiba, "Intégration des éoliennes pour la Production d'Energie Electrique en Algérie - Site Tindouf- " , 10th international Meetings on energucal physics, Journal of Scientific Research, **1**, 183-188, University of Bechar, Algeria, 2010.

[5] S. M. Shinde; K. D. Patil; S. S. Khairnar; W. Z. Gandhare, "The Role of Power Electronics in Renewable Energy Systems Research and Development", 2nd International Conference on Emerging Trends in Engineering & Technology, 16-18 December 2009, Nagpur, India, IEEE Xplore, 22 Junuary 2010.

[6] M. N. Tandjaoui, C. Benachaiba, M. Saidi, Y. Mouloudi, B. Denai and O. Abdelkhalek, " Apport des éoliennes dans les réseaux électriques-Utilisation d'un SVC", 11th international Meetings on energucal physics, University of Bechar, Algeria, 2011.

[7] F. Blaabjerg; F. Iov; R. Teodorescu and Z. Chen, "Power Electronics in Renewable Energy Systems", 12th International Power Electronics and Motion Control Conference, 30 August 2006 - 01 September 2006, Portoroz, Slovenia, IEEE Xplore: 10 February 2009.

[8] M. N. Tandjaoui, C. Benachaiba, O. Abdelkhalek, B. Denai and Y. Mouloudi, "The Impact of Wind Power Implantation in Transmission Systems", TerraGreen 13, Libanon, Energy Procedia, Elsevier, 2013.

[9] F. Lov, M. ciobotarn and F. Blaabjerg, "power Electronics control of wind energy in distribution power systems", 11th International Conference on Optimization of Electrical and Electronic Equipment, 22-24 May 2008, Brasov, Romania, IEEE Xplore: 19 August 2008.

[10] Himri, Y., Rehman, S., Draoui, B., & Himri, S, "Wind power potential assessment for three locations in Algeria", Renewable and Sustainable Energy Reviews, **12**(9), 2495-2504, 2008.

[11] F. Boukli Hacène, N. Kasbadji Merzouk, L. Loukarfi, "Analyse statistique et élaboration d'un atlas éolien de la vallée du Cheliff ", Revue des Energies Renouvelables, **10**(4), 583 – 588, 2007.

[12] Himri, Y., Rehman, S., Mostafaeipour, A., Himri, S., Mellit, A., Merzouk, M., & Merzouk, N. K, "Overview of the Role of Energy Resources in Algeria's Energy Transition", Energies, **15**(13), 4731, 2022.

[13] O. Gergaud, B. Multon, H. Ben ahmed, "Modélisation d'une chaîne de conversion éolienne de petite puissance", Electrotechnique du Futur, 17-22, Nancy 14-15 novembre 2001.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 8, No. 5, 35-54 (2023)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Condition Assessment of Medium Voltage Assets: A Review

Eilin Gómez Mesino[*,1], Joaquín Caicedo[2], Miguel Mamaní[1], David Romero Quete[3], Andrés Cerón Piamba[4], Diego García Gómez[4], Guillermo Aponte Mayor[4], José Caicedo Erazo[4], Wilmar Moreno López[5], Edward Jay[5], Andrés Romero Quete[1]

[1]*Instituto de Energía Eléctrica, Universidad Nacional de San Juan–CONICET, San Juan, Argentina*

[2]*Facultad de Ingeniería, Universidad Distrital Francisco José de Caldas, Bogotá, Colombia*

[3]*Departamento de Ingeniería Eléctrica y Electrónica, Universidad Nacional de Colombia, Bogotá, Colombia*

[4]*Grupo de investigación GRALTA, Escuela de Ingeniería Eléctrica y Electrónica, Universidad del Valle, Cali, Colombia*

[5]*Empresa de Energía del Archipiélago de San Andrés, Providencia y Santa Catalina, EEDAS S.A. E.S.P., San Andrés Islas, Colombia*

ARTICLE INFO

Article history:
Received:29 July, 2023
Accepted:25 September, 2023
Online: 30 October, 2023

Keywords:
Asset management
Condition assessment
Condition monitoring
Distribution network
Literature review
Medium-Voltage

ABSTRACT

Condition assessment of medium voltage assets is essential to ensure reliability and cost-effective operation of power distribution networks. This article presents a literature review of condition assessment of medium voltage assets related to a distribution system in a non-interconnected zone in Colombia, namely, power transformers, photovoltaic systems, switchgear, lines and cables, and instrument transformers. Advanced search rules are formulated to obtain bibliographic records of relevant academic literature from the database Scopus. The retrieved data are analyzed quantitatively to provide insights on the current state of research on the topic. Next, the most relevant academic papers for each medium voltage asset are selected and analyzed in a critical review along with more diverse literature including standards, technical reports, and white papers obtained through complementary searches. The results of the review show that several approaches have been formulated for condition assessment of medium voltage assets, ranging from traditional diagnostic methods to advanced artificial intelligence-based approaches. Moreover, research on some assets is already mature including power transformers, and photovoltaic systems, whereas other assets have been incipiently studied such as distribution and instrument transformers. Therefore, the need for deeper condition assessment research for these critical assets is highlighted. Research gaps are identified in the standardization and integration of condition assessment tools for distribution system operators.

1. Introduction

This paper extends a work presented in IEEE Biennial Congress of Argentina, ARGENCON 2022, [1]. This extended review covers other important Medium-Voltage (MV) assets operating in distribution networks. The review is a key component of a government-funded research project in Colombia aimed at implementing a digital platform for asset management of components at the archipelago San Andrés, Providencia, and Santa Catalina distribution network, in agreement with ISO 55000 [2].

In the islands, the distribution network consists of distributed generators (diesel and solar) which supply power to MV feeders at 34,5 kV and 13,2 kV, and Low-Voltage (LV) end customers at 208/120 V. The main assets in the distribution network are:

- Power Transformers (PT)
- Photovoltaic (PV) systems
- Distribution Transformers (DT)
- MV switchgear
- MV lines and cables
- Instrument Transformers (IT)

[*]Corresponding Author: Eilin Gómez Mesino, Universidad Nacional de San Juan, Av. Libertador Gral. San Martin 1109, egomez@iee.unsj.edu.ar

https://dx.doi.org/10.25046/aj080505

In this context, this paper offers a review of the main technical aspects related to condition assessment of distribution network assets. The study focuses on a non-interconnected grid.

The paper is structured as follows. The next subsections describe the methodology adopted for the literature review and the taxonomy to structure the analysis. Section 2 introduces the review of condition assessment of PT. Similarly, Section 3 reviews condition assessment of PV systems, followed by the review of DT, switchgear, lines and cables, and IT, in Sections 4 to 7. Section 8 outlines the discussion of results. Section 9 explores prospects for future research. Section 10 states the conclusions of the paper.

1.1. Methodology for the literature review

The literature review methodology is based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [3]. The advanced search rules illustrated in Figure 1, are applied in titles, abstracts, and keywords of the citation database Scopus to identify the most relevant literature related to condition assessment of MV assets. On the one hand, concepts related to condition assessment are included in each search rule and synonyms and associated concepts are also used according to Table 1. On the other hand, seven individual search rules are formulated to consider the most relevant assets of the distribution network in San Andrés Islands, i.e., (1) PT, (2) PV systems, (3) DT, (4) switchgear, (5) lines and cables, and (6) IT. In this case, synonym keywords and compound search formulas are defined according to Table 2.

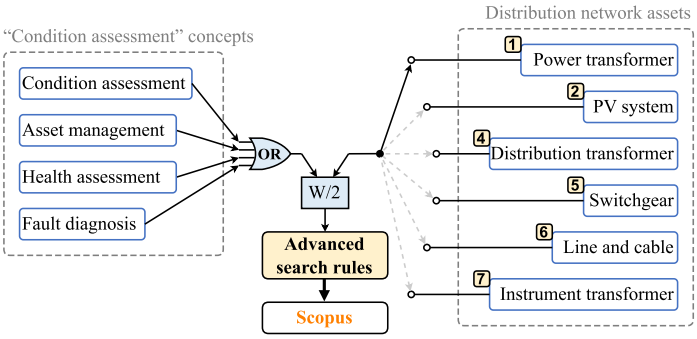


Figure 1: Advanced search rules applied in Scopus on April 1st, 2023.

Table 1: Set of synonym keywords related to condition assessment concepts.

Concept	Set of synonym keywords
Condition assessment	"condition assessment" OR "condition monitoring" OR "remaining useful life" OR "performance index"
Asset management	"asset management"
Health assessment	"health monitoring" OR "health index"
Fault diagnosis	"failure detection" OR "failure diagnosis" OR "fault detection" OR "fault diagnosis"

Table 2: Set of synonym keywords and compound search formulas for MV assets.

Asset	Set of synonym keywords / compound searches
PT	"power transformer"
PV system	"photovoltaic generat*" OR "pv generat*" OR "photovoltaic system" OR "pv system"
DT	"distribution transformer"
Switchgear	switchgear OR "power circuit breaker" OR disconnector OR recloser OR ("circuit breaker" W/2 (distribution OR OR medium-voltage OR mv))
Line / cable	(distribution OR medium-voltage OR mv) W/2 (line OR cable)
IT	"instrument transformer" OR "current transformer" OR "voltage transformer"

Figure 1 shows that the concepts related to "condition assessment" are merge with an "OR" operator, i.e., the terms are considered synonyms and any of them is useful to retrieve the bibliographic records. Moreover, "condition assessment" concepts and "distribution network assets" are associated with the operator "W/2", i.e., one of the terms can be found within a distance of two words from the other. For instance, if "condition assessment" (or "asset management", or "health assessment", or "fault diagnosis") appears in the tittle, abstract or keywords of a bibliographic record, and "power transformer" is found in the same record within a distance of two words, then the record is retrieved.

The number of records retrieved using the search rules on April 1st, 2023, are reported in Table 3. Results show extensive research related to strategic assets such as PT and PV systems, where several recently published literature reviews are found. However, limited research has been performed on condition assessment of other relevant assets such as DT and IT. In this context, this paper offers an overview of condition assessment for both widely and incipiently studied assets. It analyzes previously less explored aspects in detail to provide a comprehensive review of key technical considerations.

Table 3: Number of records retrieved with the proposed search rules.

Search rule	Jour. articles	Conf. papers	Reviews	Others	Total
PT	349	592	29	6	976
PV system	124	91	15	7	237
DT	12	37	2	0	51
Switchgear	26	53	3	0	82
Line / cable	42	68	2	1	113
IT	6	20	0	0	26

Figure 2 presents the evolution from 2003 to 2022 of the number of publications retrieved using each search rule. The figure shows an increasing trend across nearly all topics, with a particular emphasis on PV systems, which has seen a significant surge in publications over the last five years. A significant increase occurred in PT condition assessment from 2006 to 2008, and then the trend is slightly positive. Other moderately studied topics show an increasing interest such as lines and cables, switchgear, and DT. Finally, research is very incipient in IT condition assessment.

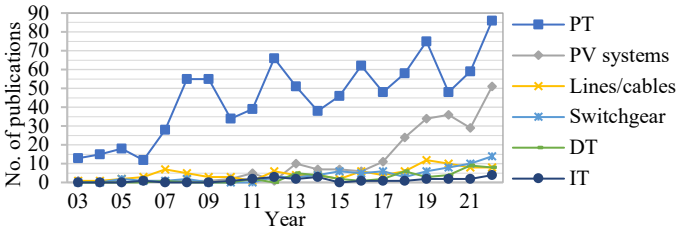


Figure 2: Publications related to condition assessment of MV assets over time.

A set of relevant publications are selected from the retrieved bibliographic records to perform the critical review of each asset. In the case of extensively studied topics, i.e., PT, and PV systems, the analysis is based on recently published reviews and a limited selection of additional relevant publications to observe aspects that required further analysis. The review of the remainder assets, i.e., DT, switchgear, lines and cables, and IT, is based on relevant publications selected by reading and assessing titles, abstracts, and conclusions of the retrieved bibliographic records. Moreover, references are complemented with related standards, technical reports, and white papers.

1.2. Taxonomy for condition assessment in distribution networks

The following structure is adopted to analyze each asset:

- A short description of the main asset's components, subsystems and features is given.
- Fault diagnosis and condition assessment approaches for each asset are described and the references are classified according to the type of test and diagnostic technique.
- Indices to monitor the condition of assets are analyzed.

The following subsections provide a broader description of the aspects covered in fault diagnosis and condition assessment of the MV assets, along with a taxonomy of test types and diagnostic techniques. Aspects related to indices are also introduced.

a) Fault diagnosis and condition assessment

Fault diagnostic and condition monitoring techniques for power assets is a wide and very dynamic research and innovative field. The variety of tests and techniques ranges from traditional and widely accepted methods, to advanced and not completely validated approaches. The main differences are based on the possibility that offers the technique to obtain a diagnosis with the asset in-service or off-service, i.e., the asset operating status. The diagnostic test nature differs according to the parameter to be analyzed, e.g., gas chromatography, which is a method from analytical chemistry commonly used to assess the insulating oil condition. In addition to chemical tests, mechanical or electrical tests can also be applied to analyze other oil features.

Once the tests or inspections have been carried out on the asset, the acquired data must be analyzed for diagnosis. This analysis can be based on expert judgment, who, like doctors in medicine, examine the data and offer a conclusion on the health status or presence of failures in the asset. This type of expert knowledge-based approaches also includes computational methods based on rules such as expert systems, and fuzzy logic.

Alternatively, analytical methods are applied by inferring a set of mathematical equations to model the phenomena involved in the asset condition as close to reality as possible. Thus, analytical models offer very accurate diagnostics, but require comprehensive understanding of the asset, in terms of specific components and operating modes, as well as the status of asset parameters.

Finally, the data acquired from tests can also be analyzed automatically with Artificial Intelligence (AI)-based techniques including Support Vector Machines (SVM), decision trees, and Artificial Neural Networks (ANN), among others. AI-based approaches provide flexible and efficient performance, which is of especial interest in large and complex systems with thousands of assets. However, these AI-based methods, also referred to as data-driven approaches, need large amounts of training and validation data, and are prone to underfitting and overfitting.

Figure 3 summarizes the process described above for fault diagnosis and condition assessment, i.e., 1) data from tests or inspections are acquired, and 2) diagnostic techniques are applied for data analysis. The figure also depicts the taxonomy of tests and diagnostic techniques discussed earlier. This taxonomy is used throughout the comprehensive review to classify tests and methods reported in the literature.

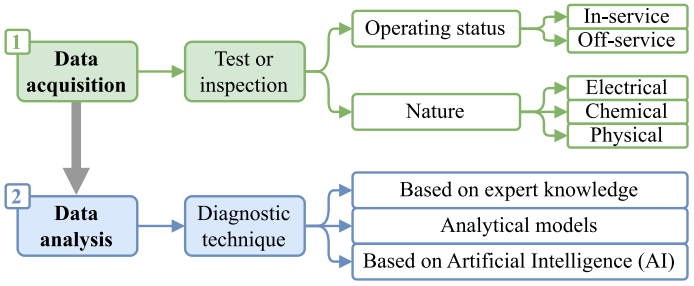


Figure 3: Taxonomy for fault diagnosis and condition assessment techniques.

b) Indices for condition assessment

In the international standard ISO 55000 [2], the main indices related to condition assessment of power assets are the Health Index (HI), consumed life and remaining life indices, performance indices, among others. Therefore, these indices are identified and described in each asset review.

2. Power Transformers (PT)

PT play a crucial role in power systems due to their strategic significance and higher costs compared to other power assets. In distribution networks, PT are usually referred to as primary DT, which connect MV networks to distributed generators, or, in interconnected systems, with High-Voltage (HV) transmission networks. Hence, the PT failure represents the impossibility of providing many customers with electrical energy. In brief, a PT comprises five subsystems. The main subsystem is the Active Part (AP) made up of windings, core, insulating paper, clamp structures, yoke, oil, and a tank containing the mentioned elements. The remainder four subsystems are (i) Bushings (BH) and terminals, (ii) On Load Tap Changer (OLTC), (iii) Conservator Tank (CTk) and breather, and (iv) the Cooling System (CS) made up of fans, pumps, controls, and radiators.

There is great interest in researching techniques for PT condition assessment. According to Table 3, a significant number of 976 bibliographic records were obtained using the search rule reported in Section 1.1. Among those publications, comprehensive and recently published literature reviews are found, e.g., [4–10]. Therefore, the following analysis summarizes the key findings of those reviews. Also, the review is complemented with standards [11–15], and selected papers [16–48].

2.1. Fault diagnosis and condition assessment of PT

Condition monitoring is an essential aspect of PT management. It involves the use of sensors and other monitoring devices to measure several parameters, such as temperature, oil level, and Partial Discharges (PD), which allows for detecting faults or changes in the asset performance. Advanced methods such as Dissolved Gas Analysis (DGA) are also employed to track the condition of transformer insulation systems. These methods allow early detection of faults and diagnosis of the PT status to take proper maintenance actions, thus reducing service interruptions and economic losses associated with its unavailability. Data sources, in both fault diagnosis and condition assessment, are acquired from the PT (primary source), power system, and environment. For this purpose, the main tests for monitoring PT parameters by subsystem and type are reported in Table 4.

Table 4: Classification of tests on PT and diagnostic techniques.

Data acquisition								Data analysis					Refs.
Test or inspection	Subsystem	Parameter	Type of test					Diagnostic technique	Purpose	Type of analysis			
			In-service	Off-service	Electrical	Chemical	Physical			Exp. know.	Analytical	AI	
DP	AP	Number of β-glucose monomers		✓		✓	✓	Viscometric method and Mark-Houwink-Sakurada equation	Direct measurement of DP value		✓		[12]
Furan analysis	AP	Furanic compound in oil (2-FAL, etc.)	✓			✓		Empirical relationships obtained in laboratory	Indirect estimation of DP value	✓			[4,9,20,39]
Load history	AP	Power (MVA)	✓		✓			Numerical models for hot-spot temperature calculation; analytical expression for pyrolysis, hydrolysis, and oxidation aging processes	Indirect estimation of DP value		✓	✓	[25,28,35]
Ambient temperature	AP, CS	Temperature (ºC)	✓				✓						
Winding/top oil temp.	AP	Temperature (ºC)	✓				✓						
Terminal voltages	AP, BH	Voltage (V)	✓		✓			Standards, guidelines, and manufacturer recommendations to assess the PT electrical performance	Detect deviations from normal, recommended, or standardized values	✓			[6,7,9,11,14]
Core to ground current	AP	Current (mA)	✓		✓								
Dissipation, power factor, capacitance	AP, BH, OLTC	Tan δ, power factor, capacitance (F)	✓	✓	✓								
Excitation current	AP	Current (A)	✓		✓								
Leakage current	AP, BH	Current (A)	✓										
Bushing voltage	BH	Voltage (V)	✓		✓								
Winding resistance	AP	Resistance (mΩ)		✓	✓								
Turns ratio	AP	Ratio		✓	✓								
Recovery voltage	AP	HV impulse (kV)		✓	✓								
Polarization index	AP	Insulation res. ratio		✓	✓								
Contact resistance	OLTC	Resistance (mΩ)		✓	✓								
DGA	AP, BH, OLTC	Gases dissolved in oil (ppm)	✓			✓		Dornenburg ratio; Key gas; Rogers ratio; trend analysis; Duval triangle; Duval pentagon; fuzzy logic	Early detection of electrical faults, insulation status assessment	✓		✓	[7,15,19,37,38]
FRA	AP	Admittance (S), impedance (Ω), transfer function	✓	✓	✓			Direct interpretation of the curve; comparison with previous measurements in the same or similar PT; analysis of harmonic components; wavelet; expert system; ANN	Detection of winding deformation and core movement	✓	✓	✓	[13,16,23,29]
DFR	AP	Transfer function (10 Hz to 1 MHz)		✓	✓			Time- and frequency-domain analysis; dielectric spectroscopy; principal component analysis; fitting models; ANN	Assessment of insulating materials frequency response	✓		✓	[17,21,27,32,36]
Oil characteristics	AP, OLTC	Moisture, acidity, color, breakdown voltage	✓			✓	✓	Direct interpretation and comparison of results from acidity, interfacial tension, dielectric, strength, and viscosity tests, Karl Fischer titration, etc., with standardized values	Evaluation of insulating oil status	✓		✓	[6,9,11,14,33]
Acoustic signals	AP, OLTC	Acoustic emission (dB)	✓				✓	Spectral analysis; time waveform analysis; pattern recognition; acoustic imaging	Health assessment of core and winding structures		✓	✓	[18,26,34,40]
PD	AP, BH	Electric charge (pC)		✓	✓			Analysis of direct PD measurements; acoustic PD measurement for source location; optical detection of the light emitted by PD	Determine the location, magnitude, and type of PD in insulation materials	✓	✓	✓	[22–24,31]
Ultra-High Frequency (UHF) PD detection	AP, BH	Electric charge (pC)		✓	✓								
Infrared thermography	AP, BH, OLTC, CTk, CS	Thermal Image (ºC)	✓				✓	The temperature variations observed in the infrared image are interpreted in the context of the PT operation and maintenance history	Identify temperature rise, loose connections, insulation failure	✓			[30,41]
Visual inspection	AP, OLTC, BH, CTk, CS	Oil levels, gas pressure, fan and pump condition, corrosion	✓				✓	The inspector records and classify observations of the components, including any signs of damage, wear, corrosion, etc. The severity of the damage is compared to standards	Identify observable anomalies	✓			[11,14]

From the tests in Table 4, it is worth noting that in-service condition monitoring allows continuous evaluation of the operation and offers the possibility of assessing the states that may impact the lifespan of the PT, whereas off-service tests require disconnecting the unit from the power grid, limiting its use to maintenance scenarios, planned inspections, or when there is a direct suspicion of failure.

Regardless the type of applied tests mentioned above, resulting data must be processed to generate information about the asset status. Consequently, diagnostic techniques are applied to assess the condition of PT and find any defects or malfunctions. For instance, these techniques include Frequency Response Analysis (FRA), which can detect winding deformation and core movement, or the analysis of tan δ, capacitance, and dielectric response measurements, among others, which can identify insulation degradation. In this regard, several diagnostic techniques have been proposed and the most relevant are also reported in Table 4 following their related tests.

Furthermore, analysis of data from traditional tests to diagnose the PT electrical performance, for instance, excitation current, power factor, and winding resistance, has rarely been addressed in recent research because of its straightforward interpretation and broad understanding and acceptance by PT managers. In contrast, most of the current research on PT condition assessment is focused on approaches or methodologies to improve the interpretation of the results, especially from FRA, Dielectric Frequency Response (DFR), DGA, acoustic signals, PD detection, and indirect Degree of Polymerization (DP) estimation. These recent approaches and

methodologies are aimed at overcoming the difficulties related to accuracy, uncertainty, interpretation, ambiguity, and discrepancies related to traditional and standard diagnosis.

For instance, in FRA and DFR, an option consists of comparing the magnitude and phase angle of the frequency response on a polar plot, which can then be analyzed by an expert. Alternatively, numerical models are used to simulate the behavior of transformers under different conditions. For instance, the effects of winding deformation, core movement, moisture content in oil and other defects on the transformer frequency response can be considered with these models. The simulated and measured frequency response are then compared to detect the presence of defects. A third alternative for analyzing the results of FRA and DFR is the use of machine learning techniques including decision trees, and ANN to identify patterns and predict the likelihood of defects or failures.

Similarly, DGA, which aims at identifying electrical faults, can be conducted with traditional analysis using Rogers ratio, key gas, Dornenburg ratio, Duval triangle and pentagon. Also based on expert knowledge, fuzzy logic allows the diagnosis of multiple failures, a more realistic representation, and quantification of the probability/severity of each damage. AI-based techniques are also applied to DGA interpretation. For instance, the use of ANN has the benefit of not requiring experts for the interpretation of DGA Approaches based on AI offer new possibilities in the application of DGA due to their automatic learning and data analysis capabilities. These approaches recognize different patterns of gas concentrations, which indicate the type of failure or problems within the transformer through the processing of large volumes of historical DGA data. These data are automatically identified and classified in real time or in certain time intervals, allowing for improvement and effectiveness in the early detection and diagnosis of failures. AI can recognize complex patterns and sudden deviations in gas concentration that are not easily noticed by experts. It is crucial to continue improving AI-based techniques to incorporate new research findings, such as the C_2O/CO ratio, for identifying failure patterns in the insulating paper.

The discussion above demonstrates that research on condition assessment of PT is a very important and fertile area of work, especially with the emergence of powerful AI-based tools to analyze large amounts of data that this type of asset generates.

2.2. Indices for condition assessment of PT

A crucial indicator in asset management of PT is the Health Index (HI) because it is required to perform fault diagnosis, life prediction and risk assessment. A PT-HI is a quantitative indicator that considers the results of asset performance observation, on-site inspections, and laboratory tests, to efficiently manage the asset. In 2017, a complete analysis of the main PT-HI calculated using mathematical equations/algorithms or expert judgment is reported in [10]. Such a paper analyzes the period 2008-2017, including the first approach published in 2008 [42]. Most of the expressions are based on the weighted summation of multiple input parameters, i.e., multi-criteria methods. In general, ranges of input parameters (collected through tests discussed above) are based on limits recommended by IEC and IEEE standards, and expert knowledge from utilities. The number of input parameters and weights varies from approach to approach. A complete list of the most used parameters is reported in [10].

Advanced techniques for PT-HI calculation are based on fuzzy logic [43–46], and ANN [47,48]. From a practical point of view, utilities prefer weighted summation approaches because of their simplicity and higher flexibility compared with advanced methods that require specialized knowledge to implement modifications.

3. Photovoltaic (PV) systems

PV systems are increasingly penetrating distribution networks as an alternative means of producing electricity motivated by the necessity of more sustainable energy sources than fossil fuels. PV systems exhibit rapid expansion within the realm of renewable energy technologies, primarily attributed to the enduring and readily accessible source of solar energy at no cost [49]. According to the International Energy Agency, the global PV capacity grew from 50 MW in 1990 to 946 GW in 2021 [49,50].

Depending on usage, PV systems can operate stand-alone used for remote power or backup applications, or grid-connected to complement the power provided to an installation from the utility grid [51]. Both stand-alone and grid-connected PV systems are mainly made up of (i) an array of solar PV panels which convert sunlight into electricity, (ii) inverters to transform DC into AC power, (iii) protection and switching equipment, (iv) supporting structure and mounting racks, (v) combiner boxes, conductors and grounding, (vi) energy meters and, optionally, (vii) an energy storage system and charge controller [49,51]. These components are exposed to deterioration and failures that make essential the use of condition assessment methods to guarantee sustainable techno-economic operation of PV systems. In the following sections, the aspects above are discussed based on the retrieved literature.

The results obtained with the search rule formulated in Section 1.1 for PV systems show that there is extensive research on the topic, i.e., 237 bibliographic records were retrieved including 15 literature reviews (see Table 3). Moreover, Figure 2 demonstrates an apparent increasing interest on the topic in the last few years. Therefore, emphasis is placed on the analysis of recent literature reviews [52–68]. In addition, the literature is complemented with standards [69–71], selected papers [72–76], and reports [77].

3.1. Fault diagnosis and condition assessment of PV systems

Condition assessment of PV systems is achieved by monitoring performance parameters of the components described above. For instance, these parameters include the operating voltage, current, and power output, which mainly depend on the proper operation of the array of solar PV panels, inverters, and the energy storage system. Integrity of combiner boxes, grounding, and insulation resistance must be also verified for condition assessment. These parameters are measured using instruments such as PV testers, solar power meters, insulation testers, battery analyzers, among other devices. In addition, testing procedures must be applied during installation, i.e., at commissioning, and regularly during the lifespan of the PV system. Standard and widely used and validated testing procedures for PV systems, as well as indication on suitable measurement instruments, are outlined by IEC 62446-1:2016 [69], and IEEE Std 1547-2018 [70], for grid-connected PV systems, and IEEE Std 1526-2020 [71], for stand-alone PV systems. Table 5 summarizes the tests applied on PV systems found in the literature,

Table 5: Classification of tests on PV systems and diagnostic techniques.

Data acquisition								Data analysis					Refs.
Test or inspection	Subsystem	Parameter	Type of test					Diagnostic technique	Purpose	Type of analysis			
			In-service	Off-service	Electrical	Chemical	Physical			Exp. know.	Analytical	AI	
Continuity	Conductors	Impedance (Ω)		✓	✓			Standards, guidelines, and manufacturer recommendations to assess electrical performance of PV systems	Detect deviations regarding normal, recommended, or standardized values	✓			[69–71]
Current (short, operating)	PV array	Current (A)	✓	✓	✓								
Open circuit voltage	PV array	Voltage (V)		✓	✓								
Voltage to ground	Grounding	Voltage (V)		✓	✓								
Insulation resistance	PV array	Impedance (Ω)	✓		✓								
Blocking diode	PV array	Voltage (V)		✓	✓								
Wet insulation resistance	PV array	Impedance (Ω)	✓		✓								
Shade evaluation	PV array	Various	✓	✓			✓	Standards, guidelines, and manufacturer recommendations to assess mechanical (physical) performance of PV systems	Verify initial and operating environmental (shading) and mechanical (connections, supports) conditions	✓			[54,69, 71]
Visual inspection	PV system	Various	✓	✓			✓						
Polarity	Comb. boxes, conductors	Voltage (V)		✓	✓		✓	Standards, guidelines, and manufacturer recommendations for electromechanical (electrical/physical) assessment of PV systems	Verify initial condition of conductors and boxes polarity, and proper operation of protections and inverters	✓			[69–71]
Functional	PV system	Various		✓	✓		✓						
Usable battery capacity	Battery	Capacity (Ah)	✓		✓			Standards, guidelines, and manufacturer recommendations to assess electrical performance of energy storage in stand-alone PV systems	Check the correct operation and requirements of the battery and charger controller	✓	✓		[71]
Recovery	Energy storage	Capacity (Ah)	✓		✓								
Battery charge	Energy storage	Autonomy (h)	✓		✓								
I-V curve	PV array	Voltage (V), current (A)	✓		✓			Handcrafted diagnosis; I-V curve tracer; real-time difference; PV and solar emulator	Identify faults or partial shading in PV arrays, and reduced efficiency in PV systems	✓			[53,57, 69,71,7 3,74]
Infrared thermography	PV array	Thermal image (°C)	✓				✓	Image processing; expert system; discriminant analysis; machine learning (SVM); deep learning (CNN, LSTM)	Detect hot spots and failures in PV panels	✓		✓	[56,62, 64,69]
Flash	PV array	Power (W)	✓		✓			Voltage, current, and power output; machine learning (SVM, GAN, transfer learning); deep learning (CNN)	Detect defects and reduced efficiency in PV panels	✓		✓	[59,60, 72]
Electroluminescence	PV array	Various	✓	✓			✓						
X-ray	PV array	Various	✓	✓			✓						
Model parameter identification	PV system	Various		✓	✓		✓	Analytical approach; machine learning (ANN); metaheuristics (particle swarm optimization, genetic algorithms, differential evolution)	Validate parameters for simulation; optimization of PV system performance; energy production forecasting		✓	✓	[55,65]
Physical fault (panel damage, defects)	PV array	Various	✓				✓	Signal processing (FFT, WT); statistics; model-based and real-time difference; machine learning (ANN, SVM, decision trees); deep learning (LSTM, CNN, DBN); fuzzy logic; handcrafted diagnosis; failures modes and effects; PV and solar emulator	Perform predictive maintenance of PV systems; early detection of faults; real-time detection and classification; monitoring; fault forecasting	✓	✓	✓	[52–54,56,5 8,61,63, 66,73,7 4]
Environmental fault (partial shading)	PV array	Irradiance (W/m^2)	✓				✓						
Electrical fault (line-line, line-ground, open circuit, blocking diode, arc)	PV array, inverters	Current (A)	✓		✓								

shows the corresponding diagnostic techniques, and indicates the PV subsystem involved in each test and the measured parameters.

Testing procedures described in standards include continuity verification of conductors and grounding, polarity check in combiner boxes and conductors, open circuit voltage, and short circuit and operational currents in the PV array, which are relevant parameters to monitor the correct operation of PV systems [69–71]. Tests are also conducted on the energy storage system such as usable battery capacity, recovery, and battery charge tests [71].

Additional testing procedures with demanding requirements for result analysis are described in the literature such as I-V curve-based approaches to study PV system performance from its I-V characteristic [53,57,73,74]. Other investigations use infrared thermography to detect hot spots in PV systems [56,62,64,69]. Moreover, flash, electroluminescence, and X-ray tests are employed to find defects in PV arrays [59,60,72]. It is worth noting that the I-V characteristic can be analyzed indirectly using thermography, flash, electroluminescence, and X-ray because there is a relation between hot spots and defects found with those tests and the voltage and current performance of the PV system. Tests and simulations are also performed for model parameter identification which is mainly intended for simulation, optimization of PV system performance, and energy production forecasting [55,65]. Moreover, modeling, simulation, and testing of faults have been widely studied. For instance, [52] presents an extensive overview of models to analyze DC faults in PV systems. The work in [53,63] extend the overview of electrical faults on the PV array, inverters, and the AC side of PV systems. In addition, [54,66] analyze not only electrical faults, but also physical, e.g., damage, cracks, and deterioration in PV panels, and environmental faults, e.g., temporary, and permanent partial shading. As an alternative to detailed modeling and simulation, PV and solar emulators are also available to perform a variety of tests [73,74]. This system type enables the emulation of faults within PV systems, as well as the evaluation of performance by means of analyzing the I-V curve.

After PV systems are tested, different diagnostic techniques are applied to analyze results. These techniques can be classified

according to the categories identified in Section 1.2 (see Figure 3), i.e., expert knowledge, analytical models, and based on AI. Table 5 also includes the classification of tests and diagnostic techniques according to the taxonomy defined in Figure 3.

A first set of approaches based on expert knowledge are standards, and manufacturer recommendations to assess electrical, mechanical, and electromechanical performance of PV systems [69–71]. In these methods, expert knowledge is required to apply tests using measurement instruments and provide a performance assessment based on the parameters established by standards. Furthermore, I-V curve tests are also analyzed through expert knowledge using handcrafted diagnosis [57], monitoring real-time difference with respect to a reference I-V curve [54,57], and using PV and solar emulators [73,74]. Also, expert systems used for interpretation of infrared thermography [62], and fuzzy logic for fault diagnosis in PV systems require expert knowledge [65].

Analytical models proposed for PV system analysis include, for instance, equations formulated to compute usable battery capacity, recovery, and battery charge in stand-alone applications [71]. Other examples include the analytical approach for model parameter identification [55,65], and the model-based difference for fault diagnosis [54], where monitored parameters of PV systems are compared with a modeled and theoretical reference to identify deviations from the correct operation.

Approaches based on AI techniques have been extensively used in the last few years. For instance, machine learning tools such as SVM, ANN, Generative Adversarial Networks (GAN), transfer learning, and decision trees are used for automatic analysis of images in infrared thermography [56,62,64,69], flash, electroluminescence, and X-ray [59,60,72], and in fault diagnosis applications [52–54,56,58,61,63,66,73,74]. Similarly, deep learning tools such as Deep Belief Networks (DBF), Long Short-Term Memory (LSTM), and Convolutional Neural Networks (CNN), have been also applied recently because of their extended capabilities for image analysis. All these AI techniques mentioned above usually require a previous step of feature extraction where statistics and signal processing techniques, e.g., Wavelet Transform (WT), and Fast Fourier Transform (FFT), are applied.

3.2. Indices for condition assessment of PV systems

In the literature, performance indices are offered for health and condition monitoring of PV systems. For instance, the world-wide association SolarPower Europe made up of hundreds of companies and policymakers from the solar sector, proposed best practice guidelines for asset management of PV projects [77]. Within the scope of these guidelines, Key Performance Indicators (KPI) are suggested to achieve technical asset management of PV power plants, e.g., specific and reference yield, performance ratio, energy performance index, among others. Moreover, equipment KPI are also indicated in the guidelines such as inverter specific energy losses, Mean Time Between Failures (MTBF), etc. Likewise, energy metrics are proposed in [76] to assess the behavior of PV systems. These metrics comprise daily and monthly energy output, array energy yields, PV array and system efficiency, capacity utilization factor, and life cycle conversion efficiency.

Furthermore, indices to analyze condition from the degradation perspective are also found in the literature. In this context, [67,68] offer a comprehensive review of degradation rates of PV systems around the world, and factors and parameters involved in the degradation process such as corrosion, encapsulant discoloration, and light and elevated temperature induced degradation. Similarly, [75] studies a health monitoring method based on the study of I-V curves and degradation modes of the PV array.

4. Distribution Transformers (DT)

DT are static, efficient, and reliable electric machines that, based on electromagnetic induction principle, supply electrical energy to end LV costumers. Distribution networks comprise thousands of DT, which are lower cost assets and more available into stock for replacement compared with PT. DT are also characterized by their rated powers that range between a few kVA to 2,5 MVA for three-phase transformers, or up to 833 kVA for single-phase transformers [78,79]. Moreover, according to the type of assembly, DT can be pad-mounted, i.e., on the ground, or pole-mounted, i.e., at the level of overhead lines.

Despite the differences, both PT and DT have many physical and constructive similarities, thus, DT consist of the following subsystems: (i) the AP including core, winding, tank, oil, and insulation paper (also referred to as insulation system), (ii) the BH, (iii) the CS, which can be dry or mineral oil immersed, (iv) the CTk, and (vi) OLTC. Moreover, the life span of their insulation systems is affected by the same factors, where DT useful life is usually around 25 years [80,81].

According to literature, over 60 % of in-service DT are aged [80]. Aged DT is a problem for electric utilities because the assets are more prone to failures and large-scale replacement of aged units is prohibitive due to high costs. Furthermore, the widespread incorporation of modern electrical loads and the unforeseen growth of energy demand result in unplanned overloading from the usual operation conditions [82], increasing probability of failures and operational costs [83].

4.1. Fault diagnosis and condition assessment of DT

According to the context described above, asset management and condition assessment methodologies are required for in-service DT to minimize failure risks, increase reliability, and coordinate maintenance actions. In this regard, Table 6 summarizes research related to maintenance and condition assessment approaches on DT that has been found in the literature.

In the current practice, condition assessment of DT is mainly based on preventive and corrective maintenance [84]. However, new DT maintenance trends include condition monitoring using sensors and measuring devices to detect incipient faults and continuously assess its operational performance. Currently, condition monitoring is not used on a large-scale by utilities on their in-service DT due to high implementation costs. However, technical benefits and improvement of DT efficiency using condition monitoring are demonstrated in the literature by several case studies on specific in-service units.

To achieve proper condition monitoring of DT, several parameters must be estimated with tests or inspections. For instance, chemical tests are used to measure moisture and oxygen content in the insulation system, dissolved gases, dielectric strength, power factor, and acidity, to estimate the remnant life of

Table 6: Classification of tests on DT and diagnostic techniques.

Data acquisition								Data analysis					Refs.
Test or inspection	Subsystem	Parameter	Type of test					Diagnostic technique	Purpose	Type of analysis			
			In-service	Off-service	Electrical	Chemical	Physical			Exp. know.	Analytical	AI	
Load History	AP	Power (kVA)	✓		✓			Analytical models from standards for hot spot temperature estimation; direct temperature measurements	Indirect estimation of DT loss of life; load balancing		✓	✓	[78–80,83,88–97]
Ambient temperature	AP, CS	Temperature (ºC)	✓				✓						
Winding/top oil temp.	AP	Temperature (ºC)	✓				✓						
Terminal voltages	AP, BH	Voltage (V)	✓		✓			Direct measurement and interpretation of DT parameters; standards and guidelines	Detect deviations from normal, recommended, or standardized values	✓		✓	[81,84,94,98,99]
Dissipation, power factor, capacitance	AP, BH, OLTC	Tan δ, power factor, capacitance (F)	✓	✓	✓								
Bushing inspection	BH	Voltage (V)	✓				✓						
Recovery voltage	AP	HV impulse (kV)		✓	✓								
DGA	AP, BH, OLTC	Gases in mineral oil (ppm), Gases in ester oil (ppm)	✓			✓		Dornenburg ratio; Key gas; Rogers ratio; data mining; Duval triangle; fuzzy logic; direct measurements	Early detection of electrical faults; insulation status assessment	✓		✓	[85,88,98–101]
FRA	AP	Impedance (Ω), transfer function	✓	✓	✓			Direct interpretation of the measured curve; comparison measurements in the same DT; expert system; ANN	Detection of winding displacement and core movement	✓	✓	✓	[85,87,102,103]
Oil characteristics	AP, OLTC	Moisture, acidity, color, breakdown voltage, dissipation factor (%), oil level	✓			✓	✓	kNN; direct measurement; direct interpretation and comparison of results from acidity, interfacial tension, and dielectric strength tests with standardized values	Assessment of insulating oil status	✓		✓	[92,96,98,99,104–106]
Acoustic signals	AP, OLTC	Acoustic emission (dB), freq. (kHz)	✓				✓	Spectral analysis; time waveform analysis; pattern recognition; acoustic imaging	Health assessment of core and winding structures		✓	✓	[85,94,96,107]
PD	AP, BH	Electric charge (pC)		✓	✓			Analysis of direct PD measurements; acoustic PD measurement to identify the source and location of PD activity; optical detection of the light emitted by PD	Determine the location, magnitude, and type of PD in insulation materials	✓	✓	✓	[85,92,93,98,99,101]
Infrared thermography	AP, BH, OLTC, CTk, CS	Thermal image (ºC)	✓				✓	The temperature variations observed in the infrared image are interpreted in the context of the DT operation	Identify temperature rise, loose connections, insulation failure	✓			[97]
Visual inspection	AP, OLTC, BH, CTk, CS	Oil level, gas pressure, fan and pump condition, corrosion	✓				✓	The inspector records and classify component observations, including any signs of damage, wear, corrosion, etc. The severity of the damage is compared to standards	Identify observable anomalies	✓			[84,94]

DT [85]. The most widely performed test for the analysis of dissolved gases is DGA, which is used for detection of incipient faults due to oil and solid insulating material degradation. DGA is conducted through in-service gas monitoring and considers the discharging rate and ratios of specific dissolved gas concentrations to determine fault types and severity in DT. To analyze results of DGA in DT, several diagnostic techniques are applied including Rogers ratio, key gas, Total Dissolved Combustible Gas (TDCG), Dornenburg ratio, and Duval triangle. In addition, AI-based tools are applied to DGA when more accurate fault diagnosis is required. Finally, DGA can be achieved using online gas monitoring, but it is expensive especially for fleets of extensive DT.

Further inspections in DT include physical and chemical tests of oil quality. These tests involve lower costs than online monitoring of gases and can be used to estimate health status of DT oil by moisture, oxygen, interfacial tension, acidity, dielectric strength, and power factor tests [86]. Results of the tests can be analyzed with expert knowledge or AI-based techniques.

Likewise, electrical tests are used to estimate deterioration of the AP. For instance, FRA is a non-invasive method in the frequency domain that detects mechanical and electrical damage such as winding deformations and displacements inside the DT. FRA is an effective test for DT monitoring, however, it can be affected by network conditions during online application [83]. Therefore, interpretation of FRA results still requires further research because of its complexity, including new computational methods such as AI-based diagnostic techniques, e.g., [87].

PD and short circuit impedance measurement are two additional electrical tests. PD measurement is a non-destructive testing procedure used to assess and monitor the condition of BH insulation and HV, LV and inter-turn winding insulation. PD can be measured off- or in-service [108]. Moreover, the short circuit impedance test is used to detect imperfect magnetic coupling, estimate short circuit parameters of DT, and comparing these values to factory test results.

The review in [85] divided miscellaneous diagnostic techniques applied to DT in Signal-Based (SB) and Data-Based (DB). SB approaches rely on processing of signals from mechanical vibration, sound, thermal or optical parameters, which is suitable for condition monitoring of in-service DT. Thus, SB methods can be classified into vibration analysis, optical fiber, acoustic emission, and thermography interpretation. For DB methods, data of DT is essential to develop statistics, standards, numerical and mathematical modeling. In general, DB approaches offer very reliable results with several sophisticated and validated models to infer DT aging.

AI methodologies integrate computational intelligence with sophisticated sensing, experience, learning, and training to enhance the precision of fault diagnosis and condition evaluation in DT. AI-based approaches applied to DT have been found in the

literature such as Bayesian networks, genetic algorithms, ANN, Adaptative Neuro Fuzzy Inference System (ANFIS), among others. The AI-based techniques mentioned above are used for HI calculation, FRA interpretation, hot spot detection, DGA and oil analysis, incipient fault detection, health condition estimation, insulation diagnosis, among other analyses applied to DT.

4.2. Indices for condition assessment of DT

The HI is an essential indicator to determine health status and a useful tool to perform proper management of DT in their life span. In this regard, different approaches to compute the DT-HI are found in literature, especially DB and AI-based techniques. DB approaches are developed using analytic methods based on specific measured data of DT parameters, usually recommended by IEEE and IEC standards, the Common Network Asset Indices Methodology (CNAIM) [109], or specific procedures proposed by researchers [94]. In such cases, the HI is calculated as the weighted sum of DT condition influencing factors, which are scored with weights according to importance [101]. Usually, estimation of weights is based on expert knowledge.

Because of high implementation costs of conventional DT condition assessment methods such as those based on DGA, FRA, and PD, novel advanced AI-based techniques have been developed. For instance, the DT-HI can be obtained using data mining [100], genetic algorithms [81], k-Nearest Neighbour (kNN) [104], Internet of Things (IoT) [105], and fuzzy logic [110].

5. Switchgear

MV switchgear are essential assets for power distribution utilities because they include protection, regulation, control, and measurement devices that allow safe operation of electrical networks in accordance with regulations. MV switchgear are also considered critical assets because their malfunction can put other assets at risk and affect the availability and quality of the electric power supply, causing financial losses to the utility. In extreme cases, this malfunction can affect safety of people, resulting in legal and financial sanctions and the deterioration of the utility's image. Thus, proper status assessment of MV switchgear is of utmost importance for early identification of possible failures and, consequently, for mitigating the negative impacts of failures on the utility. However, advances in condition monitoring of MV switchgear are incipient because of the relatively low cost of each of these assets compared with other critical assets [111], e.g., PT.

MV switchgear are available in various topologies primarily determined by their rated current (e.g., 630 A – 3150 A), and voltage (e.g., 7.2 kV – 36 kV). Moreover, MV switchgear can be classified based on their insulating medium in Air-Insulated Switchgear (AIS), Gas-Insulated Switchgear (GIS), or Oil-Insulated Switchgear (OIS). MV switchgear comprises subsystems such as dielectric, electronic, mechanical, and fluid components. Some of the sub-assets that make up an MV switchgear include (i) the primary circuit (e.g., bushings, busbars, cables), (ii) the interrupting chamber (e.g., breaker, isolating medium), (iii) mechanical linkages, (iv) the mechanism (e.g., spring, hydraulic, motor), and (v) control and auxiliary circuits (e.g., IT, protective relays, control, and monitoring systems) [112].

The breaker is widely regarded as the most critical sub-asset from both technical and economic perspectives because the protection of the entire system depends on its proper operation. Some studies into failure statistics of electrical components in MV networks have revealed that breakers represent the predominant components susceptible to failure in MV switchgear. Around 90% of these breaker issues are attributed to mechanical factors, impacting both the breaker motor and its operational mechanisms [113,114]. Consequently, this section mainly focuses on the diagnosis of the breaker, while also addressing some relevant aspects of the remainder components of the switchgear.

5.1. Fault diagnosis and condition assessment of switchgear

Under normal conditions and during the interruption of short-circuit currents, a breaker is subjected to mechanical, thermal, and dielectric stresses [115]. Over time, these efforts deteriorate the breaker's components, such as contacts, means of interruption, and operating mechanism. Deterioration can be accelerated by factors including manufacturing defects, frequency of operation, network development and evolution of the network failure rate [116].

According to [117], the main causes of breaker failures in MV networks in the United Kingdom are related to mechanical problems (30 %), partial discharges (26 %), and problems in the insulation (11 %). These statistics highlight the significance of diagnostic tests in assessing the breaker's health status.

In the report [118], the International Council on Large Electric Systems (CIGRE) provided a comprehensive review of the most widely accepted diagnostic methods for assessing the condition of MV switchgear. Some of these methods have reached a certain level of technological maturity and are therefore applicable for evaluating the health status of MV breakers. Table 7 presents a classification of some of these diagnostic tests.

Once diagnostic tests are applied, several diagnostic techniques can be applied for applications such as maintenance planning, and condition monitoring. Applying predictive maintenance in MV switchgear presents two main challenges [119]: determining the suitable sensors capable of consistently and resiliently measuring the critical physical parameters for MV switchgear management throughout its useful life; and the lack or small amount of data. For instance, continuous monitoring of temperature of MV switchgear is uncommon or nonexistent during their useful life and monitoring of the number of operations is carried out only few times per year, generally coinciding with periodic and scheduled maintenance.

This limited availability of data makes it difficult to implement AI algorithms. However, some studies have made certain advances in applying methods such as ANN for the diagnosis of MV switchgear, as reported in Table 7.

5.2. Indices for condition assessment of switchgear

a) Health Index (HI)

The HI of MV switchgear (HI_{SG}) is obtained by calculating the weighted sum of HI associated with each diagnostic test included in the model ($HI_{SG,n}$). A weight ($W_{SG,n}$) is assigned to each procedure based on its influence on the overall health status of the asset. The formula used for this weighted sum is expressed in (1).

$$HI_{SG} = \frac{\sum_{n=1}^{N} W_{SG,n} \cdot HI_{SG,n}}{\sum_{n=1}^{N} W_{SG,n}} \quad (1)$$

Table 7: Classification of tests on switchgear and diagnostic techniques.

Data acquisition								Data analysis					Refs.
Test or inspection	Subsystem	Parameter	Type of test					Diagnostic technique	Purpose	Type of analysis			
			In-service	Off-service	Electrical	Chemical	Physical			Exp. know.	Analytical	AI	
Dynamic and Static Resistance of Contacts	Breaker	Resistance (μΩ)		✓	✓			Gaussian process regression; ANN	Automatically discern pertinent predictor variables for a remaining useful life model of unmonitored, low-value power network assets, and contact temperature calculation	✓		✓	[121–124]
Timing of Operation	Breaker	ms / cycles	✓	✓			✓	Gaussian process regression	Automatically discern pertinent predictor variables for a remaining useful life model of unmonitored, low-value assets within power networks.	✓			[123,125,126]
PD (UHF, TEV, and ultrasound)	Primary circuit, breaker, IT	pC (directly / correlated)	✓		✓			Adaptive particle swarm optimization and ANN; extreme learning machine; automatic clustering and linguistic rules; CNN; scalogram representations and convolutional autoencoder	Partial discharge fault diagnosis and arcing fault detection	✓		✓	[111,124,127–134]
Insulation, dielectric integrity	Primary circuit, breaker, IT	Various		✓	✓			Optimal grade clustering method	Switchgear Insulation Condition Assessment			✓	[126,135]
Acoustic and vibrations (Accelerometers)	Mech. linkages, mechanism	Various	✓	✓			✓	Multilayer perceptron (ANN)	Fault diagnosis based on vibration signal			✓	[136–139]
Visual (as defined in IEC 62271-1)	Primary circuit, int. chamber; mech. linkages, mechanism, control and aux. circuits	Various	✓				✓	The inspector records and classify visual observations of the switchgear components, including any signs of damage, wear, corrosion, or other issues. The severity of the damage can be compared to established standards or guidelines for switchgear	Identify observable anomalies	✓			[122]
Temperature (IR cameras, Fiber guided sensors)	Primary circuit, int. chamber, mechanism	Temperature (°C, °F)	✓				✓	The temperature variations observed in the infrared image are interpreted in the context of the PT operation and maintenance history	Identify areas of abnormal temperature rise, overheating, loose connections, or insulation failure	✓			[140–142]
Monitoring number of switch operations	Breaker; Primary circuit	Various	✓				✓	Standards, guidelines, and manufacturer recommendations to assess the switchgear electrical performance	Detect deviations regarding normal, recommended, or standardized values	✓			[118,125,126]
Gas/Oil/Vacuum Quality (Mobile magnetron, Dielectric Vacuum Test)	Interrupting chamber	Various		✓		✓	✓						
Analysis of recordings from digital relays	Control and aux. circuits, int. chamber	Various	✓			✓							
Power factor	Prim. circuit, IT	Various	✓		✓								

Definition of weights $W_{SG,n}$ should consider limits and recommendations provided by manufacturers and knowledge of experts in the operation and maintenance of MV switchgear. For instance, indicators such as the number of mechanism operations and contact resistance are of great importance for estimating the health status of switchgear [116–118], so they should pose significant weight for computing HI_{SG}. However, the health status obtained through those tests ($HI_{SG,n}$) also depends on the intrinsic operating characteristics of the asset, defined by the manufacturer.

b) Remaining life

A methodology for computing the remaining life of MV switchgear based on failure mode analysis and performance analysis of the switchgear's components using existing data sources is presented in [120]. The approach suggests considering the components and subsystems of MV switchgear and estimating the overall HI as a weighted sum. The specific components and subsystems considered will depend on the type of switchgear, its technology, and its nominal values.

6. Overhead lines and underground cables

Lines and cables are highly susceptible to failures due to their long length. Given the great importance of lines and cables in the distribution network, extending their useful life, and making timely decisions to prevent failures and reduce maintenance and operational costs are crucial objectives. Lines and cables maintenance requires considering parameters such as the structural and environmental conditions, useful life, and operating environment. Despite overhead lines and underground cables have the same purpose, i.e., to transport electrical power, in practice, they are very different assets because of constructive aspects. Therefore, they are analyzed separately in the following.

6.1. Overhead lines

A MV overhead line is usually composed by (i) conductors, where the main types are Aluminum Conductor Steel Reinforced (ACSR), All Aluminum Alloy Conductor (AAAC), All Aluminum Conductor (AAC), Aluminum Conductor Composite

Core (ACCC) and Aluminum Conductor Steel Supported (ACSS), (ii) porcelain, composite, or glass insulators, (iii) fittings, mainly made of galvanized steel, and including suspension, tension, dead end, splice, and cross-arm fittings, (iv) poles made of wood, concrete, steel or, most recently, composite, i.e., a combination of fiberglass or carbon fiber reinforced polymer materials (PFRV), (v) cross-arms made of immunized wood, galvanized steel or PFRV, with rod or angle iron diagonals, (vi) sectioning and protection equipment including disconnectors, circuit breakers, and fuses, and (vii) guard lightning protection cables, communication cables and grounding systems.

To assess the MV overhead line condition, it is recommended to evaluate their components separately because the variety of materials mentioned above requires different assessment criteria. For instance, failure and aging processes in each material need to be known for each type of component.

6.2. Underground cables

Underground cables are in some cases preferred over overhead lines because of their lower visual impact and higher reliability. However, constructive requirements of underground cables make them more expensive than overhead lines. The main challenges that underground lines must overcome are to provide enough insulation so that the cables can be installed underground, and to dissipate the heat produced during operation.

The commonly used underground cable types are Cross-linked Polyethylene (XLPE), Self-Contained liquid Filled (SCFF), High Pressure, liquid Filled Pipe (HPFF), High pressure, tube Filled with Gas (HPGF), and Solid Cable. Solid cables demand lower maintenance, although monitoring and detecting insulation failures can be challenging. The size of XLPE cables increases with voltage. Typically, each circuit employs three distinct cables, which are enclosed within a buried or side-by-side conduit.

6.3. Methods for diagnosis and condition assessment of lines and cables

Tests and diagnoses are carried out to determine the MV lines and cables health with the objective of prioritizing decision-making in maintenance and to prevent service interruption due to damages. Among the main tests carried out on lines and cables are visual inspection, thermography, ultrasound, conductor analysis, grounding checking, structure verification, etc. [143–146]. Table 8 summarizes and classifies these main tests according to the taxonomy in Figure 3. Furthermore, Table 8 provides information on the corresponding diagnostic techniques, which range from traditional standard guidelines based on expert knowledge to advanced AI-based techniques such as decision trees.

6.4. Indices for condition assessment of lines and cables

Planning of maintenance, renewal or change of MV distribution lines and cables can be determined from asset management, using tools such as the HI calculation. There are some proposals that deal with this issue, for instance, [146] presents a method to compute an HI for overhead lines using visual inspection findings. An example is shown for foundations and poles, knowing that the aging of reinforced concrete is assessed by monitoring the gradual reduction in structural strength over time. Poles are evaluated within a condition range on a scale of 0 to 5, where a rating of 5 indicates a poor technical condition with an operational lifetime in its final stage, and a rating of 0 represents a state of good condition. Within this scale, the presence of longitudinal cracks with widths ranging from 0.3 to 0.6 mm corresponds to a score of 4. A table with the proposed score for different state conditions is detailed in [146].

Finally, the general HI, considering all the components of the overhead line, is computed using the weighted summation in (2).

$$HI_L = \sum_{i=1}^{M} HI_{L,i} \cdot W_{L,i} \tag{2}$$

where $HI_{L,i}$ is the index for each line component i, and $W_{L,i}$ its weighting factor, which measures the component importance and is usually defined by utility experts. It is advised to use an exponential scale to convert $HI_{L,i}$ values in (2) to achieve a suitable comparison for extreme values [146].

7. Instrument Transformers (IT)

The signals acquired by IT and subsequently processed and stored by measurement devices, are decisive state variables to control and protect power systems. Two main subsets of IT are Voltage Transformers (VT) and Current Transformers (CT). A VT produces a secondary voltage that is directly proportional to the primary voltage, while a CT generates a secondary current that is proportionate to the primary current. Thereby, VT and CT offer a high-fidelity measurement of voltage and current, respectively, while providing isolation from high voltage and current levels.

On the one hand, there is a variety of VT such as capacitive, resistive-capacitive, and inductive [147]. Inductive VT, for instance, are widely used for measurement and protection in MV networks. On the other hand, there are three types of CT including electromagnetic, capacitive, and optical. Optical CT based on the magneto-optic Faraday effect, are used for measurement and smart grid protection. They have high immunity to electromagnetic disturbances and are suitable for control and protection systems featuring an extensive measurement range [148].

For any type of IT, reliable measurements are very important. Input signals for protection relays are channeled via IT to execute the protective functions. It is important to highlight that network reliability may be compromised in the event of IT failure, particularly if the input signals for protection and control prove to be imprecise. In this context, it is relevant to identify tests and assessment approaches for IT management.

7.1. Fault diagnosis and condition assessment of IT

The tests to diagnose IT can be carried out with the asset in-service or off-service, as specified by the corresponding standards. These tests allow identifying anomalies or failures and provide some indication of the expected reliability and loss of life of the asset. It should be noted that a single electrical test may not be enough to diagnose the asset, but several tests may be necessary to identify a problem. In addition, the acceptance criteria are specific for each manufacturer.

There are three types of tests including routine tests, carried out on a sample of the equipment, e.g., short circuit withstand tests in CT and VT, and overvoltage test in CT; type tests, applied to one

Table 8: Classification of tests on underground cables and overhead lines and diagnostic techniques.

Data acquisition								Data analysis					Refs.
Test or inspection	Subsystem	Parameter	Type of test					Diagnostic technique	Purpose	Type of analysis			
			In-service	Off-service	Electrical	Chemical	Physical			Exp. know.	Analytical	AI	
Visual inspection	Conductors, insulators, fittings, cross-arms, poles	Corrosion, physical damage, loose strands, vegetation growth, etc.	✓	✓			✓	Check list; HI calculation based on predefined assessment criteria for each type of MV line component	Detect and document anomalies for further treatment; estimate the overall HI to identify critical lines; detect defects for MV line technical condition assessment; supplement onsite measurements, laboratory measurements or full-scale tests	✓		✓	[146]
Dielectric spectroscopy (over a low frequency range, i.e., few hertz)	Underground cables	Relative permittivity, dissipation factor, power factor, capacitance		✓	✓			Comparisons of measured relative permittivity with values of reference	Estimation of the breakdown voltage reduction; estimation of stage of water tree ageing (i.e., formation of microscopic tree-like structures within the insulation due to the presence of water or moisture over an extended period)	✓			[149]
PD	Underground cables	Electric charge (pC)		✓	✓			Decision tree theory	Time-domain and frequency-domain measurement and analysis of PD data			✓	[150,151]
Corona effect	Underground cables	Various	✓				✓	Method of noise reduction in the UV channel	Estimation of fault severity	✓	✓		[145]
Tests on structures	Structure	Various	✓		✓		✓	Post breaking strength	Determination of weight tolerance to support	✓			[146]
Thermography	Und. cables, poles, insulators, structure	Temperature (°C)	✓				✓	Standards, guidelines, and manufacturer recommendations to assess the lines and cables electrical performance	Detect deviations regarding normal, recommended, or standardized values	✓			[143]
Cable mechanical	Underground cables	Various		✓			✓						[152–154]
Cantilever, load (poles)	Underground cables	Various	✓	✓			✓						[152,155, 156]
Grounding	Structure	Resistance (Ω)	✓		✓								[144]
Electromechanical and mechanical failing load	Underground cables, insulator, structure	Various		✓			✓						[157,158]
Puncture voltage	Insulator	Voltage (kV)		✓	✓		✓						[157,158]
Ultrasound	Und. cables, poles, insulators, structure	Various	✓				✓						[143]
Tracking effect	Underground cables	Various	✓										[143]
Dry lightning impulse withstand voltage	Underground cables, insulator, structure	Various		✓	✓								[157,158]
Flashover	Underground cables, insulator, structure	Various		✓	✓								[157,158]
Wet power-frequency withstand voltage	Underground cables, insulator, structure	Various		✓	✓								[157,158]

or two identical or similar types of equipment to verify a specific characteristic, e.g., heating test, nominal and safety factor; and maintenance tests, carried out on equipment in-service, to verify its health status after a certain period of operation. A summary of diagnostic tests for IT based on the review is reported in Table 9.

Table 9 also reports diagnostic techniques applied on the results of tests for condition assessment of IT. It is observed that research on condition assessment of IT is still incipient (see also in Table 3 that only 26 records were retrieved with the formulated search rule). Diagnostic techniques found in the literature use, for instance, statistical models such as Weibull, or AI-based techniques including ANN and SVM, as shown in Table 9.

7.2. Indices for condition assessment of IT

Inadequate maintenance practices, environmental factors, and the transformer's current condition can lead to failures that have an adverse impact on the transformer's expected lifespan. In this context, the adoption of a predictive approach for asset management and future condition forecasting would allow avoiding failures through early detection of technical problems that may affect the transformer [159]. Research on indices for condition assessment of IT is also very incipient, however, some indicators have been formulated.

a) Loss of life

Presently, endeavors are in progress to establish guidelines for estimating the reduction in lifespan of a CT due to short-term overloading. In a related study [160], a thermal accelerated aging method is introduced and implemented on a 69 kV oil-immersed CT subjected to overloading. The laboratory-acquired life test outcomes are validated against analytical assessments involving DP and the analysis of furanic compounds in the transformer oil. The standard IEEE C57.91 is used as a reference to formulate the aging acceleration factor (F_{AA}) as expressed in (3).

$$F_{AA} = \exp\left(\frac{15000}{368} - \frac{15000}{\theta_H + 273}\right) \quad (3)$$

Likewise, the percentage of loss of life (*LOL*) is calculated resulting from the equivalent acceleration factor (F_{EQA}), the normal life of the insulation (DP = 200) and the duration of the stress (t), according to (4)-(5).

Table 9: Classification of tests on IT and diagnostic techniques.

Data acquisition								Data analysis					Refs.
Test or inspection	Subsystem	Parameter	Type of test					Diagnostic technique	Purpose	Type of analysis			
			In-service	Off-service	Electrical	Chemical	Physical			Exp. know.	Analytical	AI	
Power-frequency withstand tests on windings and between sections	CT, VT	Voltage (kV)		✓	✓		✓	Standards, guidelines, and manufacturer recommendations to assess the IT electrical performance	Detect deviations regarding normal, recommended, or standardized values	✓			[161–163]
Measurement of the radio interference voltage (RIV)	CT, VT	Voltage (mV)		✓	✓								
Mechanical test (Use loads due to wind and ice)	CT, VT	Incline degrees vs. gravity units		✓			✓						
Determination of errors (Accuracy class, current error, phase displacement)	CT, VT	Phasor degrees, error (%)	✓		✓								
Verification of terminal marking and polarity	CT, VT	Various		✓			✓						
Measurement of capacitance and dielectric dissipation factor	CT, VT	Tan δ, power factor, capacitance (F)		✓	✓								
Resistance measurements	CT, VT	Resistance (mΩ)		✓	✓								
Thermography (Infrared inspection)	CT, VT	Thermal Image (ºC)	✓				✓	ANN from the Least-Mean-Square (LMS)	Predicting the max. temperature in the low and high voltage winding			✓	[164]
Visual inspection	CT, VT	Various	✓	✓			✓	The inspector records and classify visual observations of the IT components, including any signs of damage, wear, corrosion, or other issues.	Identify observable anomalies	✓			[165], [164]
DGA and liquid content	CT, VT	Gases dissolved in oil (ppm)		✓		✓		Multi-stage classifier based on SVM	Direct measurement of DGA value		✓		[166–168]
Partial discharge measurement	CT, VT	Electric charge (pC)		✓	✓		✓	Severity Index (SI)	Indirect estimation of DP value		✓		[161–163,169]
Frequency Domain Dielectric Spectroscopy (FDS)	CT	Complex capacitances and Tan δ		✓	✓			Harverlik-Negami dielectric response function	The model was used to fit FDS	✓	✓		[170]
Flashover and inclined plane tracking test	CT, VT	Voltage (kV)						The Weibull distribution plot	The predicting the flashover and degradation initiation		✓		[171]
FRA	CT, VT	Admittance (S)/ impedance (Ω)	✓	✓	✓			Total Vector Error (TVE)	Evaluated with different input voltages to identify failures		✓		[172]
Synchrophasor measurement	CT, VT	Phasor measurement unit (PMU)	✓				✓	3Φ-SLSE (Three-Phase Current and voltage State Estimator + The Largest Normalized Residual (LNR))	Model based real-time application to identify the presence of incorrect measurements	✓		✓	[173,174]

$$F_{EQA} = \frac{\sum_{n=1}^{N} F_{AA,n} \cdot \Delta t_n}{\sum_{n=1}^{N} \Delta t_n} \quad (4)$$

$$LOL = \frac{F_{EQA} \cdot t}{normal\ insulation\ life} \times 100\% \quad (5)$$

b) Lifetime

Significant thermal aging of paper insulation stands out as a pivotal constraint on the operational lifespan of IT. In [175], research encompasses simulation studies and a comprehensive review of literature pertaining to insulation systems, moisture dynamics, and statistical modeling of IT lifetime data. The primary objective of this research was to identify potential triggers and factors that impact IT failures. Results showed the relationship between failures and age, environmental factors, and operating stress level.

c) Health Index (HI)

The HI serves as a metric for quantifying the enduring deterioration of a transformer. It combines intricate condition-related information into a singular value, offering a relative assessment of the overall state of a transformer. Most HI methodologies rely on either laboratory or on-site test data, with limited incorporation of actual operational time [159,176].

For instance, the calculation of a CT HI is formulated in [159], where oil tests (furans, acetylene, etc.) are considered, as well as thermal aging factors and other parameters including maintenance, environment, location, failure history, and observed external and internal condition. A life expectancy model of the upgraded CT is developed based on the calculated HI and the HI aging factor. The calculation is given by (6).

$$HI_{IT} = HI_{IT-0} \cdot e^{B \cdot (T_2 - T_1)} \quad (6)$$

where HI_{IT-0} is the initial HI considered as 0.05, B is the aging factor, T_1 is the year corresponding to HI_0 (T_{1-0} = expected life) and T_2 is the year under study ($T_{2-0} = 0$).

8. Discussion

A first insight offered by the quantitative review developed in this paper is that only research on condition assessment of PT shows high level of maturity in all the analyzed aspects including tests, diagnostic techniques, and indices. Moreover, condition assessment on PV systems is a topic of increasing interest because of the current relevance of renewable energy resources. Therefore, significant efforts have been put in the topic during the last five years, however, further work should be carried out on indices for condition assessment of PV systems. Regarding condition assessment of DT, switchgear, lines and cables, and IT, research is mature in terms of tests, but still incipient on diagnostic techniques and indices.

Different types of tests, mostly based on widely accepted practices and standards, are applied on the selected MV assets, and a taxonomy was formulated in this paper accordingly. In this sense, according to the operating status of the asset, tests are classified as in-service and off-service. Results of the review summarized in Tables 4-9 demonstrate that a similar amount of in-service, and off-service tests are available for the MV assets dealt within this review. Moreover, according to nature, most of those tests are electrical (as expected because of the electrical nature of the assets), e.g., short-circuit withstand tests, voltage and current measurements, FRA (in PT and DT), followed by physical tests, e.g., visual inspection (in all the analyzed assets), mechanical tests (of particular importance for switchgear), thermography (in all the analyzed assets), and chemical tests (mostly applied to PT and DT), e.g., DP, furan analysis, DGA. No chemical tests were identified in the literature for PV systems, lines, and cables.

Likewise, categories were defined in this review for diagnostic techniques, namely, expert knowledge, analytical models, and AI-based approaches. These diagnostic techniques aim at improving interpretability of test results. From Tables 4-9 and the extensive review, it is seen that most diagnostic techniques are based on expert knowledge (established in standards and widely accepted procedures and industry practice), whereas very few approaches use analytical models probably because of the complex derivation and implementation of physical models that hinders the applicability. However, analytical models are still valuable diagnostic techniques, e.g., in the analysis of PD and corona effect in lines and cables, and the analysis of FRA and acoustic signals in PT and DT.

Furthermore, AI-based diagnostic techniques have been introduced recently to overcome some limitations of conventional methods. For instance, AI models are useful to deal with large volumes of data and to represent complex phenomena with minimal requirement of knowledge and physical interpretation (black-box approaches). Examples of AI-based diagnostic techniques are ANN for multiple applications such as interpretation of FRA and DFR in PT and DT, model parameters identification in PV systems, and PD and acoustic analysis in switchgear, lines, and cables. Other AI-based diagnostic techniques have been applied including decision trees, SVM, fuzzy logic, optimization methods, expert systems, kNN, and, to a lesser extent, deep learning models such as CNN, DBF, and LSTM.

9. Prospects for future research

The results of the review demonstrate that research on condition monitoring of some critical MV assets is incipient, especially regarding indices for asset management. Therefore, the need for further research in this topic is highlighted. For instance, HI for PV systems, DT, switchgear, lines, cables, and IT, should be further investigated to obtain a proper health characterization of each asset and its subsystems according to the application, e.g., investment decision-making, network planning, and maintenance scheduling. In this context, straightforward computation of HI based on weighted sums should be explored given its applicability to real-world scenarios, and extensive study cases should be analyzed to demonstrate the benefits of condition monitoring of MV assets for cost-effective operation of distribution networks.

Root cause failure analysis, forecasting of health status, and incipient fault detection based on condition monitoring are also promising areas of research. Applications of these research lines include failure mitigation measures, and predictive maintenance to improve useful life of the assets. Several techniques have demonstrated satisfactory results and capabilities in this direction. For instance, root cause analysis using supervised AI tools such as ANN, SVM, and decision trees has been performed. However, these approaches need large volumes of labelled training data and suffer from overfitting. Moreover, Markov chains have been used for health status forecasting and informed decision-making in asset management. Nonetheless, forecasting capabilities of Markov chains are limited because of the lack of long-term dependencies or representation of trends on data. Other promising approaches for health status forecasting and incipient fault detection include supervised and unsupervised deep learning techniques such as LSTM and deep autoencoders. In this case, drawbacks are on the requirement of large datasets for training and the limited interpretability of models.

As mentioned above, the use of AI techniques for advanced applications in condition monitoring require large datasets. In this regard, new technologies are increasingly adopted by network operators for widespread acquisition of offline and real-time data, e.g., IoT, and advanced metering infrastructures. To properly exploit this huge amount of data, big data frameworks and cloud-based applications with high performance capabilities should be studied in detail, and potentialities should be envisaged.

Finally, lack of integrated tools for asset management and condition monitoring considering most critical assets of the MV network, e.g., PT, PV systems, DT, switchgear, lines, cables, and IT, is observed. Consequently, the analysis, implementation, and research on case studies with the application of comprehensive asset management and condition monitoring methodologies on distribution networks considering all the critical assets is highly encouraged to take advantage of current technology infrastructures and high availability of data to improve the efficient operation of modern power networks. In this context, the use of optimization methods for the integrated operation of distribution networks based on asset management and condition monitoring is seen as a prolific prospective area of research.

10. Conclusion

This paper extends the work presented in ARGENCON 2022 [1]. This extended review examines the literature on critical MV assets operating in distribution networks, namely, PT, PV systems, DT, switchgear, lines and cables, and IT. Moreover, this work is developed in the context of a government-funded research project in Colombia aimed at implementing a digital platform for asset management of distribution network components at the archipelago of San Andrés, Providencia, and Santa Catalina. In this context, the review offers a detailed study of the main technical aspects of each asset and its subsystems, as well as the main parameters for condition assessment. The contributions of this paper are summarized as follows:

- A systematic methodology is adopted for the review and bibliographic data retrieval, where advanced search rules are formulated, and the obtained academic publications are complemented with diverse literature including standards, reports, and relevant white papers.
- A quantitative analysis of the status of research on condition assessment of MV assets is provided, highlighting maturity on the study of PT, the increasing interest in PV systems, and the lack of research on DT, switchgear, lines and cables, and IT.
- A taxonomy for tests and diagnostic techniques used for condition monitoring of MV assets is defined and the extensive literature is classified accordingly (see Tables 4-9).
- Based on the classification above, the main technical aspects related to methods for diagnosis and condition assessment of each asset and indices related to condition assessment are developed.

Regarding the main findings and research gaps, the review identifies the types of tests and diagnostic techniques applied to MV assets, with electrical tests being the most common. Analytical models and AI-based approaches show promise in improving the capability to analyze test results. The prospects for future research lie in developing comprehensive HI for various MV assets, root cause failure analysis, forecasting health status, and incipient fault detection using advanced AI techniques such as deep learning. Leveraging big data and cloud-based applications are suggested to exploit the growing amount of data from emerging technologies, and the integration of condition monitoring and asset management methodologies is crucial for enhancing the efficient operation of modern power distribution networks.

List of symbols

B	Aging factor
F_{AA}	Aging acceleration factor
F_{EQA}	Equivalent acceleration factor
HI_{IT}	Health index of IT
$HI_{IT\text{-}0}$	Initial health index of IT
HI_{L}	Health index of line
$HI_{L,i}$	Health index of line component i
HI_{SG}	Health index of switchgear
$HI_{SG,n}$	Health index obtained in test n on switchgear
LOL	Percentage of loss of life
t	Duration of stress
T_1	Year corresponding to HI_0
$T_{1\text{-}0}$	Expected life
T_2	Year under study
$W_{L,i}$	Weighting factor of line component i
$W_{SG,n}$	Weight for test n on switchgear
θ_H	Hot Spot Temperature

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

Financial support by Fondo de Ciencia, Tecnología e Innovación (FCTeI) of Sistema General de Regalías (SGR) under project BPIN 2021000100223 entitled “Implementation of a digital platform to increase the use of the electrical network assets of the Archipelago of San Andrés and Providencia” is greatly acknowledged.

References

[1] C. Madrid-Chirinos, A.A. Romero-Quete, L. Ontiveros, A. Sarasua, “Condition Assessment of High Voltage Instrument Transformers: A Bibliometric Analysis and Literature Review,” in 2022 IEEE Biennial Congress of Argentina (ARGENCON), IEEE: 1–8, 2022, doi:10.1109/ARGENCON55245.2022.9939980.

[2] ISO-55000, Asset Management - Overview, principles and terminology, Switzerland, 2014.

[3] D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, D. Altman, G. Antes, D. Atkins, V. Barbour, N. Barrowman, J.A. Berlin, J. Clark, M. Clarke, D. Cook, R. D’Amico, J.J. Deeks, P.J. Devereaux, K. Dickersin, M. Egger, E. Ernst, P.C. Gøtzsche, J. Grimshaw, G. Guyatt, J. Higgins, J.P.A. Ioannidis, J. Kleijnen, T. Lang, N. Magrini, D. McNamee, L. Moja, et al., Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement, PLoS Medicine, 2009, doi:10.1371/journal.pmed.1000097.

[4] R.D. Stebbins, D.S. Myers, A.B. Shkolnik, “Furanic compounds in dielectric liquid samples: Review and update of diagnostic interpretation and estimation of insulation ageing,” in Proceedings of the IEEE International Conference on Properties and Applications of Dielectric Materials, 921–926, 2003.

[5] H. de Faria, J.G.S. Costa, J.L.M. Olivas, “A review of monitoring methods for predictive maintenance of electric power transformers based on dissolved gas analysis,” Renewable and Sustainable Energy Reviews, **46**, 201–209, 2015, doi:10.1016/j.rser.2015.02.052.

[6] C. AJ, M.A. Salam, Q.M. Rahman, F. Wen, S.P. Ang, W. Voon, “Causes of transformer failures and diagnostic methods – A review,” Renewable and Sustainable Energy Reviews, **82**, 1442–1456, 2018, doi:10.1016/j.rser.2017.05.165.

[7] R. Soni, B. Mehta, “Review on asset management of power transformer by diagnosing incipient faults and faults identification using various testing methodologies,” Engineering Failure Analysis, **128**, 105634, 2021, doi:10.1016/j.engfailanal.2021.105634.

[8] A.R. Abbasi, “Fault detection and diagnosis in power transformers: a comprehensive review and classification of publications and methods,” Electric Power Systems Research, **209**, 107990, 2022, doi:10.1016/j.epsr.2022.107990.

[9] L. Jin, D. Kim, A. Abu-Siada, S. Kumar, “Oil-Immersed Power Transformer Condition Monitoring Methodologies: A Review,” Energies, **15**(9), 3379, 2022, doi:10.3390/en15093379.

[10] A. Azmi, J. Jasni, N. Azis, M.Z.A.A. Kadir, “Evolution of transformer health index in the form of mathematical equation,” Renewable and

Sustainable Energy Reviews, **76**, 687–700, 2017, doi:10.1016/j.rser.2017.03.094.
[11] IEEE Std 62-1995, IEEE Guide for Diagnostic Field Testing of Electric Power Apparatus - Part 1: Oil Filled Power Transformers, Regulators, and Reactors, 1–64, 1995.
[12] IEC-60450:2004, Measurement of the Average Viscometric Degree of Polymerization of New and Aged Cellulosic Electrically Insulating Materials, 1–45, 2004.
[13] IEEE PC57.149/D9.2, IEEE Draft Guide for the Application and Interpretation of Frequency Response Analysis for Oil Immersed Transformers, 1–68, 2012.
[14] IEEE Std C57.152-2013, IEEE Guide for Diagnostic Field Testing of Fluid-Filled Power Transformers, Regulators, and Reactors, 1–121, 2013.
[15] IEEE PC57.104/D6.2, IEEE Approved Draft Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers, 1–103, 2019.
[16] S. Birlasekaran, F. Fetherston, "Off/on-line FRA condition monitoring technique for power transformer," IEEE Power Engineering Review, **19**(8), 54–56, 1999, doi:10.1109/39.780991.
[17] B. Noirhomme, E. David, H. Garbi, M.C. Lessard, R. Boissonneault, "Application of dielectric response techniques for the condition assessment of power transformers," in CEIDP '05. 2005 Annual Report Conference on Electrical Insulation and Dielectric Phenomena, 2005., IEEE: 273–276, 2005, doi:10.1109/CEIDP.2005.1560674.
[18] F.A. Wegelin, R.S. Magalhães, L.A.L. De Almeida, M. Fontana, "Condition monitoring of power transformers using acoustic signal Prony's analysis," in Conference Record - IEEE Instrumentation and Measurement Technology Conference, 1384–1387, 2005, doi:10.1109/IMTC.2004.1351324.
[19] M. Duval, "Dissolved Gas Analysis and the Duval Triangle," 1–20, 2006.
[20] Z. Abdul-Malek, N. Bashir, H. Ismail, "Furan analysis on power transformers in Malaysia: A field investigation," International Review on Modelling and Simulations, **4**(5), 2227–2233, 2011.
[21] G. Csépes, B. Németh, Cs. Vörös, "Practicable expert system for the improved interpretation of dielectric response diagnostic methods of power transformers," in 44th International Conference on Large High Voltage Electric Systems 2012, 2012.
[22] M.J. Liu, J.Z. Zou, Q.K. Zhou, Y.L. Li, N.P. Yan, "On-Line Condition Monitoring for Power Transformers Based on UHF PD Measurements," Advanced Materials Research, **614–615**, 1158–1162, 2012, doi:10.4028/www.scientific.net/AMR.614-615.1158.
[23] S. Tenbohlen, M. Heindl, M. Jovalekic, "On-site PD Diagnostics, FRA and Moisture Measurement for Power Transformers," IEEJ Transactions on Power and Energy, **132**(2), 134–138, 2012, doi:10.1541/ieejpes.132.134.
[24] H. Gago Garcia, F. Garnacho Vecino, M.A. Sánchez-Uran, J. Ortego La Moneda, I. Uliarte Ranea, "Condition assessment of power transformers in service using pd monitoring," in CIGRE Session 46, 2016.
[25] A.A. Romero-Quete, E.E. Mombello, G. Rattá, "Assessing the loss-of-insulation life of power transformers by estimating their historical loads and ambient temperature profiles using ANNs and Monte Carlo simulations," DYNA (Colombia), **83**(197), 104–113, 2016, doi:10.15446/dyna.v83n197.48134.
[26] A. Hekmati, R. Hekmati, "Optimum acoustic sensor placement for partial discharge allocation in transformers," IET Science, Measurement & Technology, **11**(5), 581–589, 2017, doi:10.1049/iet-smt.2016.0417.
[27] J. Liu, H. Zheng, Y. Zhang, H. Wei, R. Liao, "Grey Relational Analysis for Insulation Condition Assessment of Power Transformers Based Upon Conventional Dielectric Response Measurement," Energies, **10**(10), 1526, 2017, doi:10.3390/en10101526.
[28] R.D. Medina, A.A. Romero, E.E. Mombello, G. Rattá, "Assessing degradation of power transformer solid insulation considering thermal stress and moisture variation," Electric Power Systems Research, **151**, 1–11, 2017, doi:10.1016/j.epsr.2017.04.006.
[29] N. Roopnarine, A. Singh, C.J. Ramlal, S. Rocke, "Network Synthesis of Transformer Winding for FRA Interpretation," in 2016 8th International Conference on Computational Intelligence and Communication Networks (CICN), IEEE: 493–497, 2016, doi:10.1109/CICN.2016.102.
[30] N. Yao, X. Wu, "Transformer fault detection based on infrared power image," Acta Technica CSAV (Ceskoslovensk Akademie Ved), **62**(2), 237–243, 2017.
[31] S. Seifi, A. Akbari, P. Werle, H. Mohseni, A.A.S. Akmal, "A sensitivity analysis of RLC transformer winding model parameters used for a PD localization," in 2018 IEEE 2nd International Conference on Dielectrics (ICD), IEEE: 1–4, 2018, doi:10.1109/ICD.2018.8468521.
[32] M. Koch, M. Anglhuber, "Experiences with Measurement and Analysis of the Dielectric Response of Instrument Transformers," in 2018 IEEE International Conference on High Voltage Engineering and Application (ICHVE), IEEE: 1–4, 2018, doi:10.1109/ICHVE.2018.8642206.
[33] E.T. Mharakurwa, R. Goboza, "Multiparameter-Based Fuzzy Logic Health Index Assessment for Oil-Immersed Power Transformers," Advances in Fuzzy Systems, **2019**, 1–12, 2019, doi:10.1155/2019/2647157.
[34] Y.O. Shaker, "Detection of partial discharge acoustic emission in power transformer," International Journal of Electrical and Computer Engineering (IJECE), **9**(6), 4573, 2019, doi:10.11591/ijece.v9i6.pp4573-4579.
[35] A.F. Cerón, R.A. Lozano, G. Aponte, A.A. Romero, "Hot spot temperature estimation in mineral oil immersed power transformers using support vectors regression," Informacion Tecnologica, **31**(4), 35–44, 2020, doi:10.4067/S0718-07642020000400035.
[36] R. Cselkó, M. Szirtes, G. Csépes, Speciality and Applicability of Advanced Response Methods in Contrast to the Traditional Dielectric Measurements for Condition Assessment of Power Transformers, 1003–1014, 2020, doi:10.1007/978-3-030-31676-1_94.
[37] S.J.T. Shahrabad, V. Ghods, M.T. Askari, "Power transformer fault diagnosis using DGA and artificial intelligence," Recent Advances in Computer Science and Communications, **13**(4), 579–587, 2020, doi:10.2174/2213275912666190212124133.
[38] H.C. Chen, Y. Zhang, M. Chen, "Online DGA Sensor Calibration Using GANN and Data Augmentation," in Proceedings of 2022 IEEE 5th International Electrical and Energy Conference, CIEEC 2022, 1426–1430, 2022, doi:10.1109/CIEEC54735.2022.9846119.
[39] R. Emadifar, N.T. Kalantari, V. Behjat, R. Najjar, "Monitoring and Condition Assessment of Insulation Paper of Distribution Transformers With Novel Oil Spectroscopy Method," IEEE Transactions on Dielectrics and Electrical Insulation, **29**(5), 1904–1912, 2022, doi:10.1109/TDEI.2022.3199687.
[40] X. Ma, Y. Luo, J. Shi, H. Xiong, "Acoustic Emission Based Fault Detection of Substation Power Transformer," Applied Sciences, **12**(5), 2759, 2022, doi:10.3390/app12052759.
[41] M. Zulkifli, M.Z. Muthaza, I. Safaat, S.S. Yi, N. Sahari, "Thermography Study on Distribution Transformer in Substation: Transformer Core Material," International Journal of Engineering Trends and Technology, **71**(3), 9–16, 2022, doi:10.14445/22315381/IJETT-V71I3P202.
[42] A. Naderian, S. Cress, R. Piercy, F. Wang, J. Service, "An Approach to Determine the Health Index of Power Transformers," in Conference Record of the 2008 IEEE International Symposium on Electrical Insulation, IEEE: 192–196, 2008, doi:10.1109/ELINSL.2008.4570308.
[43] M. Arshad, S. Islam, A. Khaliq, "Fuzzy logic approach in power transformers management and decision making," IEEE Transactions on Dielectrics and Electrical Insulation, **21**(5), 2343–2354, 2014, doi:10.1109/TDEI.2014.003859.
[44] A.F. Cerón, D.F. Echeverry, G. Aponte, A.A. Romero, "Health index for power transformers immersed in mineral oil with voltages between 69kV and 230kV using fuzzy logic," Informacion Tecnologica, **26**(2), 107–116, 2015, doi:10.4067/S0718-07642015000200013.
[45] Juan.P. Lata, Diego.P. Chacon-Troya, R.D. Medina, "Improved tool for power transformer health index analysis," in 2017 IEEE XXIV International Conference on Electronics, Electrical Engineering and Computing (INTERCON), IEEE: 1–4, 2017, doi:10.1109/INTERCON.2017.8079656.
[46] C. Ranga, A.K. Chandel, "Expert System for Health Index Assessment of Power Transformers," International Journal on Electrical Engineering and Informatics, **9**(4), 850–865, 2017, doi:10.15676/ijeei.2017.9.4.16.
[47] A.M. Abdullah, R. Ali, S.B. Yaacob, K. Ananda-Rao, N.A. Uloom, "Transformer Health Index by Prediction Artificial Neural Networks Diagnostic Techniques," Journal of Physics: Conference Series, **2312**(1), 012002, 2022, doi:10.1088/1742-6596/2312/1/012002.
[48] D. Luo, R. Xi, L. Che, H. He, "Health condition assessment of transformers based on cross message passing graph neural networks," Frontiers in Energy Research, **10**, 2022, doi:10.3389/fenrg.2022.973736.
[49] K. Ramalingam, C. Indulkar, Solar Energy and Photovoltaic Technology, 2017, doi:10.1016/B978-0-12-804208-3.00003-0.
[50] International Energy Agency, Trends in photovoltaic applications 2022 - Report IEA PVPS T1-43:2022, 2022.
[51] L. Hernández-Callejo, S. Gallardo-Saavedra, V. Alonso-Gómez, "A review of photovoltaic systems: Design, operation and maintenance," Solar Energy, **188**, 426–440, 2019, doi:10.1016/j.solener.2019.06.017.
[52] S. Lu, B.T. Phung, D. Zhang, "A comprehensive review on DC arc faults and their diagnosis methods in photovoltaic systems," Renewable and Sustainable Energy Reviews, **89**, 88–98, 2018, doi:10.1016/j.rser.2018.03.010.

[53] A. Mellit, G.M. Tina, S.A. Kalogirou, "Fault detection and diagnosis methods for photovoltaic systems: A review," Renewable and Sustainable Energy Reviews, **91**, 1–17, 2018, doi:10.1016/j.rser.2018.03.062.

[54] D.S. Pillai, N. Rajasekar, "A comprehensive review on protection challenges and fault diagnosis in PV systems," Renewable and Sustainable Energy Reviews, **91**, 18–40, 2018, doi:10.1016/j.rser.2018.03.082.

[55] D.S. Pillai, N. Rajasekar, "Metaheuristic algorithms for PV parameter identification: A comprehensive review with an application to threshold setting for fault detection in PV systems," Renewable and Sustainable Energy Reviews, **82**, 3503–3525, 2018, doi:10.1016/j.rser.2017.10.107.

[56] Syafaruddin, D.S. Zinger, "Review on methods of fault diagnosis in photovoltaic system applications," Journal of Engineering Science and Technology Review, **12**(5), 53–66, 2019, doi:10.25103/jestr.125.07.

[57] L.B. Bosman, W.D. Leon-Salas, W. Hutzel, E.A. Soto, "PV system predictive maintenance: Challenges, current approaches, and opportunities," Energies, **16**(3), 2020, doi:10.3390/en13061398.

[58] A. Kumaradurai, Y. Teekaraman, T. Coosemans, M. Messagie, "Fault detection in photovoltaic systems using machine learning algorithms: A review," Institute of Electrical and Electronics Engineers Inc., Automotive Technology Research Group (MOBI), Vrije Universiteit Brussel (VUB), Brussels, Belgium, 2020, doi:10.1109/ICOT51877.2020.9468768.

[59] T. Berghout, M. Benbouzid, T. Bentrcia, X. Ma, S. Djurović, L.-H. Mouss, "Machine learning-based condition monitoring for pv systems: State of the art and future prospects," Energies, **14**(19), 2021, doi:10.3390/en14196316.

[60] T. Berghout, M. Benbouzid, X. Ma, S. Djurovic, L.-H. Mouss, "Machine Learning for Photovoltaic Systems Condition Monitoring: A Review," IEEE Computer Society, University of Batna2, Department of Industrial Engineering and Manufacturing, Batna, Algeria, 2021, doi:10.1109/IECON48115.2021.9589423.

[61] C. Feng, Y. Liu, J. Zhang, "A taxonomical review on recent artificial intelligence applications to PV integration into power grids," International Journal of Electrical Power and Energy Systems, **132**, 2021, doi:10.1016/j.ijepes.2021.107176.

[62] A.W. Kandeal, M.R. Elkadeem, A. Kumar Thakur, G.B. Abdelaziz, R. Sathyamurthy, A.E. Kabeel, N. Yang, S.W. Sharshir, "Infrared thermography-based condition monitoring of solar photovoltaic systems: A mini review of recent advances," Solar Energy, **223**, 33–43, 2021, doi:10.1016/j.solener.2021.05.032.

[63] M. Mansouri, M. Trabelsi, H. Nounou, M. Nounou, "Deep Learning-Based Fault Diagnosis of Photovoltaic Systems: A Comprehensive Review and Enhancement Prospects," IEEE Access, **9**, 126286–126306, 2021, doi:10.1109/ACCESS.2021.3110947.

[64] A.K. V de Oliveira, M. Aghaei, R. Rüther, "Automatic Inspection of Photovoltaic Power Plants Using Aerial Infrared Thermography: A Review," Energies, **15**(6), 2022, doi:10.3390/en15062055.

[65] M.I.S. Guerra, F.M.U. de Araújo, J.T. de Carvalho Neto, R.G. Vieira, "Survey on adaptative neural fuzzy inference system (ANFIS) architecture applied to photovoltaic systems," Energy Systems, 2022, doi:10.1007/s12667-022-00513-8.

[66] Z. Yuan, G. Xiong, X. Fu, "Artificial Neural Network for Fault Diagnosis of Solar Photovoltaic Systems: A Survey," Energies, **15**(22), 8693, 2022, doi:10.3390/en15228693.

[67] D.C. Jordan, S.R. Kurtz, "Photovoltaic Degradation Rates-an Analytical Review," Progress in Photovoltaics: Research and Applications, **21**(1), 12–29, 2013, doi:10.1002/pip.1182.

[68] O.A. Alimi, E.L. Meyer, O.I. Olayiwola, "Solar Photovoltaic Modules' Performance Reliability and Degradation Analysis—A Review," Energies, **15**(16), 5964, 2022, doi:10.3390/en15165964.

[69] IEC 62446-1:2016, Photovoltaic (PV) systems - Requirements for testing, documentation and maintenance - Part 1: Grid connected systems - Documentation, commissioning tests and inspection, 1–83, 2016.

[70] IEEE Std 1547-2018, IEEE Standard for Interconnection and Interoperability of Distributed Energy Resources with Associated Electric Power Systems Interfaces, 1–138, 2018.

[71] IEEE Std 1526-2020, IEEE Recommended Practice for Testing the Performance of Stand-Alone Photovoltaic Systems, 1–32, 2020.

[72] I. Bodnár, D. Matusz-Kalász, R.R. Boros, R. Lipták, "Condition Assessment of Solar Modules by Flash Test and Electroluminescence Test," Coatings, **11**(11), 1361, 2021, doi:10.3390/coatings11111361.

[73] V. Esen, S. Saglam, B. Oral, O. Ceylan Esen, "Toward Class AAA LED Large Scale Solar Simulator With Active Cooling System for PV Module Tests," IEEE Journal of Photovoltaics, **12**(1), 364–371, 2022, doi:10.1109/JPHOTOV.2021.3117912.

[74] S. Choi, S. Park, J. Hong, J. Won, "A Design and Validation of 400 W PV Emulator Using Simple Equivalent Circuit for PV Power System Test," Energies, **16**(4), 1561, 2023, doi:10.3390/en16041561.

[75] C. Han, H. Lee, "A field-applicable health monitoring method for photovoltaic system," Reliability Engineering & System Safety, **184**, 219–227, 2019, doi:10.1016/j.ress.2018.01.002.

[76] P. Ramanan, K. Kalidasa Murugavel, A. Karthick, "Performance analysis and energy metrics of grid-connected photovoltaic systems," Energy for Sustainable Development, **52**, 104–115, 2019, doi:10.1016/j.esd.2019.08.001.

[77] SolarPower Europe, Asset management best practice guidelines Version 2.0, 2020.

[78] IEC, IEC 60076-7. Power transformers - Loading guide for mineral-oil-immersed power transformers, Genova, Switzerland, 2018.

[79] IEC, IEC 60076-7. Power transformers - Loading guide for oil-immersed power transformers., Genova, Switzerland, 2005.

[80] F. Kabir, B. Foggo, N. Yu, "Data driven predictive maintenance of distribution transformers," IEEE Computer Society, Department of Electrical and Engineering, University of California, Riverside, CA 92507, United States: 312–316, 2018, doi:10.1109/CICED.2018.8592417.

[81] G.C. Jaiswal, M.S. Ballal, P.A. Venikar, D.R. Tutakne, H.M. Suryawanshi, "Genetic algorithm–based health index determination of distribution transformer," International Transactions on Electrical Energy Systems, **28**(5), 2018, doi:10.1002/etep.2529.

[82] IEEE, IEEE Std C57.12.00-2010 - Standard for General Requirements for Liquid-Immersed Distribution, Power, and Regulating Transformers, New York, United State, 2010.

[83] K. Azimi Hosseini, M. Hajiaghapour-Moghimi, E. Hajipour, M. Vakilian, "Advanced distribution transformer asset management embracing its relocation: A cost saving approach," IET Generation, Transmission and Distribution, **16**(12), 2370–2385, 2022, doi:10.1049/gtd2.12456.

[84] Meghalaya State Electricity Board, Manual for Maintenance of Distribution Transformers, 2010.

[85] M.M. Youssef, R.A. Ibrahim, H. Desouki, M.M.Z. Moustafa, "An Overview on Condition Monitoring & Health Assessment Techniques for Distribution Transformers," Institute of Electrical and Electronics Engineers Inc., Arab Academy for Science, Technology & Maritime Transport, Alexandria, Egypt: 187–192, 2022, doi:10.1109/ICGEA54406.2022.9791900.

[86] IEEE, IEEE Std 62-1995 - Guide for Diagnostic Field Testing of Electric Power Apparatus - Part 1: Oil Filled Power Transformers, Regulators, and Reactors., 1995.

[87] R. Behkam, H. Karami, M.S. Naderi, G.B. Gharehpetian, "Condition Monitoring of Distribution Transformers Using Frequency Response Traces and Artificial Neural Network to Detect the Extent of Windings Axial Displacements," Institute of Electrical and Electronics Engineers Inc., Amirkabir University of Technology (Tehran Polytechnic), Department of Electrical Engineering, Iran: 18–23, 2022, doi:10.1109/EPDC56235.2022.9817296.

[88] A. Sbravati, R. Breazeal, D. Robalino, "Practical Considerations for the Usage of Ester Fluids in Distribution Transformers," Institute of Electrical and Electronics Engineers Inc., Cargill Bioindustrial, Minneapolis, United States: 366–371, 2022, doi:10.1109/EIC51169.2022.9833213.

[89] R. Chumnuan, D. Rerkpreedapong, "A practicable framework for risk assessment of distribution transformers using PEA smart meter data," in: Y., K., ed., Institute of Electrical and Electronics Engineers Inc., Provincial Electricity Authority (PEA), Engineering Division Region 4, Bangkok, Thailand: 590–594, 2021, doi:10.1109/ECTI-CON51831.2021.9454682.

[90] R. Prananda, H.S. Dini, T.W.O. Putri, "Design of Load Balancing Method on Secondary Distribution Network Using Artificial Intelligence Based on Fuzzy Logic," Institute of Electrical and Electronics Engineers Inc., Institut Teknologi PLN, Faculty of Electricity and Renewable Energy, Jakarta, Indonesia: 372–377, 2021, doi:10.1109/ICHVEPS53178.2021.9600977.

[91] C.-S. Tu, C.-Y. Yang, M.-T. Tsai, "An optimal phase arrangement of distribution transformers under risk assessment," Energies, **13**(21), 2020, doi:10.3390/en13215852.

[92] Z. Yang, Y. Shen, R. Zhou, F. Yang, Z. Wan, Z. Zhou, "A transfer learning fault diagnosis model of distribution transformer considering multi-factor situation evolution," IEEJ Transactions on Electrical and Electronic Engineering, **15**(1), 30–39, 2020, doi:10.1002/tee.23024.

[93] S. Kulkarni, K. Ashok, F. Lambert, D. Divan, "Asset Monitoring using Smart Sensing and Advanced Analytics for the Distribution Network," Institute of Electrical and Electronics Engineers Inc., Center for Distributed Energy, Georgia Institute of Technology, Atlanta, GA, United States, 2019, doi:10.1109/NAPS46351.2019.9000254.

[94] G.C. Jaiswal, M.S. Ballal, D.R. Tutakne, P. Vishnu, "Intelligent condition monitoring system for distribution transformer and health status diagnosis," Institute of Electrical and Electronics Engineers Inc., Additional Executive Engineer, Maharashtra State Elec. Distr. Co. Ltd. and Res. Scholar Shri Ramdeobaba Coll. of Eng. and Mgmt., Nagpur, India: 1–6, 2018, doi:10.1109/PICC.2018.8384786.

[95] M. Mahoor, A. Majzoobi, A. Khodaei, "Distribution asset management through coordinated microgrid scheduling," IET Smart Grid, **1**(4), 159–168, 2018, doi:10.1049/iet-stg.2018.0076.

[96] A.A. Nelson, G.C. Jaiswal, M.S. Ballal, D.R. Tutakne, "Remote condition monitoring system for distribution transformer," Institute of Electrical and Electronics Engineers Inc., Dept of Elect Engg, VNIT, Nagpur, India, 2015, doi:10.1109/NPSC.2014.7103848.

[97] Y.B. Chen, Y. Sun, X.R. Sun, G.H. Sheng, X.C. Jiang, Real-time temperature on-line monitoring and analysis system for transformers, **521**, 409–413, 2014, doi:10.4028/www.scientific.net/AMM.521.409.

[98] B. Noirhomme, J. Jalbert, M.-C. Lessard, F. Brochu, S. Hachichi, "Pole mounted distribution transformers insulation condition assessment," Hydro-Québec - IREQ, Varennes, Canada: 239–243, 2013, doi:10.1109/EIC.2013.6554241.

[99] S. Carvalhosa, H. Leite, F. Branco, C.A. Sá, A.M. Moura, R.C. Lopes, M. Soares, "Survey on the Advancements of Dielectric Fluids and Experiment Studies for Distribution Power Transformers," Renewable Energy and Power Quality Journal, **19**, 97–102, 2021, doi:10.24084/repqj19.225.

[100] Y. Wang, L. Hu, J. Chen, Y. Ren, "Health status diagnosis of distribution transformers based on big data mining," Transactions on Emerging Telecommunications Technologies, **33**(2), 2022, doi:10.1002/ett.3759.

[101] N.S. Shariffuddin, N. Azis, A.M. Selva, M.S. Yahaya, J. Jasni, M.A. Talib, "Investigation on the Relationship between Failure Rates and Health Index of Distribution Transformer Population," Institute of Electrical and Electronics Engineers Inc., Universiti Putra Malaysia, Advanced Lightning, Power and Energy Research Centre (ALPER), Serdang, Malaysia: 119–122, 2021, doi:10.1109/ICPADM49635.2021.9493977.

[102] S.A. Ghani, Y.H.M. Thayoob, Y.Z.Y. Ghazali, M.S.A. Khiar, I.S. Chairul, "Condition monitoring of distribution transformer's mechanical parts using sweep frequency response analysis (SFRA)," Elsevier Ltd, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia: 469–476, 2013, doi:10.1016/j.proeng.2013.12.208.

[103] S. Zhu, C.J. Kikkert, N. Ertugrul, "A wide bandwidth, on-line impedance measurement method for power systems, based On PLC techniques," in 2014 IEEE International Symposium on Circuits and Systems (ISCAS), IEEE, School of Electrical and Electronic Engineering, University of Adelaide, Adelaide, SA, Australia: 1167–1170, 2014, doi:10.1109/ISCAS.2014.6865348.

[104] J. Javid, M.A. Mughal, M. Karim, "Using kNN Algorithm for classification of Distribution transformers Health index," Institute of Electrical and Electronics Engineers Inc., HITEC University Taxila, Department of Electrical Engineering, Taxila, Pakistan, 2021, doi:10.1109/ICIC53490.2021.9693013.

[105] R.R. Pawar, S.B. Deosarkar, "Health condition monitoring system for distribution transformer using Internet of Things (IoT)," Institute of Electrical and Electronics Engineers Inc., Department of Electronics and Telecommunication Engineering, Dr. Babasaheb Ambedkar Technological University, Lonere, Raigad, Maharashtra, India: 117–122, 2018, doi:10.1109/ICCMC.2017.8282650.

[106] P. Henault, "Detection of internal arcing faults in distribution transformers," IFD Corporation, United States, 2011, doi:10.1109/TDCLLM.2011.6042220.

[107] L. Bin, S. Sheng, L. Yuan, F. Qiang, X. Yunfeng, L. Guanke, "Spectro-Temporal Self-similarity based Identification of Corrupted Acoustic Signal of Distribution Transformer in Noisy Environment," IEEE Transactions on Power Delivery, 1–12, 2022, doi:10.1109/TPWRD.2022.3181978.

[108] IEC, IEC 60270 - High-voltage test techniques-Partial discharge measurements., 2001.

[109] Network Operators in Great Britain, DNO COMMON NETWORK ASSET INDICES METHODOLOGY Health & Criticality-Version 2.1, 2021.

[110] G.C. Jaiswal, M.S. Ballal, D.R. Tutakne, "Health index based condition monitoring of distribution transformer," Institute of Electrical and Electronics Engineers Inc., Department of Electrical Engineering, Shri Ramdeobaba College of Engineering & Management, Nagpur, India: 1–5, 2017, doi:10.1109/PEDES.2016.7914222.

[111] G.C. Montanari, R. Ghosh, L. Cirioni, G. Galvagno, S. Mastroeni, "Partial Discharge Monitoring of Medium Voltage Switchgears: Self-Condition Assessment Using an Embedded Bushing Sensor," IEEE Transactions on Power Delivery, **37**(1), 85–92, 2022, doi:10.1109/TPWRD.2021.3053658.

[112] I.A. Metwally, "Technology progress in high-voltage gas-insulated substations," IEEE Potentials, **29**(6), 25–32, 2010, doi:10.1109/MPOT.2010.939085.

[113] X. Zhang, E. Gockenbach, "Component reliability modeling of distribution systems based on the evaluation of failure statistics," IEEE Transactions on Dielectrics and Electrical Insulation, **14**(5), 1183–1191, 2007, doi:10.1109/TDEI.2007.4339478.

[114] X. Zhang, E. Gockenbach, V. Wasserberg, H. Borsi, "Estimation of the lifetime of the electrical components in distribution networks," IEEE Transactions on Power Delivery, **22**(1), 515–522, 2007, doi:10.1109/TPWRD.2006.876661.

[115] K. Niayesh, M. Runde, Power Switching Components, Springer International Publishing, Cham, 2017, doi:10.1007/978-3-319-51460-4.

[116] Power and Water Corporation, "Asset Management Plan – Distribution Switchgear," 1–54, 2019.

[117] C. Lowsley, N. Davies, D. Miller, "EFFECTIVE CONDITION ASSESSMENT OF MV SWITCHGEAR," in ea technology, 1–8.

[118] Cigré, Non-intrusive methods for condition assessment of distribution and transmission switchgear, Cigré, 2018.

[119] Y. Alsumaidaee, C. Yaw, S. Koh, S. Tiong, C. Chen, K. Ali, "Review of Medium-Voltage Switchgear Fault Detection in a Condition-Based Monitoring System by Using Deep Learning," Energies, **15**(18), 6762, 2022, doi:10.3390/en15186762.

[120] K. Perdon, L. Cavalli, M. Testa, "INNOVATIVE ANALYTICS TO ESTIMATE THE PROBABILITY OF FAILURE AND REMAINING USEFUL LIFE OF MEDIUM VOLTAGE BREAKERS," in Cired 2017, 12–15, 2017.

[121] A.A. Bhole, W.Z. Gandhare, "An Overview of Dynamic Contact Resistance Measurement of HV Circuit Breakers," Journal of The Institution of Engineers (India): Series B, **97**(2), 219–226, 2016, doi:10.1007/s40031-014-0164-2.

[122] IEC, IEC 62271-1:2017/AMD1:2021 Amendment 1 - High-voltage switchgear and controlgear - Part 1: Common specifications for alternating current switchgear and controlgear, 2021.

[123] X. Jiang, B. Stephen, T. Chandarasupsang, S.D.J. McArthur, B.G. Stewart, "A Gaussian Process Based Fleet Lifetime Predictor Model for Unmonitored Power Network Assets," IEEE Transactions on Power Delivery, **38**(2), 979–987, 2023, doi:10.1109/TPWRD.2022.3203161.

[124] J. Li, Y. Sun, N. Dong, Z. Zhao, "A Novel Contact Temperature Calculation Algorithm in Distribution Switchgears for Condition Assessment," IEEE Transactions on Components, Packaging and Manufacturing Technology, **9**(2), 279–287, 2019, doi:10.1109/TCPMT.2018.2886404.

[125] IEC, IEC 62271-100:2021 High-voltage switchgear and controlgear - Part 100: Alternating-current circuit-breakers, 600, 2021.

[126] EPRI, Plant Support Engineering : Life Cycle Management Planning Sourcebook : Medium-Voltage Switchgear, 2006.

[127] F.Á. Gómez, R. Albarracín-Sánchez, F.G. Vecino, R.G. Arrabé, "Diagnosis of insulation condition of mv switchgears by application of different partial discharge measuring methods and sensors," Sensors (Switzerland), **18**(3), 2018, doi:10.3390/s18030720.

[128] G.A. Hussain, M. Shafiq, L. Kumpulainen, F. Mahmood, M. Lehtonen, "Performance evaluation of noise reduction method during on-line monitoring of MV switchgear for PD measurements by non-intrusive sensors," International Journal of Electrical Power and Energy Systems, **64**, 596–607, 2015, doi:10.1016/j.ijepes.2014.07.057.

[129] W. Kang, Dohoon. Lee, S. Ham, J. Lee, J. Lee, D. Kim, T. Kwon, "Development of a UHF PD detection system to estimate the dielectric condition of a Medium Voltage Switchgear," in Conference Record of the 2008 IEEE International Symposium on Electrical Insulation, 377–380, 2008, doi:10.1109/ELINSL.2008.4570353.

[130] J.A. Kay, G.A. Hussain, M. Lehtonen, L. Kumpulainen, "New preemptive arc-fault detection techniques in medium-voltage switchgear and motor controls," IEEE Transactions on Industry Applications, **52**(1), 740–750, 2016, doi:10.1109/TIA.2015.2466620.

[131] S. Ishak, S.P. Koh, J.D. Tan, S.K. Tiong, C.P. Chen, C.T. Yaw, "Arcing Faults Detection in Switchgear with Extreme Learning Machine," Journal of Physics: Conference Series, **2319**(1), 012007, 2022, doi:10.1088/1742-6596/2319/1/012007.

[132] X. Zheng, Z. Wang, "Partial Discharge Fault Diagnosis of Switchgear Based on APSO-BP Algorithm," in 2021 IEEE 2nd China International Youth Conference on Electrical Engineering (CIYCEE), IEEE, Dalian

University of Technology, China: 1–7, 2021, doi:10.1109/CIYCEE53554.2021.9676811.
[133] Q. Che, H. Wen, X. Li, Z. Peng, K.P. Chen, "Partial Discharge Recognition Based on Optical Fiber Distributed Acoustic Sensing and a Convolutional Neural Network," IEEE Access, **7**, 101758–101764, 2019, doi:10.1109/ACCESS.2019.2931040.
[134] S. Barrios, D. Buldain, M.P. Comech, I. Gilbert, "Partial Discharge Identification in MV Switchgear Using Scalogram Representations and Convolutional AutoEncoder," IEEE Transactions on Power Delivery, **36**(6), 3448–3455, 2021, doi:10.1109/TPWRD.2020.3042934.
[135] H. Li, X. Tong, F. Yang, D. Huang, "Switchgear Insulation Condition Assessment with Multi-Dimensional Feature Database Based on an Optimal Grade Clustering Method," in: B., X. and K., M., eds., in 2021 IEEE 5th Information Technology,Networking,Electronic and Automation Control Conference (ITNEC), IEEE, Shanghai Marine Equipment Research Institute, Shanghai, China: 294–299, 2021, doi:10.1109/ITNEC52019.2021.9587116.
[136] Z. Zhang, H. Xia, S. Cao, C. Liu, "Research on feature extraction and fault diagnosis of mechanical vibration and sound signal of Disconnector," in 2022 IEEE 5th International Electrical and Energy Conference (CIEEC), IEEE, Shandong University, School of Electrical Engineering, Jinan, China: 1119–1124, 2022, doi:10.1109/CIEEC54735.2022.9846229.
[137] Z. Qiu, J. Ruan, D. Huang, "Mechanical fault diagnosis of outdoor high-voltage disconnector," IEEJ Transactions on Electrical and Electronic Engineering, **11**(5), 556–563, 2016, doi:10.1002/tee.22273.
[138] C. Nicolaou, A. Mansour, K. Van Laerhoven, "On-site Online Condition Monitoring of Medium-Voltage Switchgear Units," ACM International Conference Proceeding Series, 1–8, 2021, doi:10.1145/3494322.3494323.
[139] S. Yan, G. Zhang, B. Chen, C. Li, Y. Liu, J. Yang, Fault Diagnosis Technology of Low Voltage Circuit Breaker Based on Multilayer Perceptron, Springer Science and Business Media Deutschland GmbH, Power Supply Branch of Beijing Metro Operation Co., Ltd., Beijing, 100084, China: 332–344, 2022, doi:10.1007/978-981-19-1532-1_35.
[140] A.S.N. Huda, S. Taib, "Application of infrared thermography for predictive/preventive maintenance of thermal defect in electrical equipment," Applied Thermal Engineering, **61**(2), 220–227, 2013, doi:10.1016/J.APPLTHERMALENG.2013.07.028.
[141] Z. Chen, T. Yuan, M. Xu, Y. Yu, Z. Xie, Y. Liu, "Research on Fault Diagnosis of Switchgear Based on Temperature Cloud Image Technology," IOP Conference Series: Earth and Environmental Science, **1044**(1), 012007, 2022, doi:10.1088/1755-1315/1044/1/012007.
[142] Y. Yan, D. Qi, H. Gu, "A Real-Time IR-Fusion Switchgear Contact Monitoring System (SCMS)," IEEE Access, **5**, 12114–12124, 2017, doi:10.1109/ACCESS.2017.2698060.
[143] R. Salustiano, R.M. Capelini, S.R. De Abreu, M.L.B. Martinez, I.C. Tavares, G.M.F. Ferraz, M.A.A. Romano, "Development of new methodology for insulators inspections on aerial distribution lines based on partial discharge detection tools," ICHVE 2014 - 2014 International Conference on High Voltage Engineering and Application, 2014, doi:10.1109/ICHVE.2014.7035429.
[144] M. Shafiq, I. Kiitam, P. Taklaja, L. Kutt, K. Kauhaniemi, I. Palu, "Identification and Location of PD Defects in Medium voltage Underground Power Cables Using High Frequency Current Transformer," IEEE Access, **7**, 103608–103618, 2019, doi:10.1109/ACCESS.2019.2930704.
[145] N. Davari, G. Akbarizadeh, E. Mashhour, "Intelligent Diagnosis of Incipient Fault in Power Distribution Lines Based on Corona Detection in UV-Visible Videos," IEEE Transactions on Power Delivery, **36**(6), 3640–3648, 2021, doi:10.1109/TPWRD.2020.3046161.
[146] H. Manninen, J. Kilter, M. Landsberg, "Advanced condition monitoring method for high voltage overhead lines based on visual inspection," IEEE Power and Energy Society General Meeting, **2018-August**, 2018, doi:10.1109/PESGM.2018.8586498.
[147] D.R. Morais, J.G. Rolim, J. Coser, H.H. Zürn, "Reliability centered maintenance for capacitor voltage transformers," in 2006 9th International Conference on Probabilistic Methods Applied to Power Systems, PMAPS, 2006, doi:10.1109/PMAPS.2006.360281.
[148] Y. Wu, M. Liu, Z. Ji, Z. Diao, Y. Wu, F. Yang, "A Fault Diagnosis System of All-Fiber Optical Current Transformer with Correlation Detection," in Journal of Physics: Conference Series, 2020, doi:10.1088/1742-6596/1550/4/042041.
[149] P. Werelius, B. Holmgren, U. Gafvert, "Diagnosis of medium voltage XLPE cables by high voltage dielectric spectroscopy," in ICSD'98. Proceedings of the 1998 IEEE 6th International Conference on Conduction and Breakdown in Solid Dielectrics (Cat. No.98CH36132), 79–84, 1998, doi:10.1109/ICSD.1998.709231.
[150] H. van Jaarsveldt, R. Gouws, "Condition Monitoring of Medium Voltage Electrical Cables by Means of Partial Discharge Measurements," SAIEE Africa Research Journal, **105**(4), 136–146, 2014, doi:10.23919/SAIEE.2014.8538345.
[151] C.-K. Chang, C.-S. Lai, R.-N. Wu, "Decision tree rules for insulation condition assessment of pre-molded power cable joints with artificial defects," IEEE Transactions on Dielectrics and Electrical Insulation, **26**(5), 1636–1644, 2019, doi:10.1109/TDEI.2019.008208.
[152] R.M. Arias Velásquez, J.V. Mejía Lara, "Ruptures in overhead ground wire — Transmission line 220 kV," Engineering Failure Analysis, **87**, 1–14, 2018, doi:10.1016/J.ENGFAILANAL.2018.01.003.
[153] International Electrotechnical Commission - IEC, INTERNATIONAL STANDARD - IEC 61395:1998 Overhead electrical conductors - Creep test procedures for stranded conductors, 1998.
[154] American Society for Testing and Materials International - ASTM, ASTM B1008-18 Standard Test Method for Stress-Strain Testing for Overhead Electrical Conductors.
[155] American Society for Testing and Materials International - ASTM, ASTM D1036-99 Standard Test Methods of Static Tests of Wood Poles, 2017.
[156] American Society for Testing and Materials International - ASTM, ASTM C1089-19 Standard Specification for Spun Cast Prestressed Concrete Poles.
[157] International Electrotechnical Commission - IEC, NTERNATIONAL STANDARD - IEC 60383-1 Insulators for overhead lines with a nominal voltage above 1000 V - Part 1: Ceramic or glass insulator units for a.c. systems - Definitions, test methods and acceptance crite, 1993.
[158] ANSI/NEMA, ANSI/NEMA C29.1-2018 Test Methods for Electrical Power Insulators, 2018.
[159] K. Ibrahim, R.M. Sharkawy, H.K. Temraz, M.M.A. Salama, "Reliability calculations based on an enhanced transformer life expectancy model," Ain Shams Engineering Journal, **13**(4), 2022, doi:10.1016/j.asej.2021.101661.
[160] D.M.R. Vanegas, S.M. Mahajan, "Effects of thermal accelerated ageing on a medium voltage oil-immersed current transformer," Conference Record of IEEE International Symposium on Electrical Insulation, 470–473, 2008, doi:10.1109/ELINSL.2008.4570375.
[161] IEEE Standard Requirements for Instrument Transformers, IEEE Std C57.13-2016 (Revision of IEEE Std C57.13-2008), 1–96, 2016, doi:10.1109/IEEESTD.2016.7501435.
[162] International Electrotechnical Commission - IEC, INTERNATIONAL STANDARD IEC 60044-2 Instrument transformers -Part 2: Inductive voltage transformers, 2003.
[163] International Electrotechnical Commission - IEC, INTERNATIONAL STANDARD - IEC 60044-1 Instrument transformers - Part 1: Current transformers, 2003.
[164] Z. Tiancheng, Y. Ying, S. Jinhai, X. Zhihao, L. Jiashuai, "Infrared Image Segmentation Method of Current Transformer Based on DeepLabv3+ Network," in 2022 IEEE 5th International Electrical and Energy Conference (CIEEC), IEEE, State Grid Jilin Electric Power Co., Ltd., Electric Power Research Institute, Changchun, China: 3111–3115, 2022, doi:10.1109/CIEEC54735.2022.9846719.
[165] B.K. Gupta, J. Densley, A. Narang, "Review of diagnostic techniques for oil-paper insulated instrument transformers," in 2010 IEEE International Symposium on Electrical Insulation, 1–5, 2010, doi:10.1109/ELINSL.2010.5549786.
[166] A. Jeanneton, C. Perrier, A. Beroual, "Thermal Aging Study of Alternative Liquids for Oil-Paper High-Voltage Insulation in Current Transformers," Proceedings - IEEE International Conference on Dielectric Liquids, **2022-May**, 2022, doi:10.1109/ICDL49583.2022.9830941.
[167] International Electrotechnical Commission - IEC, INTERNATIONAL STANDARD IEC 60599:2022 Mineral oil-filled electrical equipment in service - Guidance on the interpretation of dissolved and free gases analysis, 2022.
[168] Ning Hao, Zhuo Dong, "Condition assessment of current transformer based on multi-classification support vector machine," in Proceedings 2011 International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE), IEEE, NORTH CHINA BAODING ELECTRIC POWER VOC. and TECH.COLLEGE, Baoding, Hebei, China: 2402–2405, 2011, doi:10.1109/TMEE.2011.6199705.
[169] O.D.L. Campos, H. Zhu, "Condition assessment of 653 instrument transformers with on-line PD testing," in 2016 International Conference on Condition Monitoring and Diagnosis (CMD), IEEE, Costa Rican Institute of Electricity, Costa Rica: 168–171, 2016, doi:10.1109/CMD.2016.7757797.

[170] G.A.T.N. Aravinda, D.C. Attaragama, I.T. Palawatte, M.A.R.M. Fernando, G.A. Jayantha, J.R.S.S. Kumara, "Condition monitoring of current transformers using frequency domain dielectric spectroscopy," in 2013 IEEE 8th International Conference on Industrial and Information Systems, IEEE, Faculty of Engineering, University of Peradeniya, Sri Lanka: 168–173, 2013, doi:10.1109/ICIInfS.2013.6731975.

[171] G. Iyer, R. Gorur, A. Krivda, P. Mahonen, "Prediction of electrical performance of medium voltage epoxy insulated equipment," IEEE Transactions on Dielectrics and Electrical Insulation, **17**(2), 334–342, 2010, doi:10.1109/TDEI.2010.5448086.

[172] T. Lei, M. Faifer, R. Ottoboni, S. Toscani, "On-line fault detection technique for voltage transformers," Measurement: Journal of the International Measurement Confederation, **108**, 193–200, 2017, doi:10.1016/j.measurement.2017.03.002.

[173] L. Zhang, H. Chen, Q. Wang, N. Nayak, Y. Gong, A. Bose, "A Novel On-Line Substation Instrument Transformer Health Monitoring System Using Synchrophasor Data," IEEE Transactions on Power Delivery, **34**(4), 1451–1459, 2019, doi:10.1109/TPWRD.2019.2905426.

[174] B. Cui, A.K. Srivastava, P. Banerjee, "Synchrophasor-Based Condition Monitoring of Instrument Transformers Using Clustering Approach," IEEE Transactions on Smart Grid, **11**(3), 2688–2698, 2020, doi:10.1109/TSG.2019.2960043.

[175] S. Tee, Q. Liu, Z. Wang, F. Hafid, P. Tournet, "Failure investigation and asset management of combined measuring instrument transformers," High Voltage, **6**(1), 61–70, 2021, doi:10.1049/hve2.12029.

[176] H. Guo, L. Guo, "Health index for power transformer condition assessment based on operation history and test data," Energy Reports, **8**, 9038–9045, 2022, doi:10.1016/j.egyr.2022.07.041.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 8, No. 5, 55-61 (2023)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Science & Technology

ASTES Journal
ISSN: 2415-6698

The Graded Multidisciplinary Model: Fostering Instructional Design for Activity Development in STEM/STEAM Education*

Mauricio Flores-Nicolás*, Magally Martínez-Reyes, Felipe de Jesús Matías-Torres

Universidad Autónoma del Estado de México – Centro Universitario Valle de Chalco, Estado de México, 56615, México

ARTICLE INFO

Article history:
Received: 01 August, 2023
Accepted: 30 September, 2023
Online: 30 October, 2023

Keywords:
21st Century Skills
Gamification
Process Model
Middle School
Educational Technology

ABSTRACT

In a challenging and increasingly technological world, it is important to promote critical thinking, multidisciplinary problem solving, and collaboration through STEAM education; however, there are important economic, administrative, and especially pedagogical management limitations for its implementation at the secondary level. Therefore, this paper presents systematic recommendations for an effective and sustainable implementation of STEAM education in educational institutions through the Gradual Multidisciplinary Model (GMM), which allows the identification and specific adaptation of STEAM knowledge through the topic of logic gates related to the representations of disjunction and conjunction in Boolean algebra (university content) to its physical representation in Minecraft (high school content). The quasi-experimental method allows to evaluate the results through the application of a pre-test designed to measure logical-mathematical thinking and a post-test designed to measure the level of understanding of practical skills and the students' perception of the learning experience. The results obtained by t-student show that there is a significantly high difference between the means and suggest that the educational intervention orchestrated by the GMM had a significant impact on the performance and skills of the participants, since different levels of understanding (gradualness) and perception of the concepts related to logic gates could be identified; while the qualitative assessment shows the group's willingness and enthusiasm to work with a practical and meaningful activity using Minecraft, which allows them to specifically apply their skills in science, technology, engineering, art and mathematics.

1 Introduction

STEAM education is an educational approach that integrates the disciplines of science, technology, engineering, arts, and mathematics into the teaching and learning process. This strategy aims to foster an interdisciplinary view of knowledge and encourages students to approach problems and challenges holistically and creatively [1]. STEAM education has now gained greater recognition and acceptance in various education systems around the world. It has been documented that schools and teachers are trying to adopt this approach to prepare students for the current and future challenges of the labor market. This is due to the promotion of 21st century skills such as critical thinking, problem solving, and collaboration, as well as preparing students for future careers that require STEAM [2, 3].

There are a number of benefits that STEAM education offers, which are described below: Focus on problem solving, this is promoted through hands-on projects and activities designed to address authentic challenges, which helps students develop critical skills and find innovative and creative solutions; Inclusion of the arts and humanities, their importance in the holistic education of students has begun to be recognized, as greater interaction between STEAM disciplines has been promoted to foster a broad and balanced understanding of the world [2, 4]; Increased use of technology, STEAM education incorporates digital tools and resources in the teaching-learning process, as it enhances experimentation, simulation and collaboration, as well as facilitates access to up-to-date information and data [5, 6]; and focus on gender equality, with the development of specific programs that promote gender equality in the participation of girls and women in the various areas of STEAM [7, 8].

In addition, STEAM education is important at all levels of education [6], especially at the secondary level, where the focus is on preparing students to work in technical fields and develop innovative solutions, as it is essential that they acquire these skills in a society increasingly driven by technology and innovation. Developing critical thinking, problem solving, and creativity is

*This is an expanded version of the article published in the 2022 IEEE Mexican International Conference on Computer Science (ENC)
*Corresponding Author: Mauricio Flores-Nicolás, México, mfloresn90@icloud.com

also important because it enables them to analyze complex situations and develop new ways of approaching problems or projects. Meaningful, hands-on experiences increase their motivation and understanding of abstract concepts and scientific and mathematical principles [4, 9, 10, 11].

Beyond the benefits and importance of STEAM education, it faces several challenges, such as the high cost of materials, the digital literacy of teachers and the technological infrastructure in educational institutions [1, 6, 12, 13], as well as the lack of guidelines or standards that allow for the development and implementation of these types of activities in a satisfactory manner depending on the educational level. For example, at the secondary level [9, 10, 14], they mention that one of the main difficulties teachers face when teaching STEAM activities is time management and the pressure to follow the curriculum. That is, teachers are squeezed between authentic learning, student engagement, and multiple pathways with lesson plan deadlines, especially because a change in the way of teaching would affect the number of ways to assess.

Similarly, teaching basic science concepts at a lower level of education can be valuable in stimulating early interest in science, preparing students for higher education, and encouraging critical thinking. This provides a broader context for how science works in practice and makes the subject relevant; however, it is essential to assess the readiness and maturity of students and to balance and tailor science-based instruction to provide appropriate motivation and support. In addition, it is important to recognize that not all students will be interested in science, so it is important to draw on diverse academic aspirations in the classroom [15, 16].

It is for this reason that the purpose of this paper is to present the GMM in its two stages: first, how to design and build a STEAM activity with the elements described in [1]; and second, to execute such activity effectively and efficiently with secondary school students with a data analysis that proves a significant learning in the students on the topic. It is emphasized that STEAM education is not limited to a specific educational level, therefore the effectiveness of the GMM will be through the topic of logic gates, which is strongly linked to the basic laws of Boolean algebra (disjunction, conjunction, and negation), addition and multiplication (which will serve as a starting point), with the aim of enabling students to develop STEAM skills. The application of this content at the secondary level is of paramount importance because of its potential to fully prepare students for the technological and globalized future. At this stage, young people are eager to discover their passions and talents, and STEAM education offers an interdisciplinary approach that allows them to explore different fields of knowledge in creative and meaningful ways.

This paper discusses a quasi-experimental field study of the implementation of STEAM education in an educational setting using the GMM. The introduction provides an overview of the STEAM approach and highlights its benefits in developing key skills and competencies in students. The materials and methods section describes the design of the study, including the selection of educational institutions and educational resources. The development section describes the pedagogical strategies, practical activities, and classroom implementation. The results present empirical evidence and student testimonies on the impact of STEAM education on the learning process. The discussion analyzes the findings in light of the existing literature, highlighting the strengths and challenges of STEAM implementation in the educational context. Finally, the conclusions synthesize the findings and provide recommendations for effective and sustainable implementation of STEAM education in educational institutions, highlighting its potential for fostering critical thinking and creativity and preparing students for an increasingly technological and changing world.

2 Material and Methodology

The previous research in [1] has provided a valuable starting point, establishing the methodological foundations on which the present study is built. The GMM is designed not only to address current challenges in education, such as the need to effectively integrate technology into the teaching-learning process, but also to promote learning through the design of activities with graded and multidisciplinary approaches. This model draws on several existing pedagogical methodologies and learning strategies, such as constructionism, conceptual change, simulations, gamification, and problem-based learning, but is not limited to these alone, as content adaptation requires simplification, contextualization and careful consideration of students' skills and maturity levels. This includes ensuring that students have a solid foundation of basic concepts related to the topic, motivating them with relevant examples and applications, and providing resources and appropriate instructional support. Similarly, the topic of logic gates in [1] is aimed at undergraduate students, so adaptation is required to understand the characteristics and needs of the target audience, simplify complex information, adapt the language, focus the content, present it clearly, and contextualize it to make it relevant, meaningful, and interesting.

For this reason, this research was carried out in two parts. First, the construction of the activity, which consists in following the triple process of the GMM that starts with a hypothetical learning path (imaginary design or plan that describes the progress and learning process that the individual could follow in a certain subject or area of knowledge [1, 6]), analyzes and adapts the available technological infrastructure (the Minecraft Education software and an IntelNUC) and concentrates the information in an educational planning (it consists in designing and structuring the curricula, contents, methodologies, resources and evaluations to achieve the educational objectives effectively). And second, the evaluation stage, which involves collecting data from secondary school students by observing the learning process and the students' performance while they participate in the activity. The pre and post-test measurement technique was used because by comparing the results it is possible to determine if there has been a significant change in a variable of interest.

In addition, the following research questions will be specifically addressed to provide a deeper and more rigorous understanding of the implementation of GMM in secondary STEAM education: 1)

What is the level of students' mathematical knowledge and competence before and after participating in a STEAM activity; 2) How can STEAM education be effectively and sustainably implemented in educational institutions; and 3) How do students perceive the STEAM learning experience and what are the perceived challenges and benefits? These research questions will serve as a fundamental guide to explore the effectiveness of GMM in teaching logic gates in their deep context (Boolean algebra and electronic symbology), as seen from its representation for high school students using Minecraft.

2.1 STEAM activity design and construction

First of all, an adaptation of the hypothetical learning trajectories proposed in [1, 6] has been carried out, since in these has already been carried out a disciplinary analysis on the subject applied to high school students, it only remains to adapt the presentation for secondary school students, remaining as follows: 1) Science, logical propositions are introduced through philosophical and logical approaches; 2) Technology, a map created in Minecraft Education helps in the exploration and construction of structures with similar function to the main firebrand; 3) Engineering, works intermediate-advanced topics through truth tables, algorithms and real applications; 4) Art, a way to develop creative skills is with the topic of schematic diagrams, allowing students to apply decision-making in the design of electronic circuits; 5) Mathematics, starting with concepts of addition and multiplication and moving to their representation in the context of Boolean algebra.

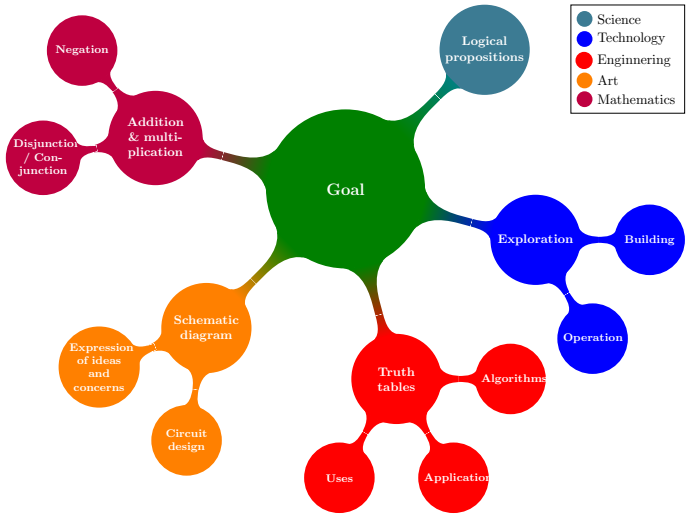


Figure 1: Hypothetical learning trajectory, classified by STEAM discipline and distributed by complexity. Own elaboration based on [1, 6].

The distribution of content is shown in figure 1, with the main theme at the center and branched into different disciplines to facilitate the implementation of STEAM activities. The objective of the second stage of the GMM is to analyze the technological infrastructure in the educational institution and align it with the pedagogical objectives, ensuring that solutions are selected that respond to specific educational needs and goals. It also optimizes their use to enhance teaching and learning and provide meaningful and effective experiences for students.

	Minecraft	Minecraft Education
Price	$6.99 – $39.99, depending on the device	$5.04 per year (institutional management required) or $12 per year (independent educator)
Does it fit in the budget?	Yes, if you have at least one compatible device	Yes, it is appropiate to carry out institutional management for the licenses
Risks	That it becomes a distraction	
Suggestion	Controlling its use	
Collaboration	Yes, in local network and online (requires additional subscriber)	Yes, in local
Trial version	14 days	26 logins
Educational elements	No	Camera, portfolio, tutors, materials (periodic table, control of actions within the map, among others)
Functioning in the institutional computer (Intel-NUC)	Best	
Additional requirements	No	Internet connection
Minimum training for its use	Included in the set	
Minimum teaching qualifications	Not available	Minecraft Learn with over 100 lessons
Examples of STEAM educational activities	Not available	Included in-game or in Minecraft Education
Can i share my own game files	Not available	Maps can be imported / exported and .pdf files of the evidence can be generated.
Are there educational resources to facilitate the development of the activity?	Not available	Yes, at: https://bit.ly/3rY76J8 (Official), How to create a map in Minecraft? & How to use Minecraft to teach high school level classes
Controllers	Joystick	Keyboard, mouse and/or joystick
Download	Xbox 360, Xbox One & Xbox series S|X, available in your online store Playstation 3, 4 y 5, available in your online store Nintendo Wii U, Switch, available in your online store Android, available in your online store	PC / macOS / Google Chromebook, Minecraft Education iPhone o iPad, available in your online store Android, available in your online store

Figure 2: Minecraft software analysis report [6].

Some of the advantages of using GMM analysis are that some of the previously created content can be reused, allowing learning to be personalized, resources and tools to be adapted to different learning styles, and leading to a more enriching and relevant education. A report of the software to be used is shown in figure 2, which provides detailed and accurate information about the operation of Minecraft and Minecraft Education IntelNUC computer equipment. Part of the information included in this report is the technical specifications, such as operating systems and compatible devices, examples of activities on similar topics, risks, suggestions, costs, and most importantly, where and how to get trained in this technology if you want to implement it in the classroom.

The third stage of the GMM provides an educational planning format that organizes the ideas, materials, goals, strategies, and pedagogical methods to achieve success during the educational process to achieve significant and lasting results. In figure 3 shows the planning format for activities in STEAM education, which aims to integrate teaching in its five disciplines with a multidisciplinary approach. In this process, teachers can design their activities and hands-on projects that foster critical thinking, problem solving, and creativity by drawing on the results of the previous two stages of the GMM. Clear learning objectives are set, and resources and digital tools are selected to indicate how and when they will intervene in the session in a way that enriches the educational experience. This format promotes collaborative work and meaningful learning,

allowing students to explore, investigate, and apply concepts and skills in real-world situations.

Project	Is it possible to add and multiply in Minecraft?
Goal	Have students build structures with the equivalent operation of logic gates or the fundamental laws of Boolean algebra. Participants can respond by using switches of buttons on the technology resource to see if the practice matches the theory.

ISTE Standards

1.3 Knowledge Builder. Students critically evaluate a variety of resources using digital tools to construct knowledge, produce creative artifacts, and develop meaningful learning experiences for themselves and others.

1.5 Computational Thinker. Students develop and employ strategies to understand and solve problems in ways that harness the power of technological methods to develop and test solutions.
- 1.5.c. Students break problems into component parts, extract key information, and develop descriptive models to understand complex systems of facilitate problem solving.

Level of technology insertion

Substitution (X)	Augmentation (X)	Modification (X)	Redefinition (X)

Pedagogic methodology	Constructionism
Learning strategy	Conceptual change; Problem-based learning; Simulation; Gamification.

How to apply it?

Science	**Technology**	**Enginnering**	**Arts**	**Mathematics**
Use and development of logical-mathematical and logical-computational thinking, basic fundamentals of circuit theory.	To favor the decision-making process with the choice of the technological tool that will allow you to solve the problem.	Know the fundamentals of electronic component manufacturing, such as design, simulation, fabrication and testing.	Designers use their creativity to create complex circuits that solve specific problems, and the circuit design itself is visually appealing process.	Theoretical verification of the operation exposed in the technological tools, used to describe the relationship between inputs and outputs of a digital circuit.

Didactic sequence

Opening (20 minutes)
1. Teacher's presentation to the class.
2. Ask students if they are familiar with Minecraft.
3. Ask students about logic gates.
4. Resolve pre-test.

Development (50 minutes)
1. The teacher presents the objectives of the activity.
2. It briefly explains about Minecraft and the basic logic gates trainer.
3. Definition and importance of logic gates.
4. Theoretical-practical definition and operation of the NOT gate in its Boolean algebra ($\neg$), mathematical ($\bar{A}$) and electronics (IC 74LS04).
5. Solve doubts and errors.
6. Theoretical-practical definition and operation of the AND gate in its Boolean algebra representation ($\wedge$), mathematical (*) and electronics (IC 74LS08).
7. Solve doubts and errors.
8. Theoretical-practical definition and operation of the OR gate in its Boolean algebra representation ($\vee$), mathematical (+) and electronics (IC 74LS32).
9. Solve doubts and errors.
10. Resolve post-test.

Closing (15 minutes)
1. Summarize what was covered in class, highlighting the importance of logic gates for everyday life.
2. Ask about their experience using Minecraft and the basic logic gates trainer in the classroom.
3. Farewell.

Materials
- Diagnostic evaluation (pre-test).
- Final evaluation (post-test).
- IntelNUC with Xbox 360 controller configured and the videogame Minecraft Education (basic logic gates map).
- Basic logic gate trainer (electrical extension, power supply, banana plug male to banana plug male cables).
- Slides.

Evaluation	Data collection through pre-test and post-test, with algebraic and mathematical operations exercises using technological tools for verification.

Figure 3: Planning format. Own elaboration based on [1, 6].

2.2 STEAM activity implementation and evaluation

The evaluation of the STEAM educational activity was carried out on a group of 43 high school sophomores in Mexico through the application of a pre-test designed to measure mathematical reasoning and a post-test designed to measure the students' level of understanding, practical skills, and perception of the learning experience. The data obtained were analyzed using a student's t-test to determine if there were statistically significant differences between the pre and post-test scores.

To design the pre-test measurement instrument, a questionnaire was developed to specifically assess students' logical-mathematical thinking prior to participating in the experience. This questionnaire includes questions on basic mathematical concepts such as addition and multiplication, problem solving skills, and logical reasoning, with the goal of establishing a baseline for the participants' level of mathematical knowledge and competence.

On the other hand, the post-test measurement tool was designed to assess the understanding and impact of the STEAM activity on the students. This questionnaire includes questions that assess the knowledge gained during the activity, the practical skills developed, and the overall perception of the STEAM learning experience. In addition, open-ended questions were included to allow students to express their opinions about the activity, its challenges, and perceived benefits, which will provide more detailed feedback on the teaching process and the level of learning achieved.

Together, these pre and post-test measurement tools provide a comprehensive view of students' progress in mathematical-logical thinking, understanding, and perception of STEAM activities. The combination of these instruments allows for analysis of the impact of the activity on the development of mathematical skills and the learning experience, which facilitates the identification of areas for improvement and informed decision making for future implementation of STEAM activities.

3 Comments on the experience

Some important observations were made while teaching logic gates:

1. Abstract concept. Logic gates are abstract concepts that may be difficult for some students to understand, especially those without a strong background in mathematics or logic. This is due to the use of symbols and logical operations that may seem unfamiliar.

2. Relationship to other disciplines. The topic caused some confusion among students who are not familiar with electronics or computer science concepts. For example, some confused logic gates with physical gates or similar concepts in logic. Also, at first it was not clear to them that the "sum" in the OR logic gate is different when there is 1 + 1 in its truth table, until it was clarified with the existence of binary numbers or simply transferring it to logic disjunction, where true $\vee$ (disjunction) true is true.

3. Supplement with examples. It was useful to provide examples of logic gates in Minecraft because, by using a different context, their translation to scientific knowledge, which is the goal of the activity, was not complex.

4. Encourage logical thinking. Teaching logic gates in secondary school is an opportunity to develop logical thinking and problem solving. Teachers can design activities and exercises that challenge students to apply their logic and reasoning skills.

Figure 5 shows the group of students excited to work together in the school lab. They have the Minecraft Education video game at their disposal, which the entire group can see on a projector. The atmosphere is dynamic and lively as the students collaborate and discuss ideas to solve STEAM activity problems. Likewise, a teacher is also present in the center, providing guidance and support to the students in their learning process. The students' expressions are one of enthusiasm and excitement as they are immersed in a meaningful, hands-on experience that allows them to apply their skills in science, technology, engineering, art, and math.

(a) Secondary school team 1/2.

(b) Secondary school team 2/2

Figure 5: Educational experience orchestrated by the GMM. Own elaboration.

4 Results and discussions

Qualitative interpretation of data from the STEAM activity on basic logic gates in high school students reveals interesting findings. After analyzing the students' responses from the pre-test, post-test, and observations during the activity, different levels of understanding (gradualness) and perception of the concepts related to logic gates could be identified.

Some students showed a solid understanding of how logic gates work and how they relate to the mathematical operations of addition and multiplication. They recognized the AND gate to a multiplication operation in mathematics and the conjunction in Boolean algebra, as well as how the OR gate relates to addition. These students were able to make clear connections between logic gates and mathematical operations, indicating a deep understanding of the concepts. On the other hand, students who remained confused in connecting logic gates to mathematics and algebra may have mixed concepts or made inaccurate comparisons between operations. These findings suggest the need to present some previous fundamentals, such as binary numbers and digital signals, during the teaching of logic gates. In addition, during the activity, it was observed that the level of participation and collaboration among the students was higher than in a traditional classroom, which means that they can work together and solve problems as a group. Overall, the qualitative data from this STEAM activity on basic logic gates provided a more holistic and detailed view of students' understanding and experience with the topic.

Now, to address the quantitative data, the Student's t-test for independent samples will be used, this is applied when there is a data set with two different samples, in this case it is so because the pre-test focuses on measuring the students' mathematical reasoning skills before the intervention, while the post-test measures the knowledge acquired after the intervention, and it is necessary to determine if there is a significant difference between the means of the two samples. The bilateral formulation of the null and alternative hypotheses in this test focuses on two possibilities, which are presented below.

$$H_0: \mu \text{ pre-test} = \mu \text{ post-test}$$

$$H_1: \mu \text{ pre-test} \neq \mu \text{ post-test}$$

The null hypothesis (H_0) states that there is no significant difference between the means, while the alternative hypothesis (H_1) states that there is a significant difference between the means. During the t-test process, a value of t is calculated and compared to a critical value to determine whether the null hypothesis is rejected. If the value of t is greater or less than the critical value, it is concluded that there is a significant difference between the means and the alternative hypothesis is accepted. On the other hand, if the value of t is not sufficiently extreme, the null hypothesis is accepted, indicating that there is no significant difference between the means.

Table 1: Values of Student's t-test assuming equal variances in the high school group. Own elaboration.

	Pre-test	**Post-test**
Mean	6.710963455	2.373285629
Variance	2.092797252	2.069731766
Observations	43	43
Pooled Variance		2.081264509
Hypothesized Mean Diff.		0
Pooled Variance		2.081264509
df		84
t Stat		13.94160851
P($T <= t$) one-tail		$7.22784e^{-24}$
t Critical one-tail		1.663196679
P($T <= t$) two-tail		$1.44557e^{-23}$
t Critical two-tail		1.988609667

The table 1 shows the results of the Student's t-test, describing the values as follows: Mean, in the pre-test, represents the average value of their logical-mathematical skills scores before the intervention, while the post-test represents the results obtained by the participants after the STEAM activity; Variance, indicates how far the individual scores deviate from the mean; Sample, there were 43 students who participated in the activity; Common variance, represents the average variance of both groups combined; Mean difference hypothesis, indicates that the difference between the pre-test and the post-test is equal to zero, meaning that there is no significant difference between the means of the two groups; Degrees

of freedom, represents the number of data that are free to vary after taking into account the restrictions imposed by the study design, in this case it is obtained by subtracting 1 from the total sum of both samples; t-statistic, with a value of 13.94, indicating that the sample means are far apart, indicating a significant difference between the pre and post-test means; P two-tailed value, is $1.44557e^{-23}$, close to zero, indicating a significant difference between the pre and post-test means; the critical two-tailed value is 1.98, which serves as a reference point to compare the t statistic, in this case the statistic value is much greater than the critical value, indicating that the null hypothesis should be rejected, and it is concluded that there is a significant difference between the pre and post-test means.

The results indicate that there is a significantly high difference between the means and there is sufficient statistical evidence to reject the null hypothesis. This indicates that the educational intervention provided by the GMM had a significant impact on the performance and skills of the participants:

1. Specific learning. During the educational intervention, students acquired specific knowledge about logic gates and their mathematical, algebraic, and electronic representation that could improve their future performance.

2. Practical Application. The STEAM activity gave students the opportunity to apply their logical-mathematical reasoning skills to practical contexts and real-world problems, which may have improved their understanding and performance on the post-test.

3. Reinforcement and practice. The educational intervention included examples and exercises that allowed students to reinforce and consolidate their skills and knowledge on the topic, which was reflected in the post-test.

It is important to note that this difference in means between the pre and post-test does not necessarily indicate that the educational intervention was ineffective. On the contrary, if students demonstrate an increase in specific knowledge and understanding related to the STEAM activity on the post-test, this could be considered a positive and desirable outcome in terms of learning. Furthermore, the results may be useful for adapting pedagogical approaches, providing additional support to students who are struggling, and promoting meaningful and effective learning of logic gates and digital logic in secondary education.

On the other hand, in the existing bibliography with the same purpose, such as [17, 18], they do not provide specific information on how these teaching-learning sequences can be adapted to different educational levels. This is done implicitly, inferring that the STEAM strategy that has different knowledge domains automatically has the capacity to adapt the content to different educational levels, which is not correct. It is important to keep in mind that the adaptation of a specific topic such as logic gates or any other topic should not be presented in its complex representation to students who have not developed a certain level of maturity, giving importance to the gradual characteristic of the GMM.

5 Conclusions

It should be noted that this study was conducted in a specific context and with a limited sample of participants. Therefore, the findings may not be generalizable to other populations or educational contexts. Although positive and significant effects of STEAM activities on participants' learning and skill development were found, it is possible that other factors not measured in this study may have contributed to these results, such as student motivation, instructional quality, technology acceptance, and content presentation. Therefore, further research using the GMM is needed to confirm the results of this study and to further explore the factors that contribute to the successful implementation of STEAM education in educational institutions.

Based on the data obtained from the t-Student test, we can affirm that the implemented STEAM activity had a positive and significant effect on the learning and skill development of the participants. The reduction in the mean of the post-test compared to the pre-test suggests that the students acquired new knowledge and skills relevant to the activity, which confirms the effectiveness of the orchestration of the STEAM activity through the GMM and its improvement of the students' performance on the assessed topic.

In addition, it was possible to determine the level of knowledge of the students through the pre-test and post-test to measure the logical-mathematical thinking, level of understanding, practical skills, and perceptions of the students after the STEAM activity. The results of these questionnaires were analyzed to determine if there were statistically significant differences between the scores. Regarding the sustainability of STEAM education in educational institutions, for the time being, further research with the GMM is suggested to further explore the factors that contribute to the success of STEAM education implementation, at least in this work, most of the challenges that have to do with the acquisition of materials can be solved with careful analysis and planning prior to an investment in technology. Finally, students expressed their opinions about the STEAM activity and perceived benefits through some open-ended questions, highlighting a sense of active involvement and relevance in their learning, as these activities are usually practical and meaningful, just as they appreciated the opportunity to apply academic concepts in a real context, which allows them to see the usefulness and concrete application of what they are learning. These attitudes, along with collaboration and team problem solving, also fostered a sense of community and an appreciation of diverse perspectives; overall, the activities orchestrated by the GMM tend to generate a lasting enthusiasm for learning and exploration in these areas, which can have a positive impact on students' future academic and career choices.

Future work Given that the STEAM activity had a positive and significant effect on learning, we make the following recommendations for future work: 1) Expand the sample, increasing the size will allow us to obtain greater representativeness and robustness in the results, providing greater confidence in generalizing the findings to a larger population of students; 2) Conduct long-term follow-up, this will allow us to evaluate whether the positive effects of

the STEAM intervention are maintained over time or if there are long-term learning effects; 3) Conduct more interventions using the GMM; analyzing and comparing different types of STEAM interventions using the GMM would help determine if they are effective with the five disciplines in terms of learning and specific skill development, as well as help identify more efficient and effective pedagogical approaches.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment We are very grateful to the Escuela Secundaria Técnica 34 "Luis V. Massieu" (09DST0034Q), for allowing us to work with the high school group.

References

[1] M. Flores-Nicolás, M. Martínez, "The Graded Multidisciplinary Model: Fostering instructional design for activity development in STEM/STEAM education," in 2022 IEEE Mexican International Conference on Computer Science (ENC), 1–7, 2022, doi:10.1109/ENC56672.2022.9882917.

[2] D. Aguilera, J. Ortiz-Revilla, "STEM vs. STEAM Education and Student Creativity: A Systematic Literature Review," Education Sciences, **11**(7), 2021, doi:10.3390/educsci11070331.

[3] M. Conde, Rodríguez, C. Fernández-Llamas, J. Gonçalves, J. Lima, F. J. García-Peñalvo, "RoboSTEAM Project Systematic Mapping: Challenge Based Learning and Robotics," in 2020 IEEE Global Engineering Education Conference (EDUCON), 214–221, 2020, doi:10.1109/EDUCON45650.2020.9125103.

[4] G. Ozkan, U. Umdu Topsakal, "Exploring the effectiveness of STEAM design processes on middle school students' creativity," International Journal of Technology and Design Education, **31**, 95–116, 2021.

[5] A. Leavy, L. Dick, M. Meletiou-Mavrotheris, E. Paparistodemou, E. Stylianou, "The prevalence and use of emerging technologies in STEAM education: A systematic review of the literature," Journal of Computer Assisted Learning, **39**(4), 1061–1082, 2023, doi:https://doi.org/10.1111/jcal.12806.

[6] M. Flores-Nicolás, M. Martínez-Reyes, "A Theoretical-Practical Case Study on the Graded Multidisciplinary Model: Training of High School Students Through STEAM Education," in Streamlining Organizational Processes Through AI, IoT, Blockchain, and Virtual Environments, 194–218, IGI Global, 2023, doi:10.4018/978-1-6684-8639-9.ch010.

[7] L. Ma, H. Luo, X. Liao, J. Li, "Impact of Gender on STEAM Education in Elementary School: From Individuals to Group Compositions," Behavioral Sciences, **12**(9), 2022, doi:10.3390/bs12090308.

[8] S. Aydın Gürler, O. Kaplan, "Attitudes Towards STEAM, Critical Thinking Disposition and Decision-Making Skills: Mediation and Gender Moderation," Journal of Contemporary Educational Research, **10**(1), 210–223, 2023.

[9] N. Arís, L. Orcos, "Educational Robotics in the Stage of Secondary Education: Empirical Study on Motivation and STEM Skills," Education Sciences, **9**(2), 2019, doi:10.3390/educsci9020073.

[10] Y. Rahmawati, A. Ridwan, T. Hadinugrahaningsih, Soeprijanto, "Developing critical and creative thinking skills through STEAM integration in chemistry learning," Journal of Physics: Conference Series, **1156**(1), 012033, 2019, doi:10.1088/1742-6596/1156/1/012033.

[11] E. Yilmaz Ince, M. Koc, "The consequences of robotics programming education on computational thinking skills: An intervention of the Young Engineer's Workshop (YEW)," Computer Applications in Engineering Education, **29**(1), 191–208, 2021, doi:10.1002/cae.22321.

[12] D. Herro, C. Quigley, "Exploring teachers' perceptions of STEAM teaching through professional development: implications for teacher educators," Professional Development in Education, **43**(3), 416–438, 2017, doi:10.1080/19415257.2016.1205507.

[13] S. Belbase, B. R. Mainali, W. Kasemsukpipat, H. Tairab, M. Gochoo, A. Jarrah, "At the dawn of science, technology, engineering, arts, and mathematics (STEAM) education: prospects, priorities, processes, and problems," International Journal of Mathematical Education in Science and Technology, **53**(11), 2919–2955, 2022, doi:10.1080/0020739X.2021.1922943.

[14] C. F. Quigley, D. Herro, A. Baker, Moving Toward Transdisciplinary Instruction: A Longitudinal Examination of STEAM Teaching Practices, 143–164, Springer International Publishing, Cham, 2019, doi:10.1007/978-3-030-04003-1_8.

[15] Z. Yang, "The challenges of personalized learning and their solutions," Chinese Science Bulletin, **64**(5–6), 493–498, 2018.

[16] M. Á. Conde, F. J. Rodríguez-Sedano, C. Fernández-Llamas, J. Gonçalves, J. Lima, F. J. García-Peñalvo, "Fostering STEAM through challenge-based learning, robotics, and physical devices: A systematic mapping literature review," Computer Applications in Engineering Education, **29**(1), 46–65, 2021.

[17] I. M. Greca, J. Ortiz-Revilla, I. Arriassecq, "Diseño y evaluación de una secuencia de enseñanza-aprendizaje STEAM para Educación Primaria," 2021, doi:10.25267/Rev_Eureka_ensen_divulg_cienc.2021.v18.i1.1802.

[18] M. Queiruga-Dios, E. López-Iñesta, M. Diez-Ojeda, M. Sáiz-Manzanares, J. Vázquez-Dorrío, "Implementation of a STEAM project in compulsory secondary education that creates connections with the environment (Implementación de un proyecto STEAM en Educación Secundaria generando conexiones con el entorno)," Journal for the Study of Education and Development, **44**(4), 871–908, 2021, doi:10.1080/02103702.2021.1925475.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 8, No. 5, 62-69 (2023)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Science & Technology

ASTES Journal
ISSN: 2415-6698

Modeling Control Agents in Social Media Networks using Reinforcement Learning

Mohamed Nayef Zareer*, Rastko Selmic

Department of Electrical and Computer Engineering, Concordia University, Montreal, H3G1M8, Canada

ARTICLE INFO

Article history:
Received: 01 August, 2023
Accepted: 29 September, 2023
Online: 30 October, 2023

Keywords:
Markov Decision Process
Multi-agent Systems
Q-Learning
Opinion Dynamics
Social Network Analysis

ABSTRACT

Designing efficient control strategies for opinion dynamics is a challenging task. Understanding how individuals change their opinions in social networks is essential to countering malicious actors and fake news and mitigating their effect on the network. In many applications such as marketing design, product launches, etc., corporations often post curated news or feeds on social media to steer the users' opinions in a desired way. We call such scenarios opinion shaping or opinion control whereby a few selected users, called control users, post opinionated messages to drive the others' opinions to reach a given state. In this paper, we are interested in the control of opinion dynamics in social media using a combination of multi-agent systems and Q-learning. The social media environment is modeled with flexible multi-agent opinion dynamics that can capture the interaction between individuals in social media networks using a two-state updating mechanism. The environment is formulated as a partially observable multi-agent Markov decision process. We propose using intelligent reinforcement learning (RL) agents to control and shape the social network's opinions. We present the social media network as an environment with different kinds of individuals and connections and the influencing agent as an RL agent to influence the network.

1 Introduction

Social media has received widespread attention due to its rise as one of the most essential tools for societal interaction and communication. Social media platforms have significantly contributed to the rapid dispersion of news and information and the facilitation of communication between different groups worldwide. Nevertheless, in the past few years, there has been a notable surge in the spread of misinformation, a trend propelled and magnified by the influence of social media platforms such as Facebook and Twitter. Due to this, it has become increasingly vital to understand opinion dynamics in social networks to minimize the detrimental effects of malicious agents, fake news, and other polarizing factors. This work builds on the work presented in [1], where we discussed how updating dynamics on social media and competition for influence affect the overall opinion and how intelligent agents decrease polarization and disagreement in social media networks.

Extensive research has been conducted to examine and understand this issue of misinformation in social media networks [2]–[4]. The prevailing trends to combat this issue are automated tools for detecting fake news and misleading posts [5]. However, this presents multiple challenges, such as defining what is fake and true. For example, simply defining what is true and false can generate a lot of discussion and controversy. This is compounded by the availability of artificial intelligence tools that can mimic a human's voice or generate fake video recordings. Additionally, misinformation in social media can result in polarization, economic impacts, time and resource wastage, and cybersecurity concerns.

Furthermore, understanding the effect of RL on influencing opinions in a social media environment is vital for several avenues such as public opinion management [6], policy implementation [7], commercial interests [8], platform design [9], and Ethical Considerations [10].

This paper proposes an RL-based method for influencing and controlling opinions in a social media environment. Unlike what is presented in the literature, where the model or algorithm is static and does not react to changes in the social network, our method proposes that each control agent can react and take action using Deep Q-Networks (DQN) to influence and persuade the other agents in the network to adopt a similar opinion. This paper expands the work in [1] with the following contributions:

1. The control agents can change their expressed opinions to influence the social media network.

2. The environment is a two-state expressed and private opinion dynamics model with asynchronous and synchronous updat-

*Corresponding Author: Mohamed Zareer, Canada, 514-3782524 & mngasem1990@gmail.com

ing dynamics that emulate the interactions of individuals in social media networks.

3. The agents work to influence the opinions of others in the network to a set goal.

4. Simulation results that demonstrate our approach for multiple social media networks.

The rest of the paper is organized as follows. Related works and preliminaries are provided in Section 2. We introduce the Markov decision process (MDP), partially hidden MDPS, and RL in Section 3. In Section 4, we formulate the problem and present the paper's main results. The experimental results are given in Section 5. Concluding remarks are given in Section 6.

2 Related Work

Social media networks have become an integral part of our society. This resulted in widespread attention to opinion dynamics and social network analysis. Opinion dynamics studies how opinions, beliefs, and attitudes form, evolve, and interact within social networks and communities. Opinion dynamics encompasses a variety of disciplines, such as sociology, political science, complex systems analysis, psychology, and multi-agent systems. One of the essential assumptions in opinion dynamics is that opinions in social networks are mainly influenced by others.

Many opinion dynamics have been proposed to study the evolution of opinions in social networks [11]–[13]. One of the most popular opinion dynamics models is agent-based models [14]. In agent-based models, individuals are depicted as agents, and their opinions on a specific topic are captured as evolving real values over time. The underlying communication network in agent-based models is represented by a graph, where a node represents an agent, and an edge represents communication between two individuals.

One of the earliest agent-based models is the French-DeGroot model, commonly referred to as the DeGroot model [15]. The model assumes that the opinion of each individual evolves as a result of integrating the opinions of their connected neighbors with agents' own opinions using weighted averaging (modeled using a differential or a difference equation). The DeGroot model was experimentally validated in [16, 17]. An extension of the DeGroot model is the Friedkin-Johnsen model, which simulates strong diversity due to stubborn agents by introducing a variable that measures the agent's susceptibility to social influence [18]. This model has been experimentally verified for small and medium-sized networks [19]–[21]. A model that encodes the effects of social pressure on the agents in the network, termed the expressed and private opinion (EPO) dynamics model, was introduced in [22]. The EPO model is based on Asch's conformity experiments and Prentice and Miller's field experiments on pluralistic ignorance [23, 24]. An extension of the EPO model is the asynchronous and synchronous expressed and private opinion dynamics model (EPOAS) [25]. This model introduces different updating dynamics for the agent's expressed and private opinions, which emulates the interaction of individuals in social media networks.

These models tend to reach a consensus if all the agents in the network agree on one opinion on a specific topic. However, it can be observed that in most social networks, the presence of stubborn (who insist on their own opinion) or controlling agents has a significant impact on the opinions of the other agents in the network [26]. This trend is evident in social media marketing, economics, and political campaigns [27]. Moreover, numerous political and economic entities utilize data mining techniques and social science principles to strategically engage specific individuals within social networks, aiming to enhance profit margins [28].

There have been many attempts to study and simulate opinion control in social networks. The control of the DeGroot model under the influence of a leader was investigated in [29]–[31]. In [32], the authors study the optimal placement of control agents in a social network to influence other individuals to reach a consensus where the control agents have a fixed common state. Another control approach using noise to affect the opinions of individuals in the network was investigated in [33]. A control strategy based on the degree of connection each agent has shown that it is possible to drive the overall opinion toward a desired state even if we control only a suitable fraction of the nodes was presented in [34].

In [35], the control of public opinion using social bots was investigated. In this approach, an agent's opinion is modeled as a static value, based on the approach described by Sohn and Geidner [36]. The study demonstrates that, depending on the density and position within the network, a mere $2\% - 4\%$ of bots are sufficient to influence all the opinions in the network.

3 Markov Decision Process

An MDP is a mathematical framework designed to model decision-making where sequential actions are involved, and the outcomes of each action are partially random and partially under the control of the decision-making agent [37]. An MDP was developed to model decision-making problems where the outputs are probabilistic and are affected by the agent's actions. An MDP is modeled by a tuple (S, A, P_a, R_a) where:

1. States (S) encapsulate all potential scenarios the agent could encounter at any given time step.

2. Actions (A) embody all the options or decisions accessible to the agent in any given state. Agents select an action to transition from one state to another.

3. Transition probabilities (P_a) specify the likelihood of transitioning to a new state. These probabilities represent the dynamics of the environment and determine how the agent's actions influence the next state.

4. Rewards (R_a) quantify the desirability of taking a specific action in the current state. The agent's goal is to maximize the rewards it receives to reach its desired goal efficiently.

The goal of an MDP is to have the agent learn a good policy for decision-making. A policy $\pi(S)$ can be defined as a strategy that maps states to actions, indicating the action the agent should choose in each state to maximize the agent's reward and help the agent reach their goal.

3.1 Partially Observable MDP

Partially Observable Markov Decision Processes (POMDPs), Figure 1, provide a robust framework for effectively modeling and resolving decision-making challenges in contexts marked by uncertainty and partial observability. A POMDP represents a decision-making scenario for an agent. In this scenario, it is assumed that the system's behavior is governed by a MDP. However, the agent is unable to perceive the inherent state of the system directly.

Decision-making in real-world environments often involves inherent uncertainty and partial observability, where agents lack complete information about the underlying states and face uncertain outcomes from their actions. The framework of POMDPs is sufficiently versatile to represent a wide array of sequential decision-making scenarios encountered in the real world.

In POMDPs, an agent makes decisions based on a belief state (a probability distribution over all possible states) rather than the actual state. The agent's belief state is updated based on the actions it takes and the observations it receives. The agent's goal is to choose actions over time to maximize its expected cumulative reward.

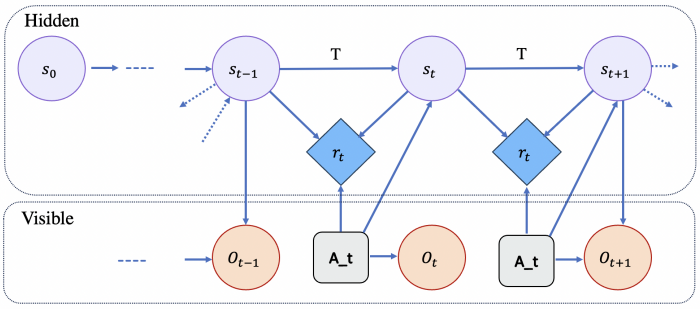


Figure 1: Partially Observable Markov Decision Processes (POMDPs).

3.2 Reinforcement Learning (RL)

One of the most powerful and widely used machine learning tools in data science is RL. The ability of RL algorithms to learn about the environment and generate a suitable policy, which can be improved through trial and error, has made them a popular choice for tasks such as game playing, autonomous driving, robotics, and resource management, among others. One of the most essential RL algorithms is the Q-learning algorithm. This can be attributed to the algorithm's versatility, simplicity, and robustness.

Q-learning is a model-free RL algorithm, meaning it does not require knowledge of the environment's dynamics to generate an efficient policy for solving the problem. This characteristic makes the algorithm particularly suitable for problems where a model of the environment is challenging to obtain or the environment is non-deterministic.

The essence of Q-learning is learning the action-value function (Q-function). The Q-function evaluates the value of taking a specific action in a given state

$$Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma\, max_a Q(s',a') - Q(s,a)], \qquad (1)$$

where s represents the agent's present state, s' represents the agent's next state, and a is the action taken by the agent in the current time step. The Q-value of the action a for the current state s is denoted by $Q(s,a)$, which is the state-action value. The reward the agent receives at each time step is represented by r, and γ is the discount factor that reduces the value of future rewards over time.

The value $Q(s,a)$, for the current state s, is updated at every time step based on a blend of the existing value and the equation that identifies the optimal action in the current state. Initially, the Q-value table is randomly populated for each state and potential action. The Q-learning process continues by updating the Q-value for each state using (1). The policy is then updated using the highest Q-values for each state-action pair. After the agent performs an action a in state s, it transitions to the next state s'. This procedure is repeated multiple times until the overall Q-values reach a point of convergence [38]. The algorithm for Q-learning is described below.

Algorithm 1: Q-learning for estimating $\pi \approx \pi^*$

Algorithm parameters: step size $l_r \in (0,1]$ ε, γ;
Initialization: $Q(s,a)$ *for all* $s \in S, a \in A$
for *episode* $\leftarrow$ *episodes* **do**
  initialize s
  while *not done* **do**
    Choose a from s using policy derived from Q (e.g., $\varepsilon - greedy$) ;
    Take action a and observe the reward r, and next state s';
    $Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma\, max_a Q(s',a') - Q(s,a)]$,
    $s \leftarrow s'$
  end
end

The learning rate l_r needs to be fine-tuned over time to solidify the learned policy. Discount factor γ is a numerical value between 0 and 1 that determines the importance of future rewards compared to immediate ones. If γ is close to 0, then the agent will prioritize immediate rewards and largely ignore future rewards. However, if γ is close to 1, then the agent will consider future rewards almost as important as immediate ones.

4 Problem formulation

Consider a group of n agents interacting in a social media environment, where agents form and evolve their opinions at each time step by interacting and exchanging information with other connected agents.

An independent agent or a group of agents is arbitrarily selected to control and influence the network to a desired outcome. The agents can have the same goal and work together to achieve that goal or they can have different goals and compete for influence in the network.

4.1 Modeling the Agents

We model n agents as nodes interacting in a multi-agent opinion dynamics environment. The agents are interconnected via an underlying graph $\mathcal{G}[W]$, which maps out the relationships between these agents. The underlying graph $\mathcal{G}[W]$ is a directed graph that encodes the flow of information from one agent to another. Each individual

i in the network has two states: one state represents the individual's expressed opinion $\hat{y}_i(t)$, and the second state is the individual's private opinion $y_i(t)$. Each agent within the network starts with an initial private opinion, denoted as $y_i(0)$, and an initially expressed opinion, represented as $\hat{y}_i(0)$ (for the remainder of this paper, the terms 'state' and 'opinion' will be used interchangeably). The initial private and expressed opinions of each agent can either coincide ($y_i(0) = \hat{y}_i(0)$) or differ ($y_i(0) \neq \hat{y}_i(0)$). This mirrors real-world scenarios where individuals may choose to voice their genuine opinions or, due to external influences, conceal their private views by expressing differing opinions. At every time step t, each individual i in the network has two choices encoded by the variable α. The individual can express his/her opinion ($\alpha_i = 1$) or hide the opinion ($\alpha_i = 0$). If an individual decides to hide his/her opinion, then the last expressed opinion made by that individual will appear as his/her current expressed opinion to his/her neighbors. Note that for control agents $\alpha(t) = 1 \forall t$. At every time step, all individuals in the network, except control agents, update their private opinions using the following dynamics

$$y_i(t+1) = \lambda_i w_{ii} y_i(t) + \lambda_i \sum_{j \neq i} w_{ij} \hat{y}_j(k) + (1-\lambda_i) y_i(0), \tag{2}$$

where $\lambda_i \in [0,1]$ represents the coefficient of susceptibility of agent i, $w_{ij} \geq 0$ represent the weights assigned by agent i to agent j, $w_{ii} \geq 0$ represents the weight assigned by agent i to his/her opinion.

The expressed opinion is updated asynchronously, utilizing the information in the vector $\alpha(t) \in [0, N]$. This vector encodes the decisions made by each agent at the time step t. If agent i opts to express their opinion, then the update of their expressed opinion adheres to the subsequent dynamics

$$\hat{y}_i(k+1) = \begin{cases} \phi_i y_i(t) + (1-\phi_i)\hat{y}_{avg}(t) & \alpha_i = 1 \\ \hat{y}_i(t) & \alpha_i = 0 \end{cases}, \tag{3}$$

where ϕ represents the resilience of the agent to social pressure and $\hat{y}_{avg}(t)$ represents social pressure or the prevailing opinion (*public opinion*) throughout the network. The average opinion, as observed by agent i, is given by the following dynamic

$$\hat{y}_{i,avg}(t) = \sum_{j \neq i} m_{ij} \hat{y}_j(t), \tag{4}$$

where $m_{ij} \geq 0$ satisfy $\sum_{j=1}^{n} m_{ij} = 1$. In many instances, w_{ij} and m_{ij} are not identical. This occurs as an individual's viewpoint can be shaped and influenced by a certain group of individuals, all while concurrently feeling the urge to align with the expectations of the overall network.

Note that control agents do not follow the same updating dynamics and do not have a private opinion. The expressed opinions of the control agents are controlled by the Q-learning dynamics based on the optimal policy learned by the agent.

4.2 Communication Topology

The communication topology for n individuals (agents) can be represented as a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}[W], W)$, where $\mathcal{V} = \{v_1, \ldots, v_n\}$ is the set of nodes (which represent agents or individuals in the network), and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of ordered edges, $\mathcal{E} = \{e_1, \ldots, e_n\}$. Each edge in the network is denoted as $e_{ij} = (v_i, v_j) \in \mathcal{E}$. With N_i we denote the neighbor set for the agent i (set of agents connected to the agent i).

The relative influence matrix of the network is modeled by W. The influence matrix W encodes how much trust or weight each agent has in the opinions of his/her connected neighbors. It is assumed that the influence matrix is static and connected. However, the communication network changes at each time step based on the updating dynamics of the agents in the network.

Each agent can only observe the expressed opinions of other connected agents. This means that the Q-table depends on the number of individuals connected to the control agents rather than the size of the network. The agent can have an idea of the overall opinion of the network based on the M weight matrix in (4). However, in this work, we assume that $M = W$. The observation space depends on the connections (neighbors) of the control agents. For example, if the control agent has three connections, the agent will have an observation space of three expressed opinions, as shown in Figure 2.

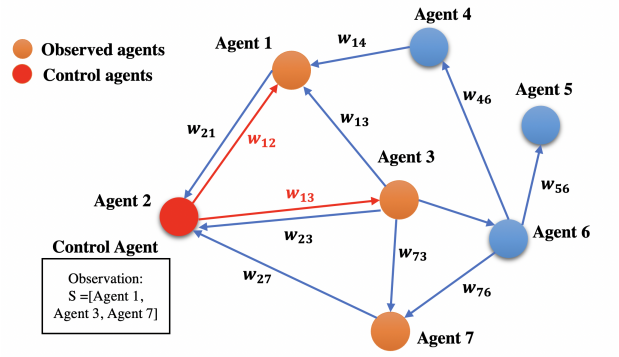


Figure 2: Observation of a control agent.

Some key information about the agent can be summed up in the following table.

Table 1: Agent information.

Action Space	Discrete(5)
Observation Shape	(number of neighbors,)
Observation High	1.0
Observation Low	0.0

4.3 The Environment Dynamics

The environment is based on the asynchronous and synchronous expressed and private opinion dynamics model. Consider a connected network of n agents. Let $y(t) = [y_1, y_2, \ldots, y_n]^T \in \mathbb{R}^n$, and $\hat{y}(t) = [\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n]^T \in \mathbb{R}^n$ be a vector representing the private, and expressed opinions (states) of all the agents in the network, respectively. Next, we define $Y(t) = [y(t)^T, \hat{y}(t)^T]^T$, as the vector of all opinions (states) among the individuals at time step t. The vector of initial private opinions (prejudices) is defined as $u = [u_1, u_2, \ldots, u_n]^T \in \mathbb{R}^n$. We define M as a matrix that encodes the weights for calculating social pressure. This matrix can be similar to the influence matrix $M = W$ (which is followed in this work)

or it can be different depending on the source of the social pressure (prevailing opinion in the network).

We use $W = [w_{ij}]$ as the influence matrix of the network which is a stochastic matrix. We define $\tilde{W} = \text{diag}(w_{ii})$ as the diagonal matrix containing the diagonal values of W which represents the self-confidence of each agent in the network, and $\hat{W}$ as the exact matrix as W with 0 in its diagonal ($\hat{w}_{ij} = w_{ij}$ for all $j \neq i$). The influence matrix can be rewritten as $W = \tilde{W} + \hat{W}$.

Additionally, we define $\lambda = [\lambda_1, \lambda_2, \ldots, \lambda_n]$ as a vector that encodes the agent's susceptibility to social influence, while $\phi = [\phi_1, \phi_2, \ldots, \phi_n]$ is a vector that encodes the agent's resilience to social pressure. The matrix $\Lambda = \text{diag}(\lambda)$ is a susceptibility matrix that encodes the susceptibility of the agents in its diagonal, and the matrix $\Phi = \text{diag}(\phi)$ is a resilience matrix encoding the resilience of the agents in its diagonal.

We designate $\alpha(t) = [\alpha_1(t), \alpha_2(t), \ldots, \alpha_n(t)]$ as a choice vector of zeros and ones that encode the actions agents have chosen by the agents at time t, where a zero value indicates that a particular agent has chosen not to express their opinion at the time step t, and a value of one signifies the agent's choice to share their opinion with their neighbors. Let $E = \text{diag}(\alpha(t))$ be a zero matrix with the choices of the agents encoded in its diagonal, and $T = I_n - E$ be an identity matrix where the value of the activated agents are set to zero.

To demonstrate the dynamics of the system we define $P_{\alpha(t)}$ and C as follows

$$P_{\alpha(t)} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} = \begin{bmatrix} \Lambda\tilde{W} & \Lambda\hat{W} \\ E\Phi & T + E(I_n - \Phi)M \end{bmatrix} \in \mathbb{R}^{2n \times 2n}, \quad (5)$$

$$C = \begin{bmatrix} (I_n - \Lambda) \\ \mathbf{0}_{nxn} \end{bmatrix} \in \mathbb{R}^{2n \times n}, \quad (6)$$

where $\mathbf{0}_{nxn}$ is a square zero matrix, and I_n is the identity matrix. The dynamics of the environment are updated using the following equation

$$\begin{bmatrix} y(t+1) \\ \hat{y}(t+1) \end{bmatrix} = P_{\alpha(t)} \begin{bmatrix} y(t) \\ \hat{y}(t) \end{bmatrix} + \begin{bmatrix} (I_n - \Lambda) \\ \mathbf{0}_{nxn} \end{bmatrix} u. \quad (7)$$

This equation can be rewritten as follows

$$Y(t+1) = P_{\alpha(t)} Y(t) + Cu. \quad (8)$$

The matrix $P_{\alpha(t)}$ changes with each time step t based on the changes that occur in the choice vector $\alpha(t)$. Each control agent in the environment has the following available actions.

Table 2: Action space

Action	Opinion value	Meaning
0	0.1	Strong disagreement
1	0.3	Slight disagreement
2	0.5	Neutral
3	0.7	Slight agreement
5	0.9	Strong agreement

The following algorithm is used to update the dynamics of the system at each time step.

Algorithm 2: Step function in system dynamics

```
time step = time step +1
Agents_Activation = random(size=number of agents)
E = diag(Agents_Activation)

for i in range(length(control agents) do
    E[control agents[i]] = 1
    Agents_Current_Exp_Opinions[control agents[i]] =
      action values[actions[i]]

    T = I_n − E
    P11 = Λ W_wave
    P12 = Λ W_hat
    P21 = E Φ
    P22 = T + E(I_n − Φ)M
    c = (I_n − Λ) initial_private_states)

    Ypvt = P11 Agents_Current_Pvt_Opinions + P12
      Agents_Current_Exp_Opinions + c

    Yexp = P21 Agents_Current_Pvt_Opinions + P22
      Agents_Current_Exp_Opinions

    Agents_Current_Pvt_Opinions = Ypvt
    Agents_Current_Exp_Opinions = Yexp
    controlling agent observation = []

    for agent in control agents do
        a = []
        for i in control agent observations[agent] do
            a.append(Agents_Current_Exp_Opinions[i])
        end
    end

    controlling agent observation.append(a)
    check reward(Agents_Current_Exp_Opinions)
    check terminal(Agents_Current_Exp_Opinions)

    return controlling agent observation, reward, terminal
end
```

4.4 Reward Function

Every control agent in the network aims to influence the individuals' opinions to a specific value (opinion). The opinions of the agents in the network range from 0, which represents strong disagreement, to 1, which represents strong agreement. Let ref_i be the opinion goal or reference of the control agent i.

We define the disagreement $d(i, j)$ between two individuals i, j as the squared difference between their opinions at equilibrium: $d(i, j) = w_{ij}(y_i^* - y_j^*)^2$, and the total disagreement is defined as $D_G = \sum_{(i,j)\in\mathcal{V}} d(i, j)$, [39].

The overall goal of the control agents is to influence the network to a predetermined opinion. To encourage the agents to reach their goal within the desired number of steps, we define $O_i(t)$ as the vector containing the expressed opinions of the agents connected to agent i at time step t. The reward function for each control agent is given

by

$$r_i(t) = -\sum_{j \in N_i} (ref_i - O_{ij}(t))^2. \quad (9)$$

The control agent is given an increasingly negative reward for each connected agent that does conform to the goal opinion to encourage it to influence as many individuals as possible.

5 Experimental Results

We tested the effectiveness of the RL agent in influencing and controlling the opinions of individuals in social media networks. A graph was randomly generated using the Erdos-Renyi random graphs model, preferential attachment model, or the stochastic block model. The weights agents assigned to themselves and their neighbors were generated randomly.

The susceptibility and resilience values were randomly generated to ensure the agent learned under different conditions. Then one or more control agents were randomly selected to influence the rest of the network to one specific opinion. At each time step, the agent observed the expressed state of their connected neighbors and took action to influence their opinion. The reward was calculated at each time step to motivate the agent to complete the task in a timely manner. Note that the agents start with no knowledge of the underlying network. In addition, the agents do not have any prior knowledge about other agents in the network.

The agents started with a very high exploration factor ε that decayed gradually until the value reached 0.01, which means that the agents reverted to an almost purely greedy algorithm when choosing their action. The factor ε is decayed by a factor of 0.0001 every episode. The discount factor for the agents was selected as 0.99, which means that the agents developed long-term planning to achieve their goal rather than focus on immediate rewards.

The agent was trained for 500, 000 episodes, and the reward was logged for every 1, 000-th episode. Additionally, for every 1, 000-th episode, we log the episode's length to measure how fast the agent completes the assigned task. Figure 3 shows the evolution of an agent's learning process, starting with a randomly defined knowledge base and no experience over 250, 000 iterations. The rewards (orange line) show the 500 rolling average of the rewards for a more clear illustration.

Figure 3 shows an agent's learning in a random network. The agent starts to receive high rewards after 150, 000 iterations. However, there are still episodes where the control agents struggle to efficiently control the network due to the difference in agent personality (susceptibility and resilience) or opinions.

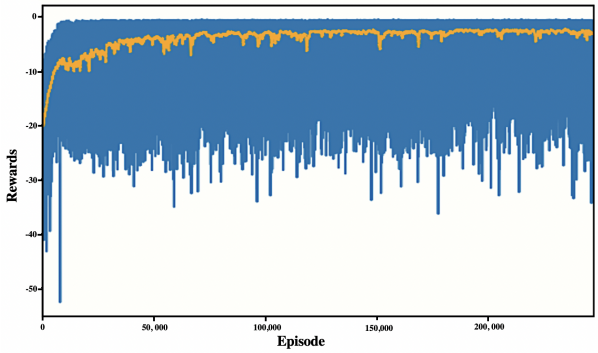


Figure 3: Training reward of a control agent over 250,000 training episodes with data collected every 1,000 episodes.

Figure 4 shows the reward over 250, 000 training episodes with the reward logged every 1, 000-th training episode and a 500 rolling average. The agent starts with a large negative reward when they are exploring ways to influence the network, then the agent starts getting consistent rewards close to 0, which is the optimal reward, after 150, 000 training iterations. The figure shows that the agent is able to influence the system where the reward of the network reaches 0 meaning that the opinions of the agents in the network converge to the opinion desired by the control agent. However, increasing the training time after 250, 000 iterations does not have much effect on the reward received by the agent.

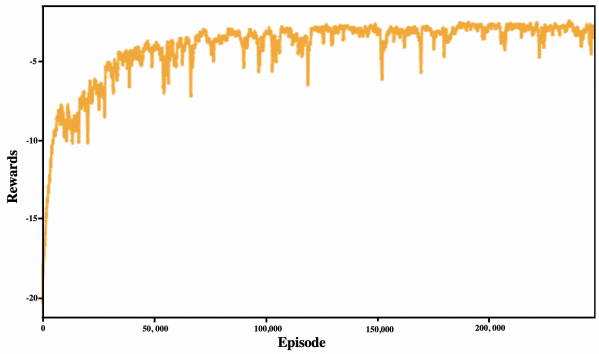


Figure 4: Training reward of a control agent over 250,000 training episodes (with 500 rolling average).

Figure 5 shows the length of every episode and how long the agent takes to reach an efficient policy. It can be seen that with more training, the agent can influence the individual's opinion in a much smaller period of time. It can also be observed that for the first 90 training episodes, the agent cannot influence the network in less than 30 time steps. However, this improves as the agent learns more techniques to gather influence and control the network more efficiently. Any training over 150, 000 episodes did not generate drastically different results.

Figure 6 shows a random group of 4 connected agents interacting in a social media environment with one control agent (Agent 1). The control agent can observe the expressed opinion of the other three individuals in the network. However, the agent can only influence two of these individuals (the size of the nodes indicates the level of influence each individual has). The network shows that all other agents have more influence than the control agent. The

agents in the network randomly expressed their opinions at each time step.

The evolution of opinions of the agents in Figure 6 is observed in Figure 7. The control agent aims to have all others in the network reach an opinion value of 0.9 (strong agreement) on the discussed topic. The control agent changes their opinion to influence the rest of the group. At the end of the discussion, the other agents in the network adopt an opinion similar to the goal of the control agent. Additionally, we can see that occasionally the control agent changes their opinion to match those of his/her neighbors in order to increase their overall reward.

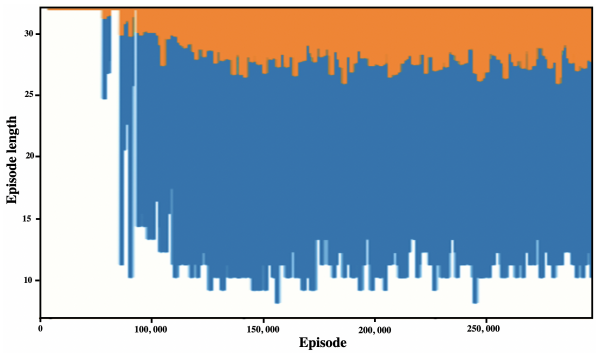


Figure 5: Episode length of a control agent over 250,000 training episodes.

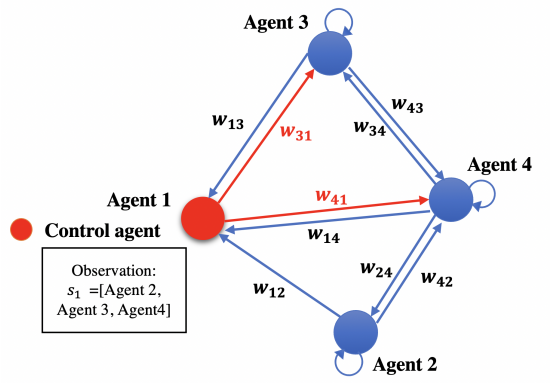


Figure 6: A network of 4 agents with one control agent (red) (size of the agent indicates their connections).

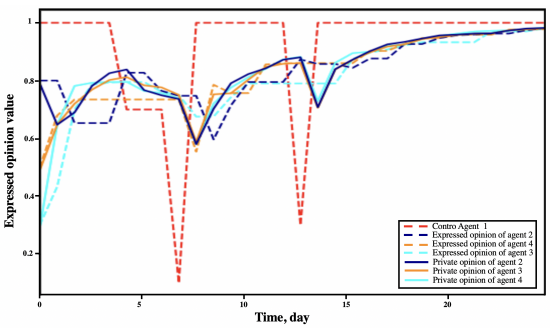


Figure 7: The evolution of private and expressed opinions of a network of 4 agents with one control agent (red).

6 Conclusion and Future Work

In this paper, we proposed an approach based on RL to solve the problem of disagreement between agents in multi-agent systems as well as social network control. To do so we used the expressed and private opinion dynamics model with asynchronous and synchronous updating dynamics to create an environment that closely resembles a social media network. Additionally, we used independent RL control agents to influence and control the network to a desired output. The method was validated using a random social media environment where agents interact and update their opinions randomly at every time step.

For future work, we are planning to investigate the interaction of control agents that are working cooperatively in very large networks. Additionally, we will explore the performance of controlling agents on evolving networks, and on networks with stubborn agents.

References

[1] M. N. Zareer, R. R. Selmic, "Modeling Competing Agents in Social Media Networks," in 2022 17th International Conference on Control, Automation, Robotics and Vision (ICARCV), 474–480, 2022, doi:10.1109/ICARCV57592.2022.10004247.

[2] H. Allcott, M. Gentzkow, "Social media and fake news in the 2016 election," Journal of economic perspectives, **31**(2), 211–236, 2017.

[3] A. Fourney, M. Z. Racz, G. Ranade, M. Mobius, E. Horvitz, "Geographic and Temporal Trends in Fake News Consumption During the 2016 US Presidential Election." in CIKM, volume 17, 6–10, 2017, doi:10.1145/3132847.3133147.

[4] A. Guess, B. Nyhan, J. Reifler, "Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign," 2018.

[5] E. Tacchini, G. Ballarin, M. L. Della Vedova, S. Moret, L. De Alfaro, "Some like it hoax: Automated fake news detection in social networks," arXiv preprint arXiv:1704.07506, 2017, doi:10.48550/arXiv.1704.07506.

[6] D. M. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, et al., "The science of fake news," Science, **359**(6380), 1094–1096, 2018.

[7] M. Howlett, Designing public policies: Principles and instruments, Routledge, 2019.

[8] P. Kotler, H. Kartajaya, I. Setiawan, Marketing 4.0: moving from Traditional to Digital, John Wiley & Sons, 2016.

[9] J. Van Dijck, T. Poell, M. De Waal, The platform society: Public values in a connective world, Oxford University Press, 2018.

[10] L. Floridi, J. Cowls, "A unified framework of five principles for AI in society," Machine learning and the city: Applications in architecture and urban design, 535–545, 2022, doi:10.1002/9781119815075.ch45.

[11] M. Saburov, "Reaching a consensus in multi-agent systems: a time invariant nonlinear rule," Journal of Education and Vocational Research, **4**(5), 130–133, 2013, doi:10.22610/jevr.v4i5.110.

[12] Y. Dong, W. Liu, F. Chiclana, G. Kou, E. Herrera-Viedma, "Are incomplete and self-confident preference relations better in multicriteria decision making? A simulation-based investigation," Information Sciences, **492**, 40–57, 2019, doi:10.1016/j.ins.2019.04.015.

[13] S. E. Parsegov, A. V. Proskurnikov, R. Tempo, N. E. Friedkin, "Novel multidimensional models of opinion dynamics in social networks," IEEE Transactions on Automatic Control, **62**(5), 2270–2285, 2016, doi:10.1109/TAC.2016.2613905.

[14] B. D. Anderson, M. Ye, "Recent advances in the modelling and analysis of opinion dynamics on influence networks," International Journal of Automation and Computing, **16**(2), 129–149, 2019, doi:10.1007/s11633-019-1169-8.

[15] M. H. DeGroot, "Reaching a consensus," Journal of the American Statistical association, **69**(345), 118–121, 1974.

[16] J. Becker, D. Brackbill, D. Centola, "Network dynamics of social influence in the wisdom of crowds," Proceedings of the national academy of sciences, **114**(26), E5070–E5076, 2017, doi:10.1073/pnas.1615978114.

[17] A. G. Chandrasekhar, H. Larreguy, J. P. Xandri, "Testing models of social learning on networks: Evidence from two experiments," Econometrica, **88**(1), 1–32, 2020, doi:10.3982/ECTA14407.

[18] E. Johnsen, "Social influence and opinions," J. Math. Sociology, **15**(3-4), 193–205, 1990.

[19] N. E. Friedkin, E. C. Johnsen, Social influence network theory: A sociological examination of small group dynamics, volume 33, Cambridge University Press, 2011.

[20] H. Kreft, W. Jetz, "Global patterns and determinants of vascular plant diversity," Proceedings of the National Academy of Sciences, **104**(14), 5925–5930, 2007.

[21] C. C. Childress, N. E. Friedkin, "Cultural reception and production: The social construction of meaning in book clubs," American Sociological Review, **77**(1), 45–68, 2012, doi:10.1177/0003122411428153.

[22] M. Ye, Y. Qin, A. Govaert, B. D. Anderson, M. Cao, "An influence network model to study discrepancies in expressed and private opinions," Automatica, **107**, 371–381, 2019, doi:10.1016/j.automatica.2019.05.059.

[23] S. E. Asch, "Effects of group pressure upon the modification and distortion of judgments," Groups, leadership, and men, 177–190, 1951.

[24] D. A. Prentice, D. T. Miller, "Pluralistic ignorance and alcohol use on campus: some consequences of misperceiving the social norm." Journal of personality and social psychology, **64**(2), 243, 1993.

[25] M. N. Zareer, R. R. Selmic, "Expressed and private opinions model with asynchronous and synchronous updating," in 2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2840–2846, IEEE, 2021, doi:10.1109/SMC52423.2021.9659269.

[26] S. Moscovici, M. Zavalloni, "The group as a polarizer of attitudes." Journal of personality and social psychology, **12**(2), 125, 1969.

[27] R. Felix, P. A. Rauschnabel, C. Hinsch, "Elements of strategic social media marketing: A holistic framework," Journal of business research, **70**, 118–126, 2017, doi:10.1016/j.jbusres.2016.05.001.

[28] S. Issenberg, "Cruz-connected data miner aims to get inside US voters' heads," Bloomberg, https://www. bloomberg. com/news/features/2015-11-12/is-the-republicanparty-s-killer-data-app-for-real, 2015.

[29] J. Shao, J. Qin, A. N. Bishop, T.-Z. Huang, W. X. Zheng, "A novel analysis on the efficiency of hierarchy among leader-following systems," Automatica, **73**, 215–222, 2016, doi:10.1016/j.automatica.2016.07.007.

[30] F. Dietrich, S. Martin, M. Jungers, "Control via leadership of opinion dynamics with state and time-dependent interactions," IEEE Transactions on Automatic Control, **63**(4), 1200–1207, 2017, doi:10.1109/TAC.2017.2742139.

[31] J. Shao, W. X. Zheng, T.-Z. Huang, A. N. Bishop, "On leader–follower consensus with switching topologies: An analysis inspired by pigeon hierarchies," IEEE Transactions on Automatic Control, **63**(10), 3588–3593, 2018, doi:10.1109/TAC.2018.2797205.

[32] A. Clark, B. Alomair, L. Bushnell, R. Poovendran, "Leader selection in multi-agent systems for smooth convergence via fast mixing," in 2012 IEEE 51st IEEE Conference on Decision and Control (CDC), 818–824, IEEE, 2012, doi:10.1109/CDC.2012.6426323.

[33] W. Su, X. Chen, Y. Yu, G. Chen, "Noise-Based Control of Opinion Dynamics," IEEE Transactions on Automatic Control, **67**(6), 3134–3140, 2022, doi:10.1109/TAC.2021.3095455.

[34] G. Albi, L. Pareschi, M. Zanella, "On the optimal control of opinion dynamics on evolving networks," in System Modeling and Optimization: 27th IFIP TC 7 Conference, CSMO 2015, Sophia Antipolis, France, June 29-July 3, 2015, Revised Selected Papers 27, 58–67, Springer, 2016, doi:10.1007/978-3-319-55795-3_4.

[35] B. Ross, L. Pilz, B. Cabrera, F. Brachten, G. Neubaum, S. Stieglitz, "Are social bots a real threat? An agent-based model of the spiral of silence to analyse the impact of manipulative actors in social networks," European Journal of Information Systems, **28**(4), 394–412, 2019, doi:10.1080/0960085X.2018.1560920.

[36] D. Sohn, N. Geidner, "Collective dynamics of the spiral of silence: The role of ego-network size," International Journal of Public Opinion Research, **28**(1), 25–45, 2016, doi:10.1093/ijpor/edv005.

[37] R. S. Sutton, A. G. Barto, Reinforcement learning: An introduction, MIT press, 2018.

[38] M. Langlois, R. H. Sloan, "Reinforcement learning via approximation of the Q-function," Journal of Experimental & Theoretical Artificial Intelligence, **22**(3), 219–235, 2010, doi:10.1080/09528130903157377.

[39] C. Musco, C. Musco, C. E. Tsourakakis, "Minimizing polarization and disagreement in social networks," in Proceedings of the 2018 world wide web conference, 369–378, 2018, doi:10.1145/3178876.3186103.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 8, No.5, 70-76 (2023)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Resonance Coil Design for a Novel Battery Cell Balancing with using Near-Field Coupling

Juhyeon Jeon, Dongho Lee*

Department of Electrical and Control Engineering, Mokpo National University, Naju, 58277, Republic of Korea

ARTICLE INFO

Article history:
Received:26 June, 2023
Accepted:18 October, 2023
Online: 30 October, 2023

Keywords:
Battery cell balancing
Near-field coupling
Wireless power transfer
Coil design

ABSTRACT

In this paper, we delve into the pressing necessity for proficient battery cell balancing, an imperative in the context of the escalating adoption of renewable energy and electric vehicles. While traditional methodologies, including the passive technique, offer a straightforward and cost-effective solution, they compromise on efficiency. The active technique, though superior in efficiency, is hindered by its innate restriction of transferring energy solely to proximate cells, thus prolonging the balancing process. To address these limitations, we introduce a novel near-field coupling method centered on a meticulously designed resonant coil with an emphasis on achieving a larger Q-factor, a pivotal factor for enhanced battery cell balancing. This augmented Q-factor not only propels our approach past the passive method in efficiency but also catalyzes rapid balancing by enabling wireless energy transfer to cells regardless of their relative positioning. Validating our theoretical insights, we developed physical coil prototypes and adopted a Series-Parallel circuit configuration, steered by the resonant coil's Q-factor. Preliminary experiments with three batteries substantiate our claim, showcasing that our proposed technique achieves cell balancing with approximately double the speed of conventional strategies.

1. Introduction

This paper is an extension of a work originally presented in ITC-CSCC 2022 [1]. Through the Paris Climate Change Agreement, major countries around the world aim to achieve Net-zero, setting their carbon emissions to '0' by 2050. Among the various policies being pursued by South Korea to achieve these goals, the expansion of renewable energy and electric vehicles plays a major role [2-3]. Renewable energy serves as an alternative to fossil fuels and nuclear power, with solar and wind power at its core. However, such energy generation is intermittent depending on weather conditions, necessitating the use of an Energy Storage System (ESS). The ESS is composed of large-capacity battery packs[4], and each of these battery cells demonstrates slight discrepancies in the State of Charge (SOC) during manufacturing. This deviation tends to increase during charging and discharging processes, leading to problems in battery life and efficiency[5-6]. Battery cell balancing is a technique used to rectify these issues. On the other hand, as the proportion of electric vehicles continues to increase globally, a new environmental issue, namely the 'problem of battery disposal,' has been highlighted [7]. One solution to this problem that is actively researched involves recycling spent batteries to create ESS [8-9]. In this context, the battery cell balancing technology is recognized as playing a crucial role in maximizing the efficiency and lifespan of ESS.

However, the conventional battery cell balancing, as discussed in Section 2, has limitations. As the proportion of renewable energy and electric vehicles continues to expand, there is a growing need for faster and more efficient battery cell balancing techniques. Therefore, in this paper, we propose a novel battery cell balancing method and evaluate its performance.

2. Conventional Battery Cell Balancing

Traditional cell balancing can be divided into passive battery cell balancing and active battery cell balancing. The passive type connects resistors to the battery to expend energy for balancing, while the active type connects capacitors, inductors, or transformers to the battery to transfer energy to adjacent batteries for balancing [10-13].

2.1. Passive Battery Cell Balancing

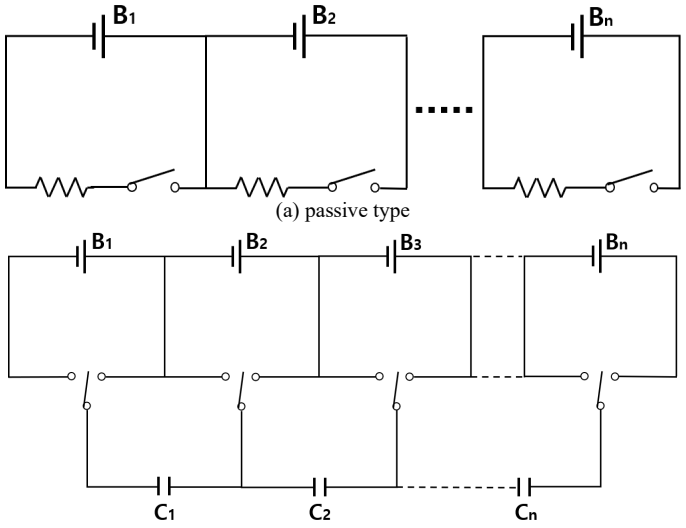
The passive type, which exhausts energy from the batteries with resistors, is based on the battery with the lowest State of Charge (SOC). This method is widely used due to its simple circuit and low cost. However, it has several drawbacks: it wastes energy leading to low efficiency, it generates heat making it prone to fire,

*Corresponding Author: Dongho Lee, dongho.lee864@gmail.com

and it can't perform balancing during discharge operation. Therefore, it is not suitable for high-efficiency battery cell balancing. Figure 1 (a) illustrates a switching shunting resistor method of the passive type, where all switches are controlled identically or individually depending on the SOCs of the batteries.

2.2. Active Battery Cell Balancing

The active type transfers energy from cells with a relatively high SOC to those with a lower SOC. While it's more efficient than the passive type, its circuit is complex, control is difficult, and it only allows balancing between adjacent cells, which slows down the balancing speed. Hence, it is not suitable for high-speed battery cell balancing. Figure 1 (b) shows a switching capacitor method of the active type. This method is a simpler and cheaper option within the active type, but if the cells with the highest and lowest SOCs are assumed to be at the ends of the circuit, more switches need to be controlled, leading to a longer time needed for balancing. Other methods to overcome this issue have complicated circuit configurations or high costs.



(a) passive type

(b) active type

Figure 1: Conventional battery cell balancing

2.3. Proposal Battery Cell Balancing

As such, conventional battery cell balancing methods are not suitable for high-speed, high-efficiency applications due to their respective disadvantages. Table 1 summarizes the characteristics, advantages, and disadvantages of typical conventional battery cell balancing methods. It shows that passive type is not suitable for high efficiency, and active type is not suitable for high speed. Therefore, a novel battery cell balancing method, which is more efficient than the passive type and faster than the active type, is needed to overcome the disadvantages of conventional methods. This paper proposes a novel battery cell balancing method that uses near-field coupling for balancing.

Table 1: Summary of typical conventional battery cell balancing

category	passive type		active type		
method	shunt resistor	switching resistor	switching capacitor	inductor	Proposal
devices used	n resistors	n switches and resistors	n-1 capacitors and 2n switches	PWM	N+1 coils and N Rectifiers
advantage	simple circuit and low cost		simple circuit and low cost	reasonable balancing speed	Fast balancing & Reasonable efficiency
disadvantage	low efficiency and safety issue		slow balancing speed	low efficiency & high cost	high cost

3. Near-Field Coupling Battery Cell Balancing

The new battery cell balancing method proposed in this paper utilizes 'Near-Field Coupling'. This method enables selective cell balancing by transmitting energy to specific cells. 'Near-Field Coupling' is a method of transmitting energy between objects in close proximity using a magnetic resonance method. The resonance frequency selected for this method is 13.56MHz in the ISM band (Industrial Scientific Medical Band). Figure 2 illustrates the circuit of the proposed method. When the State of Charge (SOC) of battery B1 is lower than that of B2 and B3, the transmitter's capacitor is variably adjusted to achieve impedance matching with the receiver of B1's battery. The energy transmission continues until the SOC of B2 and B3 matches that of B1.

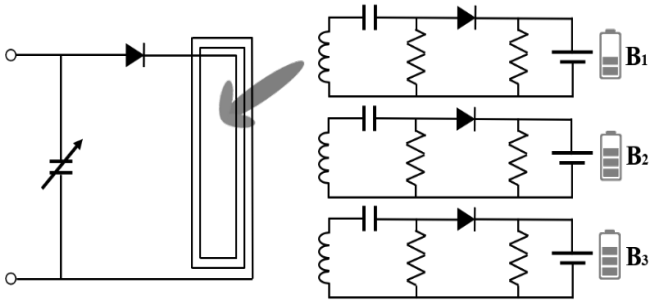


Figure 2: Proposed battery cell balancing with using near-field coupling

This proposed method is more efficient than the conventional passive type. The passive type consumes energy, leading to low efficiency, particularly because the balancing is based on the battery with the lowest SOC. However, the proposed method is more efficient because it supplements energy rather than wasting it. The energy efficiency of wireless power transmission can be divided into transmitter efficiency, receiver efficiency, and coil efficiency. With the transmitter efficiency being less than 80%, receiver efficiency at 95%, and coil efficiency less than 85%, the combined overall efficiency can reach approximately 64.6%[14-15]. This makes the proposed method more suitable for high efficiency than the passive type.

Additionally, the proposed method is faster in balancing speed than the active type. The primary reason for the slower balancing speed of the active type is because it can only transfer energy to adjacent batteries. On the other hand, our proposed method allows energy transfer even to relatively distant cells by variably adjusting the transmitter's capacitor to achieve impedance matching with a specific receiver. This results in a faster balancing speed than the active type.

4. Impedance Matching

Impedance serves as a measure of electrical flow resistance, and impedance matching refers to all methods aimed at reducing reflection caused by differences in impedance between the input and output terminals. That is, the more closely the impedances of both terminals are matched, the more the reflection due to the impedance difference reduces, leading to less energy loss and higher efficiency[16]. Such impedance matching techniques enable selective cell balancing. By designing so that the impedance matches with a specific battery and differs from other batteries, it is possible to deliver energy only to the desired cells. The proposed method to implement this involves differentiating the mutual inductance and coupling coefficient between the transmitter and each receiver depending on the location of the receivers, based on the transmitter. This leads to unique impedance differences, enabling the selection of specific receiver cells.

4.1. Mutual Inductance

Mutual inductance is a value that represents the magnetic field interaction between two coils. As shown in Figure 3, When the current flowing through a coil changes, the magnetic field passing through that coil also changes.

This change induces an electrical change in an adjacent coil, and this interaction is referred to as mutual inductance. This varies according to distance, coil orientation, and position, leading to each battery cell having a unique value.

4.2. Coupling coefficient

The coupling coefficient is calculated using the mutual inductance and the inductance of the coil itself, and represents the efficiency of energy transfer between two coils. The higher the coupling coefficient, the higher the energy transmission efficiency, so it's crucial to adjust the coupling coefficient appropriately. This also varies according to cell distance, similar to mutual inductance.

$$k = \frac{M}{\sqrt{L_1 L_2}} \tag{1}$$

Through such impedance matching, the energy transmission efficiency between battery cells can be increased, and based on this, efficient cell balancing can be implemented.

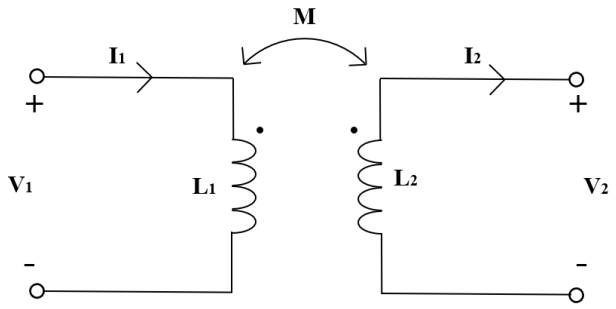


Figure 3: Mutual inductor circuit

5. Transmitter and Receiver Coil Design

Coil patterns can be designed in various forms such as square, circular, or hexagonal. However, in this study, we used the square pattern for easy modeling, based on the previous research which states that the difference between patterns is not significant at the frequency of 13.56MHz utilized in this research[1]. Furthermore, we decided on an LC circuit configuration suitable for selective battery cell balancing according to the Q-factor.

5.1. Q-factor

As shown in Figure 4, the Q-factor, defined by equation (2), represents the ratio of the resonance frequency ω_0 to the 3dB bandwidth, defined as the difference between the higher frequency ω_h and the lower frequency ω_l, situated 3dB below ω_0. For effective cell balancing application, a sufficiently large Q-factor is indispensable, signifying reduced insertion loss at specific frequencies and significant increase in insertion loss with minor frequency deviations. This narrow bandwidth impedance matching is vital for the proposed selective cell balancing approach. Analyzing equations (3) and (4), it's evident that for a consistent resonance frequency, a series LC circuit demands a larger inductance for a heightened Q, while a parallel LC circuit requires a smaller inductance for the same effect.

Given the physical constraints like battery size and coil diameter, the receiver coil was designed with a parallel configuration, where a reduced inductance yields a higher Q. Conversely, to ensure efficient energy transmission via impedance matching, the transmitter coil, designed in a series configuration, incorporates a larger inductance for an enhanced Q. This culminated in the LC circuits for both transmission and reception units being structured as Series-Parallel (SP).

$$Q = \frac{\omega_0}{\omega_h - \omega_l} \tag{2}$$

$$Q(series\ LC) = \omega_0 \frac{2W_m}{P_{loss}} = \frac{\omega_0 L}{R} = \frac{1}{\omega_0 RC} \tag{3}$$

$$Q(shunt\ LC) = \omega_0 \frac{2W_m}{P_{loss}} = \frac{R}{\omega_0 L} = \omega_0 RC \tag{4}$$

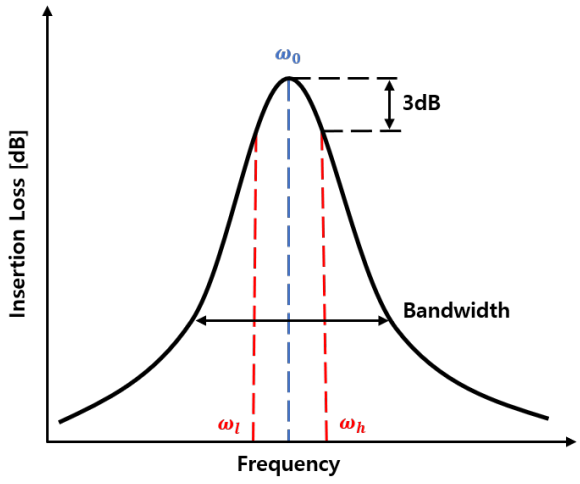


Figure 4: Q-factor in resonance circuit

Building on the crucial role of the Q-factor, the transmitter was crafted with an increased inductance, embodying a series LC circuit, fine-tuned for selective cell balancing. In parallel, the receiver was architected with diminished inductance, aligning with

a parallel LC circuit, to cater to the same balancing purpose. Factors like pattern spacing, internal and external diameters, and thickness play pivotal roles in determining coil inductance. For instance, in Figure 5, pattern spacings for the transmitter and receiver coils were fixed at 1mm and 2mm respectively, echoing past research indicating augmented inductance with reduced spacing. Coil measurements in this study were not merely based on theoretical derivations; instead, they were corroborated through actual coil measurements and simulations, ensuring accuracy by accounting for potential production discrepancies.

5.2. Coil Design

By leveraging the Q-factor, it was ascertained that the inductance of the transmitter coil can be designed to be larger, ensuring that the serial LC circuit configuration is suitable for selective cell balancing. Conversely, by designing the inductance of the receiver coil to be smaller, it was confirmed that the parallel LC circuit configuration is suitable for selective cell balancing. Figure 6 presents the blueprint of the transmitter and receiver coils designed using 3D EM Simulation, as well as the fabricated PCB. (a) Represents the transmitter coil, and (b) represents the receiver coil.

The inductance of the coil is determined by various variables as shown in Figure 6, and the inductance changes even if only one of these conditions changes. The smaller the pattern spacing 's' of the transmitter and receiver coil, the larger the inductance; therefore, the transmitter was designed with 1mm and the receiver with 2mm. Similarly, the smaller the internal diameter 'Di', the larger the inductance; thus, the transmitter was designed with 1mm and the receiver with 5mm. As can be seen from Figure 6, considering the size of the battery, the receiver's outer diameter 'Do' was 20x19mm, while the transmitter, considering the size of the receiver, was 81x31mm. The thickness 'w' was designed to be the same at 1mm. There are various papers that infer inductance values using these variables[17-18], but in this paper, the inductance is inferred by comparing the actual measurement results of the coil with the simulation.

5.3. De-embedding

Circuits that use RF (Radio Frequency), such as 13.56MHz, cannot be measured for voltage and current like normal circuits. This is because as the frequency increases, the size of the wavelength decreases, and the voltage and current measurement values change depending on the circuit location. Therefore, RF circuits should use S-parameters, which interpret the circuit as a ratio of output voltage to input voltage, and in this paper, the proportion of the input signal reaching the output port was measured for efficiency verification.

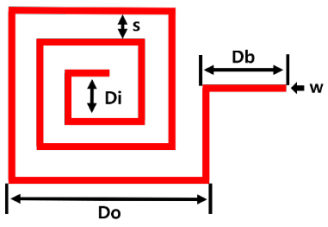
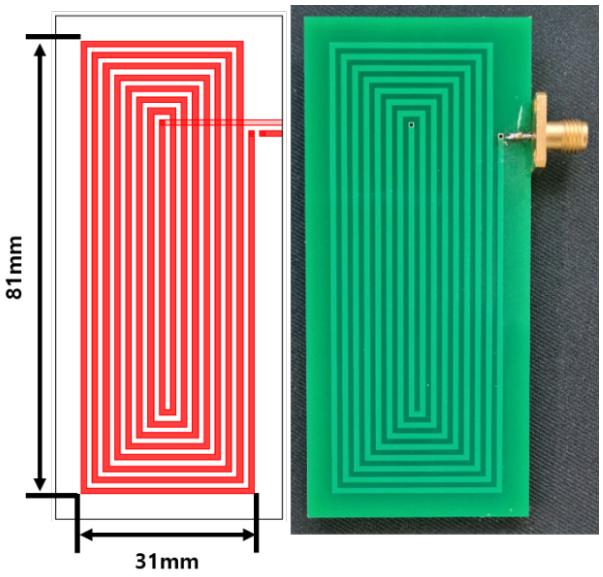
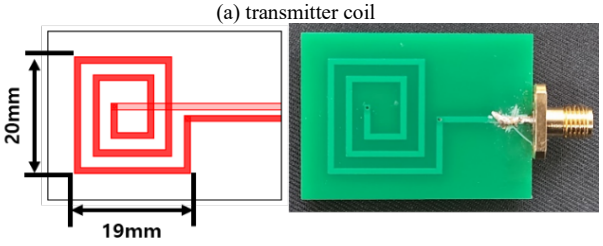


Figure 5: Coil variables



(a) transmitter coil



(b) receiver coil

Figure 6: Transmitter and receiver coil designed using 3D EM Simulation, as well as the fabricated PCB

To measure S-parameters, an NA(Network Analyzer) was used. For the accuracy of the measurement results, the NA's own calibration function is used to remove error components, and there are standardized methods such as SOLT(Short, Open, Load, Through), TRL(Through, Reflect, Line). In this paper, an SOLT calibration kit was used, and components such as phase delay that are not removed by general correction should be further corrected through de-embedding. De-embedding is not a standardized method like ordinary correction, so it must be conducted according to the research environment. When the circuit in the open state is represented as a Smith chart, it is indicated in the center to the right like the red dot in Figure 7, and the degree of phase delay can be understood by arbitrarily opening the fabricated transmitter and receiver PCB coil and comparing it with the measurement. The transmitter and receiver PCB coil data measured with the NA was adjusted in the simulation so that the electrical length of the ideal Transmission Line came out as in Figure 7. Figure 8 shows the result, with the blue plot being the result measured with the NA and the red plot being the result adjusted for electrical length. Figure 9 is the result of the transmitter and receiver de-embedding, and including up to the transmission line, it becomes an inductor model with phase delay removed. (a) is the result of transmitter de-embedding with an electrical length of –17.4° at 1GHz. (b) is the result of receiver de-embedding, changing the electrical length to –18.5°. Therefore, by applying the electrical length obtained through the de-embedding of the transmitter and receiver to the measurement results, accurate results with phase delay removed can be obtained.

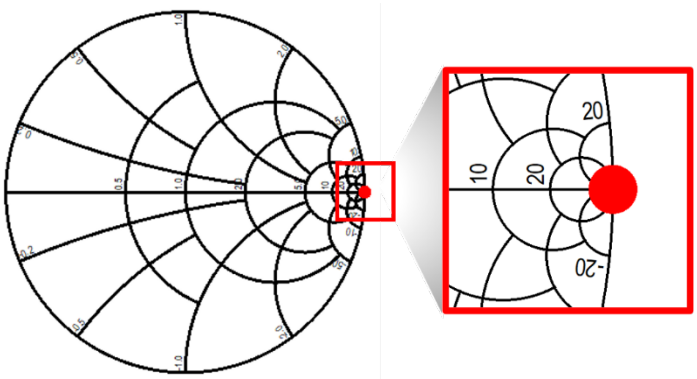
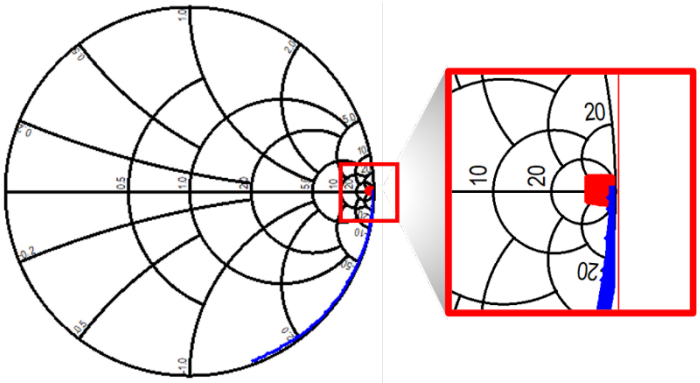


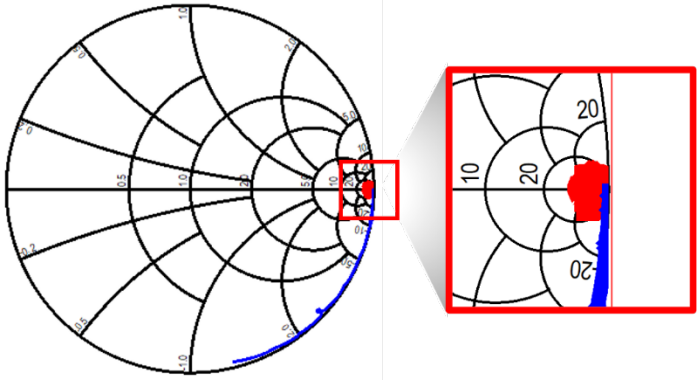
Figure 7: Open circuit on Smith chart

5.4. Inductance inference

Figure 10 shows the results of comparing the corrected measurement results through the transmitter's de-embedding, 3D EM Simulation, and Schematic 2600nH. By comparing the phase using the s-parameter and Smith chart, there was an error of less than 3% at 13.56MHz, which makes it possible to represent the coil with one inductor of the Schematic. The receiver was compared with 130nH, and the results were similar to the transmitter. Figure 10 shows the results of comparing the measurement results of the LC circuit determined through modeling of the transmitter and receiver and the inferred LC circuit. The schematic circuit of the resonant circuit in this application is shown in Figure 11.

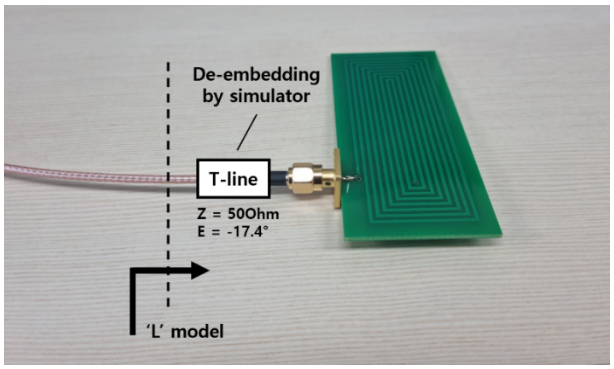


(a) impedance of transmitter on Smith chart

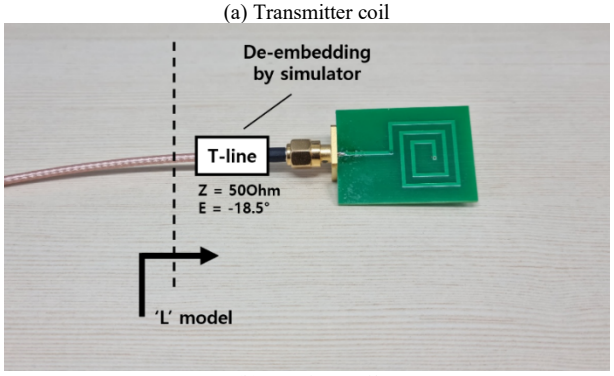


(b) Impedance of receiver on Smith chart

Figure 8: Transmitter and receiver coil open state measurement results(blue) electrical length adjusted results(red)



(a) Transmitter coil



(b) Receiver coil

Figure 9: De-embedding of transmitter/receiver coil

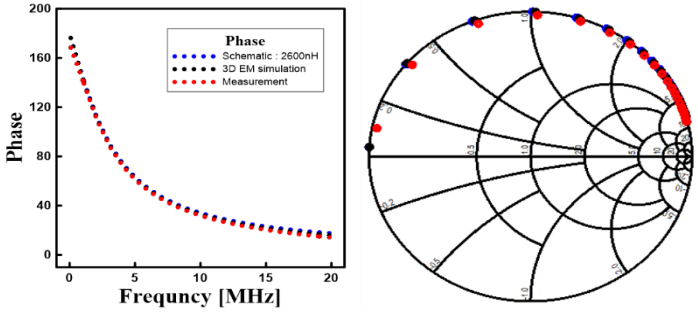


Figure 10: Transmitter de-embedding, 3D EM Simulation and Schematic 2600nH comparison results, error of less than 3% at 13.56MHz

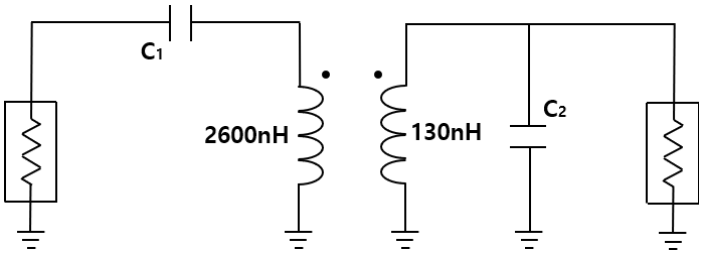


Figure 11: The schematic of measured resonance coils

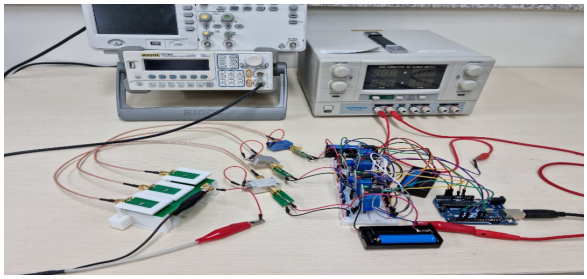


Figure 12: Cell balancing measurement set-up

5.5. Cell balancing measurement

Figure 12 shows an experiment with three battery cells with the proposed near-field coupling. One of the three cells has less charge than the other two. Figure 13 compares the regular balancing results with the balancing results obtained by implementing Near-Field Coupling using the coils and inferred capacitance modeled through this paper. The X-axis is time and the Y-axis is the voltage difference with the reference battery. The red graph represents the regular balancing results, and the blue graph represents the balancing results with added Near-field. Assuming that balancing was successfully achieved when the voltage difference with the reference battery decreased to less than 1%, the regular balancing took about 248 minutes, while the balancing with added Near-Field took about 121 minutes, nearly twice as fast as the former.

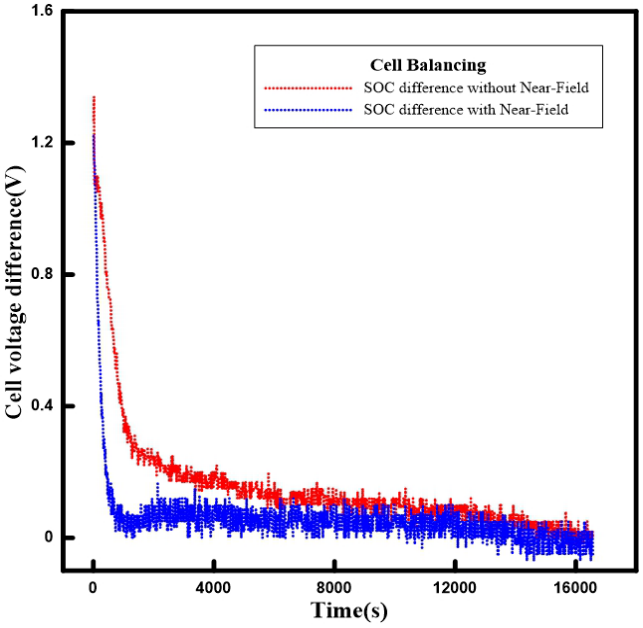


Figure 13: Cell voltage difference with and without near-field coupling

6. Conclusion

Our study confirmed that configuring the transmitter with a serial LC circuit and the receiver with a parallel LC circuit, as determined through the Q-factor, is more suitable for selective battery cell balancing. The coil of the receiver was constructed taking into account the battery size, which consequently determined the size of the transmitter coil. After the coil construction, uncorrected parasitic components were eliminated through additional correction processes such as de-embedding. Additionally, the electrical length of the transmitter coil's t-line was adjusted to -17.4° at 1GHz and the receiver coil to -18.5° to remove phase delay.

A comparison of the more accurate measurement results obtained through de-embedding, 3D EM Simulation, and Schematic with the s-parameter and Smith chart revealed less than 3% error at 13.56MHz. As the Schematic was compared using a single inductor, it was confirmed that the coil to be produced could be represented by a single inductor. Accordingly, we inferred that the transmitter coil was 2600nH and the receiver coil was 130nH. A simple test of cell balancing with three battery cells using the proposed method confirms that cell balancing is achieved compared to no cell balancing. The results of this experiment confirm that there is not enough unwanted coupling between the coils

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This Research was supported by Research Funds of Mokpo National University in 2022

References

[1] J. Jeon, W. Park, S. Pyo and D. Lee, "Coil design and de-embedding for a novel cell balancing circuit using near-field coupling," 2022 37th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), 546-548, 2022, doi: 10.1109/ITC-CSCC55581.2022.9895073.

[2] A. Qazi et al., "Towards Sustainable Energy: A Systematic Review of Renewable Energy Sources, Technologies, and Public Opinions," in IEEE Access, **7**, 63837-63851, 2019, doi: 10.1109/ACCESS.2019.2906402.

[3] S. Thangavel, D. Mohanraj, T. Girijaprasanna, S. Raju, C. Dhanamjayulu and S. M. Muyeen, "A Comprehensive Review on Electric Vehicle: Battery Management System, Charging Station, Traction Motors," in IEEE Access, **11**, 20994-21019, 2023, doi: 10.1109/ACCESS.2023.3250221.

[4] X. Xu, M. Bishop, D.G. Oikarinen and C. Hao, "Application and modeling of battery energy storage in power systems," CSEE journal of power and energy system, **2**(3), 82-89, 2016, doi: 10.17775/CSEEJPES.2016.00039

[5] J. Shen, S. Dusmez and A. Khaligh, "Optimization of Sizing and Battery Cycle Life in Battery/Ultracapacitor Hybrid Energy Storage Systems for Electric Vehicle Applications," in IEEE Transactions on Industrial Informatics, **10**(4), 2112-2121, 2014, doi: 10.1109/TII.2014.2334233.

[6] Q. Ouyang, W. Han, C. Zou, G. Xu and Z. Wang, "Cell Balancing Control For Lithium-Ion Battery Packs: A Hierarchical Optimal Approach," in IEEE Transactions on Industrial Informatics, **16**(8), 5065-5075, 2020, doi: 10.1109/TII.2019.2950818.

[7] D. L. Thompson, J. M. Hartley, S. M. Lambert, M. Shiref, G. D. J. Harper, E. Kendrick, P. Anderson, K. S. Ryder, L. Gaines and A. P. Abbott, "The importance of design in lithium ion battery recycling – a critical review" Green Chem, **22**, 7582-7584, 2020, doi: 10.1039/D0GC02745F

[8] M. A. Hannan, M. M. Hoque, A. Hussain, Y. Yusof and P. J. Ker, "State-of-the-Art and Energy Management System of Lithium-Ion Batteries in Electric Vehicle Applications: Issues and Recommendations," in IEEE Access, **6**, 19362-19378, 2018, doi: 10.1109/ACCESS.2018.2817655.

[9] E. Hossain, D. Murtaugh, J. Mody, H. M. R. Faruque, M. S. Haque Sunny and N. Mohammad, "A Comprehensive Review on Second-Life Batteries: Current State, Manufacturing Considerations, Applications, Impacts, Barriers & Potential Solutions, Business Strategies, and Policies," in IEEE Access, **7**, 73215-73252, 2019, doi: 10.1109/ACCESS.2019.2917859.

[10] J. Qi and D. Dah-Chuan Lu, "Review of battery cell balancing techniques," 2014 Australasian Universities Power Engineering Conference (AUPEC), 1-6, 2014, doi: 10.1109/AUPEC.2014.6966514.

[11] M. Daowd, N. Omar, P. Van Den Bossche and J. Van Mierlo, "Passive and active battery balancing comparison based on MATLAB simulation," 2011 IEEE Vehicle Power and Propulsion Conference, 1-7, 2011, doi: 10.1109/VPPC.2011.6043010.

[12] Das, Utpal Kumar, et al. "Advancement of lithium-ion battery cells voltage equalization techniques: A review." Renewable and Sustainable Energy Reviews **134**, 110227, 2020, doi: 10.1016/j.rser.2020.110227.

[13] Diao, Weiping, et al. "Active battery cell equalization based on residual available energy maximization." Applied energy **210,** 690-698, 2018, doi: 10.1016/j.apenergy.2017.07.137

[14] A. N. Laskovski and M. R. Yuce, "Class-E oscillators as wireless power transmitters for biomedical implants," 2010 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL 2010), 1-5, 2010, doi: 10.1109/ISABEL.2010.5702913.

[15] G. Lee, B. H. Waters, B. J. Mahoney, J. R. Smith and W. S. Park, "An investigation of cross-coupling for magnetically coupled wireless power transfer," 2013 Asia-Pacific Microwave Conference Proceedings (APMC), 80-82, 2013, doi: 10.1109/APMC.2013.6695197.

[16] T. C. Beh, M. Kato, T. Imura, S. Oh and Y. Hori, "Automated Impedance

Matching System for Robust Wireless Power Transfer via Magnetic Resonance Coupling," in IEEE Transactions on Industrial Electronics, **60(**9) 3689-3698, 2013, doi: 10.1109/TIE.2012.2206337.

[17] K. Chen and Z. Zhao, "Analysis of the Double-Layer Printed Spiral Coil for Wireless Power Transfer," in IEEE Journal of Emerging and Selected Topics in Power Electronics, **1(**2), 114-121, 2013, doi: 10.1109/JESTPE.2013.2272696.

[18] U. -M. Jow and M. Ghovanloo, "Design and Optimization of Printed Spiral Coils for Efficient Transcutaneous Inductive Power Transmission," in IEEE Transactions on Biomedical Circuits and Systems, **1(**3), 193-202, 2007, doi: 10.1109/TBCAS.2007.913130.

Advances in Science, Technology and Engineering Systems Journal Vol. 8, No.5, 77-93 (2023)

www.astesj.com

Special Issue on Innovation in Computing, Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Design and Prototyping of a 3DOF Worm-drive Robot Arm

Ian Spencer Howard*

SECAM, University of Plymouth, Plymouth, PL4 8AA UK

ARTICLE INFO

Article history:
Received:03 August, 2023
Accepted:28 October, 2023
Online: 30 October, 2023

Keywords:
Key 3DOF Revolute arm
Key Worm-drive actuation
Key 3D printing

ABSTRACT

Many designs for robot arms exist. Here we present an affordable revolute arm, capable of executing simple pick-and-place tasks. The arm employs a double parallelogram structure, which ensures its endpoint angle in the plane of the upper arm remains fixed without the need for additional actuation. Its limbs are fabricated from circular tubes made from bonded carbon fiber, to ensure low moving mass while maintaining high rigidity. All custom structural elements of the arm are produced via 3D printing. We employ worm-drive DC motor actuation to ensure that stationary configurations are maintained without the necessity of continuous motor power. Our discussion encompasses an analysis of the arm's kinematics. A simulation of the arm's operation was carried out in MATLAB, revealing key operational metrics. In conclusion, we achieved extrinsic endpoint position tracking by implementing its inverse kinematics and PID control using a microcontroller. We also demonstrate the arm's functionality through simple movement tracking and object manipulation tasks.

1. Introduction

1.1. Extension of previous work

This paper builds upon work on a prototype 2D planar arm first published in the proceedings of the 2022 International Conference on System Science and Engineering (ICSSE) [1]. We explicitly highlight the extensions of previous work and novel contributions of this study at the beginning of the discussion section.

1.2. Overview

Brought about by a combination of a dwindling agricultural workforce and advancements in robotics technology, the deployment of robots for harvesting fruits and vegetables has become an increasingly important area of development [2]. As a result, there is a growing demand for low-cost robots to not only reduce the cost of harvesting but also to address the shortages in agricultural workers willing to perform the task.

As a typical application area, we consider the development of robotic arms designed to harvest berries from plants. For such an application, robotic arms are normally attached to a computer-controlled mobile robot base. These systems typically operate in structured environments, such as greenhouses, and must autonomously identify, pick, and collect ripe berries. The design of suitable robotic arms involves trade-offs, as multiple requirements need to be met [3].

1.3. Requirements

To offer a realistic alternative to employing human workers, any useful arm design needs to perform its picking tasks effectively with no greater, and preferably lower, overall operating costs. Therefore, keeping the cost of arm construction low is an important design consideration.

In terms of performance, the workspace of the arm must be sufficient for it to reach the berries chosen for picking, and it must be able to do so with reasonable accuracy (of the order of millimeters). Ideally, it should accomplish this with minimal correction during movement, avoiding the need for visual servoing [4]. The arm must also generate sufficient force to hold and operate a suitable end effector mechanism to pick the berries. Pulling off a raspberry can require up to 10N [5], although alternative removal techniques, such as cutting, are also viable [6]. Force generation requirements place demands on the joint torques needed, as well as on the stiffness and strength of the arm structure. Clearly, a balance must be struck between arm link length and stiffness, given that arm stiffness decreases rapidly as a function of link length.

The actuation and control mechanisms of the arm should prioritize power efficiency, as mobile picking systems generally

*Corresponding Author: Ian Spencer Howard, SECAM, University of Plymouth. UK, ian.howard@plymouth.ac.uk

operate from batteries. To address limitations in the use of batteries, researchers are exploring the possibilities of harnessing solar energy to recharge the batteries when daylight conditions permit [7].

The arm must ensure safe operation in the proximity of people and guidelines for safe robot operations are specified by the ISO 10218 standard [8]. To do so, it must effectively manage occasional, inevitable collisions with items in its environment, which could include greenhouse plant support infrastructure or even curious human workers standing too close to the robot. Paramount attention must be paid to this issue, to prevent any injury or damage resulting from inadvertent collisions.

Several strategies can facilitate safe robotic arm operation. One approach involves building inherent compliance into their designs [9,10]. One such design, the Gummiarm, achieves this by making use of compliant non-linear tendons driven by Dynamixels [11,12]. However, incorporating passive compliance can often compromise performance.

Another way to enhance robot safety is to limit the arm's kinetic energy due to movement. This can be achieved by restricting the maximum operational speed and ensuring the arm has a lightweight structure, thereby reducing potential damage arising from a collision.

1.4. Collaborative robots

Much work has been carried out to make industrial robots safe around people [13–15]. A collaborative robot, (also known as a cobot), is a specialized robot, engineered to work in tandem with humans within a shared workspace.

Cobots are often equipped with an array of sensors, including cameras, force sensors, and proximity sensors, which enable them to sense and interact with their surroundings. These sensors allow cobots to identify human presence, avert collisions, and react to environmental changes. Furthermore, cobots incorporate safety features, such as force-limiting and speed-limiting controls, which contribute to accident prevention and the protection of human co-workers. A popular safety strategy employed in cobots, is to monitor their force generation [16]. If force levels exceed expected task values, the robot identifies this anomaly as a probable fault condition and safely deactivates the robot in a controlled manner.

Numerous companies, researchers, and engineers have embraced the cobot concept. Universal Robots, a Danish company, was among the first to introduce a cobot, and Rethink Robotics was also an early pioneer in the cobot domain. Cobots are gaining traction across a broad spectrum of industries, including manufacturing, logistics, and healthcare.

1.5. Lightweight arms

Several lightweight arms are commercially available. The Barrett WAM, popular among researchers, features low weight and employs a cable drive. Its actuation exhibits very low backlash and low friction, while the arm itself demonstrates a low effective mass [17].

The ANYpulator is another example of a lightweight robotic arm. It employs force control and utilizes lightweight carbon fiber links in its construction to minimize moving mass [18].

The DLR lightweight robot provides another example of high-quality arms capable of safe operation [15]. Kuka has commercialized its design, which also serves as the basis for the Franka Emika arm [19].

1.6. Existing low-cost robots

Most of the commercially available robots mentioned earlier are quite expensive. However, several low-cost robotic arm designs also exist. For instance, Reachy employs off-the-shelf components and fully 3D-printed limb sections [20].

The BioRob-Arm [21] is a lightweight design that utilizes compliant actuation. It is based on an antagonistic, series-elastic actuation that employs cable transmission [22].

Another compliant, 7-DOF (Degrees of Freedom) robotic arm was developed at Stanford University. To keep costs down, the design makes use of stepper motor drive and adopts series-elastic transmission in the lower arm and base. It adopts Dynamixels for actuation in the remaining three [23].

More recently, the Berkeley Blue 7-degree-of-freedom arm has been developed with the intention of making a low-cost, force-controlled robot available to the wider AI community [24].

2. Arm design

2.1. Design and construction overview

In this section, we focus on designing a lightweight arm driven by worm-gear motors, via HTD 5M x 15mm wide timing belts, enabling it to maintain static configurations without requiring constant motor power. The arm design is based on a revolute configuration. See Figure 1 for a photograph of the arm, Figure 2 for a close-up of the pulley actuation, and Figure 3 for labelled schematics indicating important parts of the robot design.

The design of the arm was carried out using Autodesk Fusion 360. To ensure minimal weight, the upper and lower parallelogram arm sections employed carbon fiber tubing, with dimensions of 16x14mm in diameter and 250mm in length.

All parts that were custom-designed were initially exported in the STL file format, subsequently sliced employing Ultimaker Cura software, and 3D printed utilizing PLA+ material with a Creality CR-6 SE printer. To optimize stiffness and strength, arm joint components were printed with a 100% infill [25]. Additional components were printed with a 35% infill to strategically minimize their moving mass. Keeping the weight of the components low was particularly important for the upper arm mechanism support pillars, as they rotate during the operation of the arm.

Loctite Precision superglue was used to attach the carbon fiber tubes to their corresponding 3D printed PLA+ parts.

2.2. Rotary base

The first arm joint is achieved using a platform that rotates around a vertical axis. This platform supports a planar arm mechanism, which provides an additional two degrees of freedom.

The top of the rotary base assembly is fabricated from plate sections. These are both 20mm-thick and are 3D-printed. They are supported by two side pillars resting on a single, similarly thick

3D-printed base plate. These components are secured together using four long bolts. The top two base plates tightly secure a deep groove ball bearing in place. It has a large internal bore of 50mm to accommodate a 3D-printed shaft, which passes through the bearing. The shaft is affixed to the bearing assembly, using a 3D-printed clamping collar. This provides the shaft with radial and axial support. The collar incorporates an integrated 40-tooth HTD 5M x 15mm timing pulley. The shaft represents the arm's first degree of freedom. The upper end of the main shaft features a flat plate with mounting holes, so the upper arm assembly can be bolted onto it.

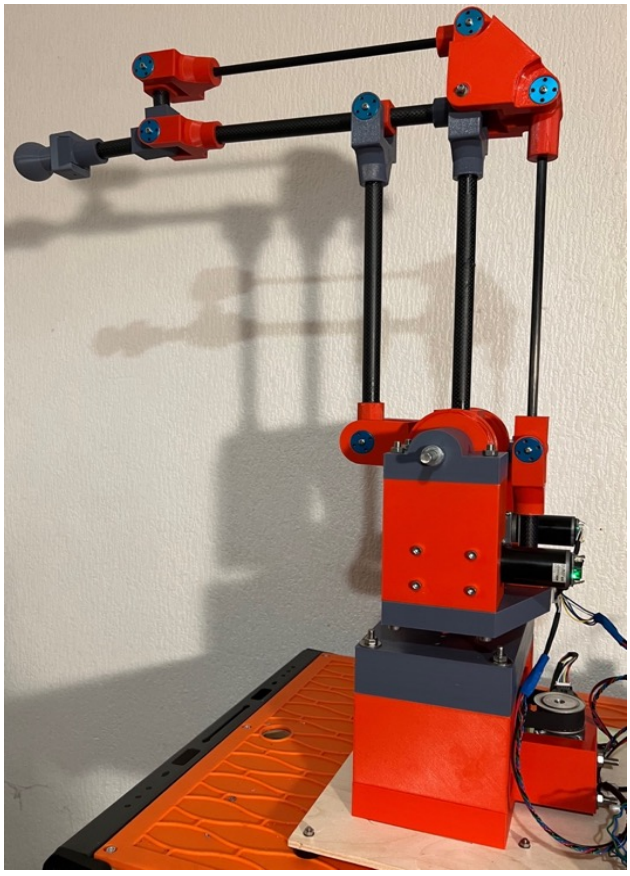


Figure 1: 3 DOF worm-gear motor robotic arm. The base mechanism serves as the initial rotational joint of the arm. Incorporating two degrees of freedom (2DOF), the upper segment adopts a parallelogram design.

The main shaft timing pulley on the collar is driven by a timing belt. A worm-gear motor, fitted with a smaller aluminum 24-tooth HTD 5M x 15mm timing pulley, drives the other end of the timing belt in order to rotate the shaft (Figure 2B). The motor itself is located on the lower base plate.

2.3. Two-degrees of freedom parallelogram design

The upper part of the arm mechanism is mounted on the main axis and is rotated by the base assembly, as depicted in Figure 3. Its design incorporates a mechanism to maintain the orientation of the endpoint in the plane of the upper arm, eliminating the need to control this linkage explicitly.

The main upper arm link L2 is supported and driven by a 3D-printed component that incorporates a 40-tooth HTD 5M x 15mm wide timing pulley. The component incorporates 2 x 8mm ball bearing races and is located on the main 8mm-diameter parallelogram mechanism shaft.

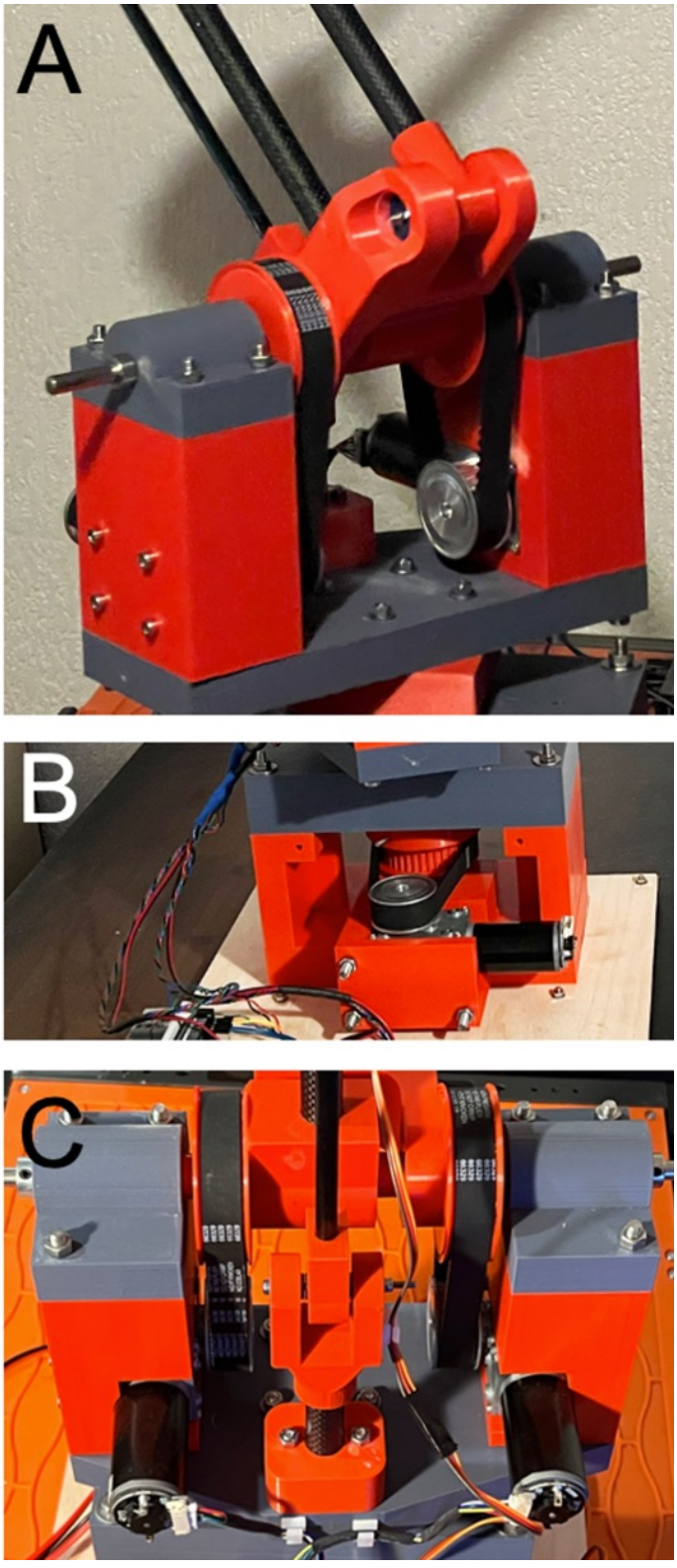


Figure 2. Close-up view of the rotating base platform mechanism. Panel A: Front view of the pulleys drive connected to the parallelogram arm mechanism. Panel B: The horizontal pulley drive for J1 for the first rotary axis of rotation. Panel C: Rear view of the parallelogram pulley drive showing worm-gear motors.

The secondary upper arm link arm is supported and driven by a 3D-printed mini-arm component, which also incorporates a 40-

tooth HTD 5M x 15mm wide timing pulley. This component also incorporates 2 x 8mm ball bearing races and is likewise located on the main 8mm-diameter parallelogram mechanism shaft.

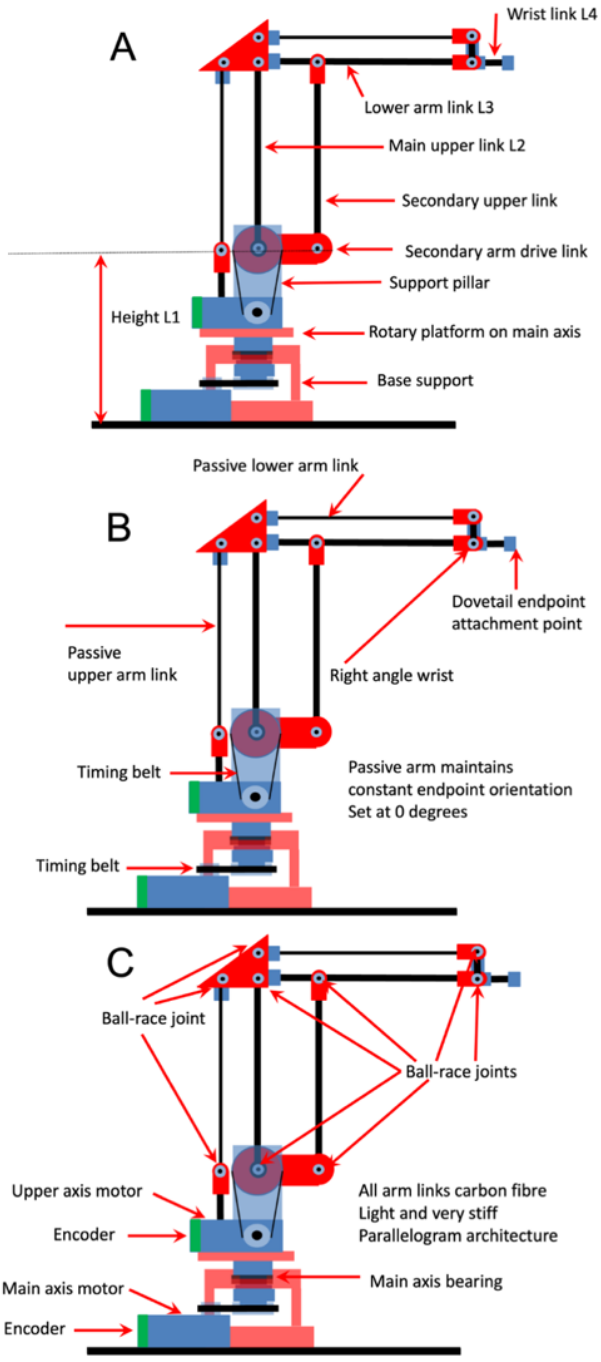


Figure 3: Robot viewed from the side. Panel A: Depicts the main components of the base and upper parallel arm links. Panel B: Shows the passive arm components and drive belts for the main arm. Panel C: Highlights the locations of the bearings, motors, and encoders.

The lower link L3 is driven by the main arm link L2 and secondary link L2 via custom made 3D-printed joint components (see Figure 3A).

To maintain the consistent orientation of the final arm link (wrist link L4), and consequently the endpoint, within the plane of the parallel arm mechanism, a passive orientation system is strategically situated behind the main arm. This ensures stable positioning and alignment throughout the operational movements of the arm mechanism. This additional parallelogram structure makes use of solid 8mm diameter carbon fiber rods in its construction. Figure 4A illustrates the endpoint alignment rocker mechanism, and the endpoint alignment mechanism and end-effector attachment point are shown in Figure 4B. This configuration enables the planar arm mechanism to preserve the endpoint orientation angle within the plane of the parallel arm without necessitating additional explicit actuation. As a result, the upper arm demands only two actuators for control, instead of three.

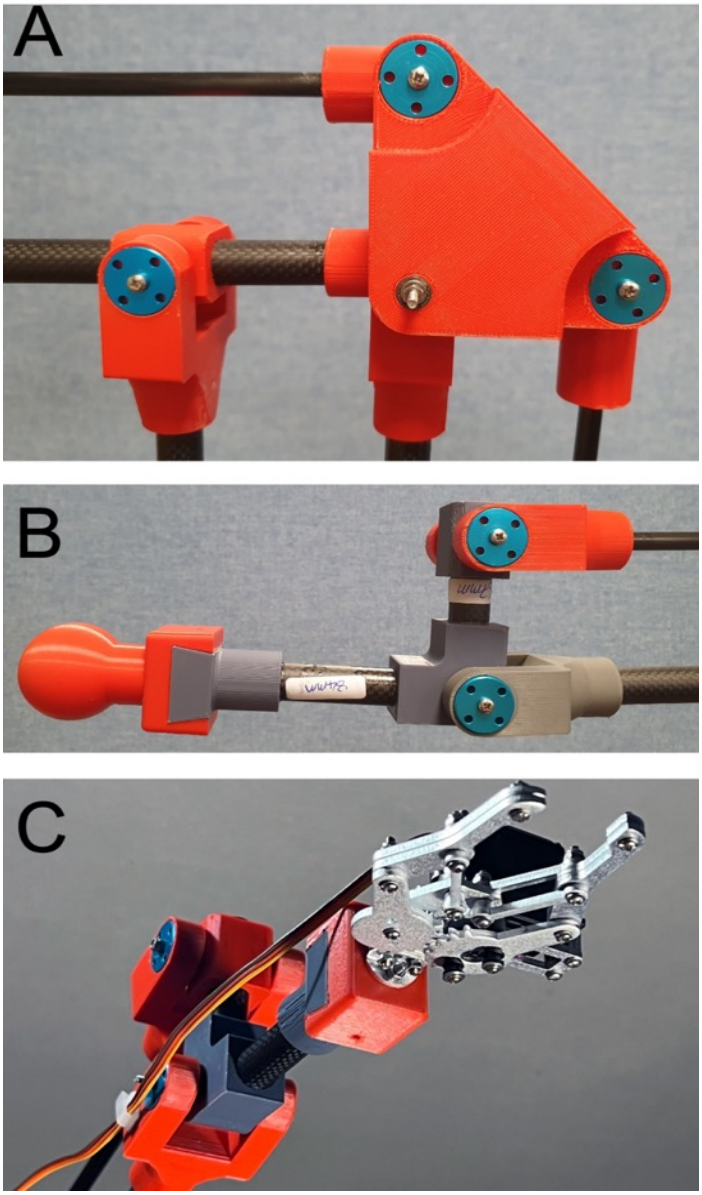


Figure 4: Panel A: The 2D upper arm features an endpoint alignment joint component located at the elbow. Panel B: The endpoint orientation mechanism. A dovetailed profile at the lower end of the last limb segment provides a connection point for end effectors. Here, a round knob is attached for safe demonstration purposes. Panel C: Simple commercially available RC-servo gripper mechanism attached to arm to demonstrate pick-and-place operations.

2.4. Arm joint components

The arm design relies on the use of joints, which utilize miniature ball-races to minimize friction. Figure 5 shows arm joint

components and their assembly is depicted in Figure 6. Each fork utilizes two lightweight aluminum flanges to support a shaft passing through the bearings, making use of the enhanced mechanical properties of aluminum compared to plastic. A long M3 screw is used as the joint shaft and is held in its support fork by means of a nylon insert hex locknut. Washers are strategically positioned on the screw shaft to achieve joint clearance between the joint bearing and the aluminum flanges.

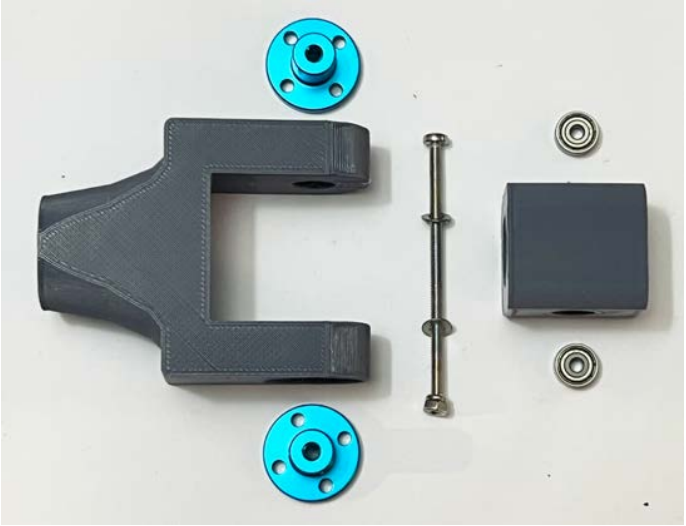


Figure 5: Components of the arm joints. The bearing component on the right-hand side hold miniature ball bearing races with a 3mm internal diameter. An M3 bolt is used as the joint shaft. When assembled, the joint bearing component locates between the two-sided support, illustrated on the left. This employs aluminum guides to hold the shaft. Washers are used to obtain suitable joint clearance.

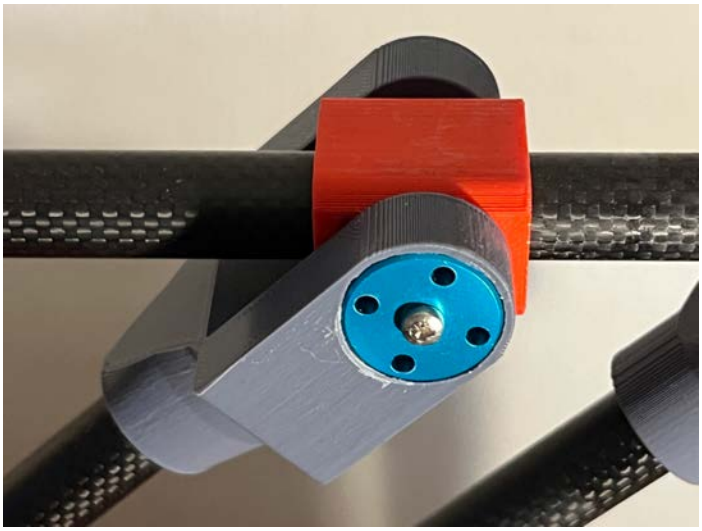


Figure 6: 2D Arm joint mechanism. Fundamental to the arm mechanism design, the arm joints utilize small ball bearings to ensure high precision and low friction operation.

2.5. Actuation using worm-drive motors

We utilize worm-gear motors for the arm's actuation. Opting for brushed worm-drive DC motors proves advantageous in constructing a robust servo drive system due to their wide availability, cost-effectiveness, and capacity to deliver high torque. Furthermore, the inherent non-back drivable characteristic of these drives guarantees the maintenance of a static joint configuration without necessitating continuous motor power [26].

The selected worm-gear motors run from a 24V supply, drawing up to 1A under full load. The worm gearing achieves a rotary speed reduction of 280, and the motor units achieve an unloaded output shaft speed of 28 RPM and an operational torque of 5Nm. These motors are equipped with Hall sensor encoders, located at the back, which deliver 11 pulses per revolution. This configuration allows the motor output shaft angle to be estimated with high accuracy and an output angular resolution of 3080 pulses per revolution was achieved. This figure can also be quadrupled further by utilizing the pulse edges.

As previously described, the motor output shafts are connected to the arm joints' drive axes utilizing 24-tooth HTD 5M x 15mm timing pulleys. The maximum angular speed of each driven joint, when using the selected worm-gear motors, is dictated by both the maximum motor rotational speed, converted from RPM to radians per second, and the mechanical advantage offered by the pulley drive. Consequently, given that all robot axis joints employ identical pulley ratios, the following value is obtained:

$$\omega_{max} = 28 * \left(\frac{24}{40}\right) * \left(\frac{2\pi}{60}\right) = 1.76\ Rads^{-1} \quad (1)$$

The maximum joint torque that can be achieved is given by the maximum continuous torque motor scaled up by the mechanical advantage of the pulley drive:

$$\tau_{max} = \left(\frac{40}{24}\right) * 5 = 8.33\ \text{Nm} \quad (2)$$

The robot arm incorporates three worm-drive motors in total. The main vertical revolute joint axis assembly, which integrates a 3D-printed timing pulley, is actuated by the first worm-gear-driven timing belt assembly. Moreover, two additional worm-drive motors mount within the pillars of the rotating vertical axis platform. The second joint motor engages a pulley to drive the main upper arm parallelogram arm link, while the third engages a timing pulley within the secondary arm drive link. Figure 2 illustrates the worm-drive motors, pulleys, and timing belts on the physical robot.

3. Gripper end effector for the arm

3.1. Gripper designs

Grippers are crucial end-effectors for robotic arms, enabling them to execute a myriad of tasks. Numerous designs have been proposed [27,28]. They are also an integral component of robotic harvesting technologies [29,30]. In addition to gripper designs that utilize hard materials, there is a significant contemporary interest in using soft materials for gripper construction [31,32].

3.2. RC servo operated gripper

Although pneumatic operation can offer a high-power density and a clean method of actuation for grippers, we chose to use RC servo actuation for easier integration into the current system. To demonstrate straightforward pick-and-place procedures with the arm, we affixed a low-cost, commercially available aluminum alloy RC-servo gripper mechanism to the arm via its dovetail connection point (refer to Figure 4C). This lightweight gripper assembly, weighing 160g, is capable of supporting simple pick-and-place tasks under direct human operator control.

We also designed and built a prototype RC servo-driven gripper with soft Ninja Flex cups at the ends of its fingers, specifically for assisting in gripping berries (refer to Figure 7). This gripper is lightweight, weighing 202g, and is attached to the arm through its dovetail connection point. We utilized this gripper to showcase the robotic arm's ability to detach berries from an artificial raspberry bush [33]. Refer to the results section for links to YouTube videos demonstrating the operation of both grippers.

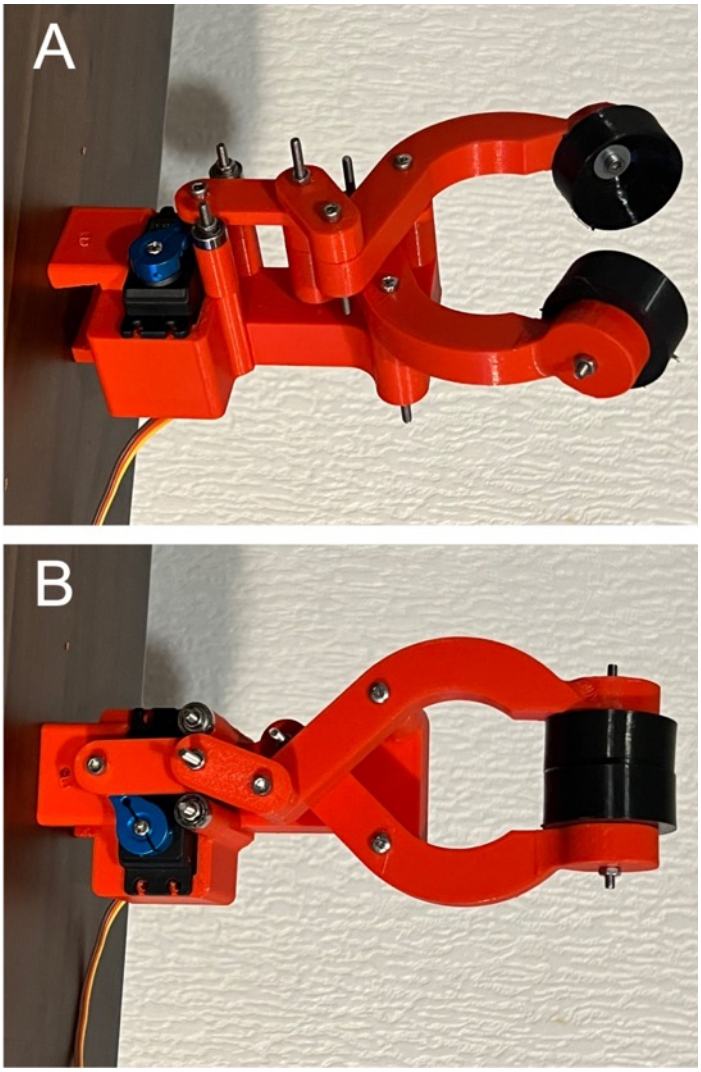


Figure 7. PLA+ 3D printed RC servo actuated gripper with soft Ninja flex endpoint cups. The gripper is shown in open (panel A) and closed (panel B) configurations.

4. Forward kinematics

4.1. Kinematic analysis of the arm

In robotic arms that operate with actuators capable of torque control, the examination of arm dynamics is an important and valuable exercise. In our design, we utilize worm-drive motors. Due to the high levels of friction introduced by this type of drive mechanism, which essentially results in a non-back-drivable system, worm drive motors are not conducive to building systems that effectively implement torque control. As a result, in our design, we do not employ torque control for the motors. Instead, we directly implement kinematic positional control of joint angles. For this reason, our analysis focuses on the forward and inverse kinematics of the mechanism as well as the Jacobian-based relationships between motor rotational velocity and endpoint velocity. Additionally, since the arm is lightweight and operates at a relatively slow speed, we also present an analysis of static force relationships using the Jacobian transpose.

4.2. Denavit-Hartenberg

Numerous textbooks on robotics offer kinematic analyses of arm structures [34–36]. For reader convenience, here we provide straightforward derivations.

Figure 8 illustrates the structure of the four links within the arm. While a straightforward derivation of the kinematics from the arm's geometry is possible, we have chosen to employ classical Denavit-Hartenberg (DH) analysis, a method often adopted in the field of robotics. Figure 9 displays the appropriate frames for DH analysis. The corresponding DH table for the arm are presented in Table 1 and the operating range limits of the robot joints are also included.

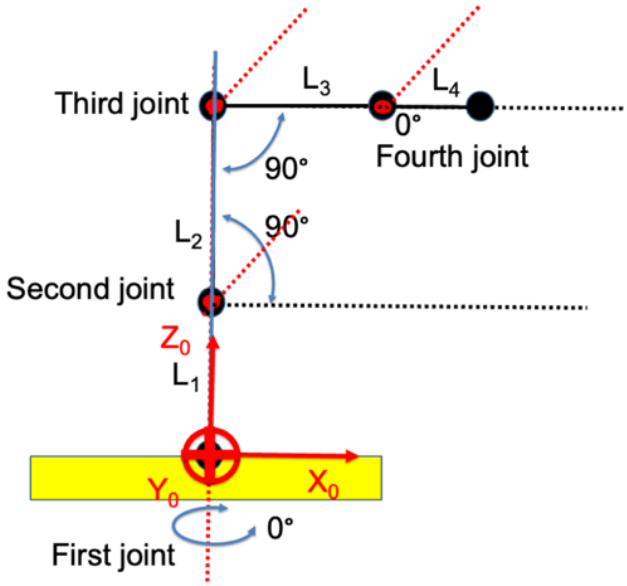


Figure 8: Default configuration.

Table 1: Denavit-Hartenberg parameters for the arm mechanism

Link	Angle θ [°]	Min θ [°]	Max θ [°]	Angle α [°]	Radius a [mm]	Offset d [mm]
1	θ_1	-45°	45°	0	0	300
2	θ_2	35°	145°	0	300	0
3	θ_3	-145°	-35°	0	300	0
4	θ_4	-110°	110°	0	175	0

4.3. Forward kinematics

The homogeneous transformation matrix for a classical DH-frame is given by the expression:

$$ {}^{i-1}_{i}H = \begin{bmatrix} c\theta_i & -s\theta_i c\propto_i & s\theta_i s\propto_i & a_i c\theta_i \\ s\theta_i & c\theta_i c\propto_i & -c\theta_i s\propto_i & a_i s\theta_i \\ 0 & s\propto_i & c\propto_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3) $$

where *'i'* is the frame, 'c' denotes cosine and 's' sine. We note that here we use the suffix notation for a homogeneous transformation H between link frames, indicating mapping *from the frame* identified by the i-suffix location *to the frame* identified in the i-1 suffix location, that is ${}^{to}_{from}H$.

In our case, the overall forward kinematic transformation for the arm is accomplished by multiplying the homogeneous matrices for the links together in the appropriate serial order, leading to the expression:

$$ {}^{0}_{4}H = {}^{0}_{1}H\ {}^{1}_{2}H\ {}^{2}_{3}H\ {}^{3}_{4}H \tag{4}$$

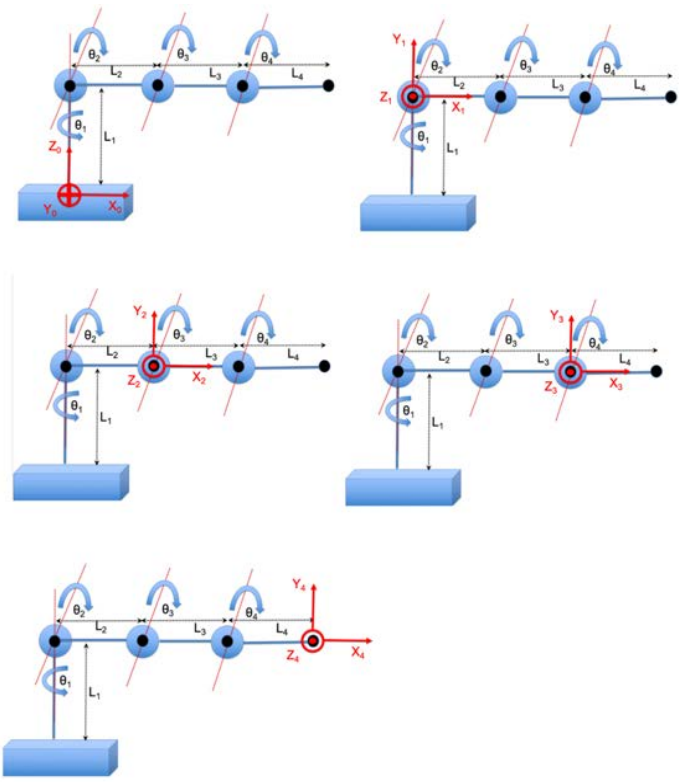


Figure 9. DH coordinate frames.

We utilize the MATLAB symbolic toolbox to evaluate the expression given in Eqn. (4), leading to the overall homogeneous transformation:

$$ {}^{0}_{4}H = \begin{bmatrix} c_{234}c_1 & -s_{234}c_1 & s_1 & c_1(L_3c_{23} + L_2c_2 + L_4c_{234}) \\ c_{234}s_1 & s_{234}s_1 & -c_1 & s_1(L_3c_{23} + L_2c_2 + L_4c_{234}) \\ s_{234} & c_{234} & 0 & L_1 + L_3s_{23} + L_2s_2 + L_4s_{234} \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5}$$

In this context, L_1, L_2, L_3, and L_4 denote the four link lengths, and we introduce the following notational simplifications for clarity: $c_1 = \cos(\theta_1)$, $s_1 = \sin(\theta_1)$, $c_{23} = \cos(\theta_2 + \theta_3)$, $s_{23} = \sin(\theta_2 + \theta_3)$, $c_{234} = \cos(\theta_2 + \theta_3 + \theta_4)$, $s_{234} = \sin(\theta_2 + \theta_3 + \theta_4)$.

In our arm design, the angle θ_4 is dependent on the final link orientation angle θ_{end} as well as joint angles θ_2 and θ_3. That is:

$$ \theta_{\text{end}} = \theta_2 + \theta_3 + \theta_4 \tag{6}$$

The expression below just provides the endpoint location of the final link, together with its orientation angle θ_{end}:

$$ \begin{bmatrix} x_{\text{end}} \\ y_{\text{end}} \\ z_{\text{end}} \\ \theta_{\text{end}} \end{bmatrix} = \begin{bmatrix} c_1(L_3c_{23} + L_2c_2 + L_4c_{\text{end}}) \\ s_1(L_3c_{23} + L_2c_2 + L_4c_{\text{end}}) \\ L_1 + L_3s_{23} + L_2s_2 + L_4s_{\text{end}} \\ \theta_2 + \theta_3 + \theta_4 \end{bmatrix} \tag{7}$$

We define a neutral arm posture, as illustrated in Figure 8. Configuration of the arm can be specified by control angles that deviate from this posture. Clearly, in this state, these angles Φ_1, Φ_2, and Φ_3 have values of zero. By inspection of Figure 8, it can be seen that Φ_1 follows θ_1 directly, whereas Φ_2 and Φ_3 are offset from θ_2 and θ_3. More precisely, we calculate the control angles as follows:

$$ \Phi_1 = \theta_1 \tag{8}$$

$$ \Phi_2 = \theta_2 - 90° \tag{9}$$

$$ \Phi_3 = \theta_3 + 90° \tag{10}$$

4.4. Workspace of the arm

The robotic arm's workspace is confined not only by the lengths of its links but also by the operational range of the angular joints, which are restricted due to inherent physical limits.

The primary vertical joint of the robot arm, denoted as J1, is theoretically capable of rotating from -180° to 180°. However, due to practical constraints arising from the wiring of the upper motors and encoders we chose to limit its rotation to a range between -45° and 45°. Joint J2, associated with the main upper arm link, is intrinsically able to rotate across a range from 0° to 180°. Nevertheless, physical limitations within the end-point orientation mechanism confine its rotational scope to between 35° and 145°.

Similarly, the rotation of joint J3, linked to the secondary arm drive, is constrained by the endpoint orientation mechanism, permitting rotation only within an approximate range of -35° to -145°.

To illustrate the arm's workspace, we first generate a test dataset of random configurations of the arm. To achieve this, for each arm joint angle, we independently draw 20,000 samples from a uniform distribution with lower and upper angular limits set by the operating range of the respective joint. We then use these 20,000 random angular configurations to calculate the arm end-point positions using forward kinematics, as given by Eqn. (7). The resulting arm end-points are then plotted in 3D and displayed in three different views for clarity, as illustrated in Figure 10.

5. Differential kinematic analysis

5.1. Exploring joint and endpoint velocity relationships

The 4x4 Jacobian matrix of the arm can be used to relate its end-point velocity to joint velocity. To formulate the Jacobian requires that we find the partial derivatives of its end point position and orientation given in Eqn. (7), with respect to the arm joint angles. The Jacobian is given by the matrix expression:

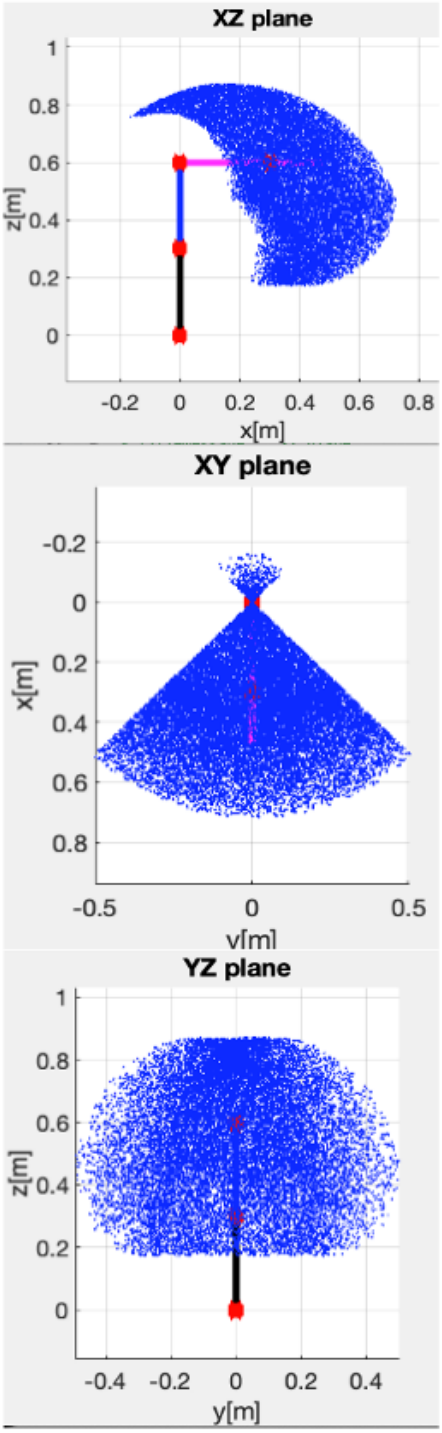


Figure 10. Workspace of the robot arm, also showing the arm displayed in its standard canonical configuration. Panel A: Side view showing the XZ axes. Panel B: Top view showing the XY axes. Panel C: Front view showing the YZ axes.

$$J = \begin{bmatrix} \frac{\partial x}{\partial \theta_1} & \frac{\partial x}{\partial \theta_2} & \frac{\partial x}{\partial \theta_3} & \frac{\partial x}{\partial \theta_4} \\ \frac{\partial y}{\partial \theta_1} & \frac{\partial y}{\partial \theta_2} & \frac{\partial y}{\partial \theta_3} & \frac{\partial y}{\partial \theta_4} \\ \frac{\partial z}{\partial \theta_1} & \frac{\partial z}{\partial \theta_2} & \frac{\partial z}{\partial \theta_3} & \frac{\partial z}{\partial \theta_4} \\ \frac{\partial \theta_{\text{end}}}{\partial \theta_1} & \frac{\partial \theta_{\text{end}}}{\partial \theta_2} & \frac{\partial \theta_{\text{end}}}{\partial \theta_3} & \frac{\partial \theta_{\text{end}}}{\partial \theta_4} \end{bmatrix} \quad (11)$$

This leads to the expression:

$$\begin{bmatrix} -s_1(L_3c_{23} + L_2c_2 + L_4c_{234}) & -c_1(L_3s_{23} + L_2s_2 + L_4s_{234}) & -c_1(L_3s_{23} + L_4s_{234}) & -c_1L_4s_{234} \\ c_1(L_3c_{23} + L_2c_2 + L_4c_{234}) & -s_1(L_3s_{23} + L_2s_2 + L_4s_{234}) & -s_1(L_3s_{23} + L_4s_{234}) & -s_1L_4s_{234} \\ 0 & L_3c_{23} + L_2c_2 + L_4c_{234} & L_3c_{23} + L_4c_{234} & L_4c_{234} \\ 0 & 1 & 1 & 1 \end{bmatrix} \quad (12)$$

The Jacobian relates the rotational velocities at the robot's joints to extrinsic velocity and rotational velocity of the endpoint through the simple expression:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ \dot{\theta}_{end} \end{bmatrix} = J \begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \\ \dot{\theta}_3 \\ \dot{\theta}_4 \end{bmatrix} \quad (13)$$

Examination of the Jacobian is informative regarding the singularities of the arm mechanism. Here we consider the case when the base rotation is zero. In this case the determinant of the Jacobian can be expressed as the relationship:

$$det(J) = -L_2L_3s_3(L_4 + L_3c_{23} + L_2c_2) \quad (14)$$

We observe that the determinant reaches zero when θ_3 is either 0 or π, occurring when the third link aligns with, or rotates backwards onto the second arm link. However, such configurations are not within the operating range of the arm. Consequently, on the basis of Eqn. (13), we can employ the inverse Jacobian to compute the joint velocities as follows:

$$\begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \\ \dot{\theta}_3 \\ \dot{\theta}_4 \end{bmatrix} = J^{-1} \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ \dot{\theta}_{end} \end{bmatrix} \quad (15)$$

5.2. Static force relationships

We can construct the 3×4 Jacobian by omitting the last row in Eqn. (12). Utilizing the transpose of the 3×4 Jacobian, we can establish a relationship between static joint torques - denoted $[\tau_1 \quad \tau_2 \quad \tau_3 \quad \tau_4]^T$ - and static endpoint forces - expressed as $[f_x \quad f_y \quad f_z]^T$ - through the following relationships:

$$\begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \end{bmatrix} = J^T \begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix} \quad (16)$$

6. Inverse kinematics

6.1. Main rotary axis

To determine the arm's inverse kinematics, we first define the endpoint location in the base frame coordinate system as the point (x_{end}, y_{end}, z_{end}). We note that in the base frame coordinate system, the arm mechanism's first joint constitutes a rotational axis around its z-axis, specified the angle θ_1. This rotation is the sole contributor to the end-effector's displacement along the base frame coordinate system's y-axis, as illustrated in Figure 11. Thus, determining θ_1 is relatively straightforward, given that:

$$\theta_1 = \text{atan}\left[\frac{y_{end}}{x_{end}}\right] \quad (17)$$

6.2. Planar arm mechanism

Our attention now shifts to the remaining planar arm mechanism. We define a 2D coordinate frame in the plane of the planar arm, situating its base at the origin (0,0), and defining the end effector location by the 2D point (x_{p3}, y_{p3}), as depicted in Figure 12.

Referring once more to Figure 12, we observe that the x-displacement, denoted as x_{p3}, is given by:

$$x_{p3} = \sqrt{x_{end}^2 + y_{end}^2} \quad (18)$$

We note that the direction of y-axis in the planar arm frame corresponds to the direction of z-axis in base frame. However, the 2D location of the end effector, denoted as y_3, must take the base height into account: Therefore, we have:

$$y_{p3} = z_{end} - L_1 \quad (19)$$

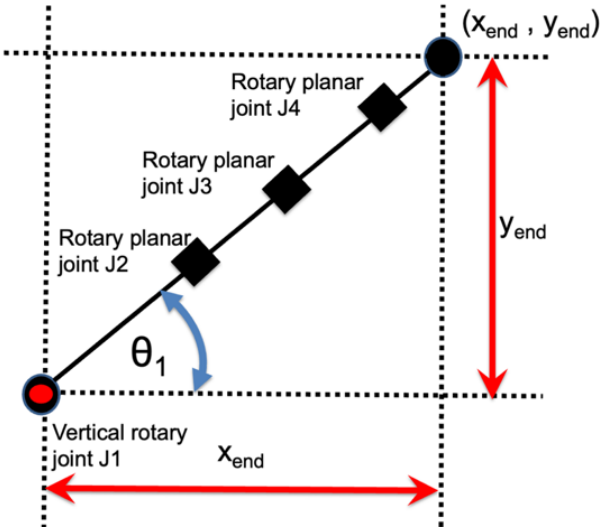


Figure 11: Top-down view of the robot arm: This perspective demonstrates the simplicity in calculating the first joint angle θ_1.

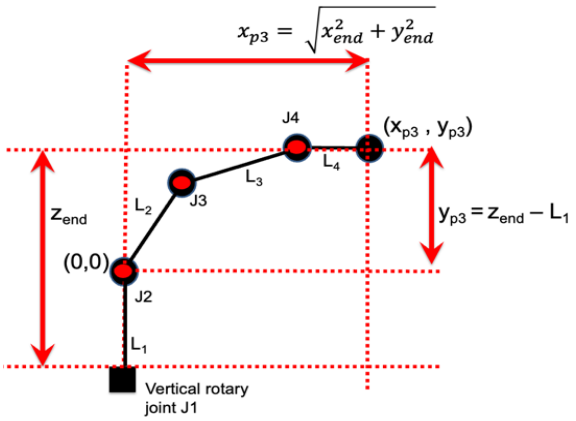


Figure 12: Side view of arm when main vertical axis joint angle is zero. The end-point of the arm in this coordinate frame is specified as the point (x_p, y_p), and its origin (0, 0) is located at J2.

We will shortly use the point (x_{p3}, y_{p3}) to calculate the 2D point (x_{p2}, y_{p2}), which we now consider obtaining the angles θ_2 and θ_3 of the remaining 2-link planar arm assembly. The cosine rule can be used to analytically derive this structure's inverse kinematics.

As depicted in Figure 13A for the 2-link planar arm mechanism, the elbow joint is situated at point (x_{p1} y_{p1}), while the upper arm joint is located at (x_{p2}, y_{p2}). Utilizing the Pythagorean theorem yields an expression for the length of the hypotenuse r:

$$r^2 = x_{p2}^2 + y_{p2}^2 \quad (20)$$

Employing the cosine rule, we see that:

$$r^2 = L_2^2 + L_3^2 - 2L_2L_3\cos(\alpha) \quad (21)$$

$$\Rightarrow \cos(\alpha) = \frac{L_2^2 + L_3^2 - x_{p2}^2 - y_{p2}^2}{2L_2 L_3} \quad (22)$$

From Figure 13A, it is noted that:

$$\theta_3 = \pi - \alpha \quad (23)$$

And from the trigonometric relationship, we have:

$$\cos(\pi - \alpha) = -\cos(\alpha) \quad (24)$$

$$\Rightarrow \cos(\theta_3) = \frac{x_{p2}^2 + y_{p2}^2 - L_2^2 - L_3^2}{2L_2 L_3} \quad (25)$$

From Figure 13B, we see:

$$\tan(\beta) = \frac{L_3\sin(\theta_3)}{L_3\cos(\theta_3) + L_2} \quad (26)$$

$$\Rightarrow \beta = \text{atan}\left[\frac{L_3\sin(\theta_3)}{L_3\cos(\theta_2) + L_2}\right] \quad (27)$$

From Figure 13C, we see that:

$$\theta_2 = \gamma - \beta \quad (28)$$

And that:

$$\gamma = \text{atan}\left[\frac{y_{p2}}{x_{p2}}\right] \quad (29)$$

Therefore, we have:

$$\theta_2 = \text{atan}\left[\frac{y_{p2}}{x_{p2}}\right] - \text{atan}\left[\frac{L_3\sin(\theta_3)}{L_3\cos(\theta_3) + L_2}\right] \quad (30)$$

To obtain a direct expression for the inverse kinematics of the arm assembly in terms of the point (x_{p3} y_{p3}), we first consider the constant orientation of the end link. In order to account for its effect, which is only a translation along the x-axis and none along the y-axis, we subtract its length from the end effector's position x_{p3} and directly use its y value:

$$\begin{bmatrix} x_{p2} \\ y_{p2} \end{bmatrix} = \begin{bmatrix} x_{p3} - L_4 \\ y_{p3} \end{bmatrix} \quad (31)$$

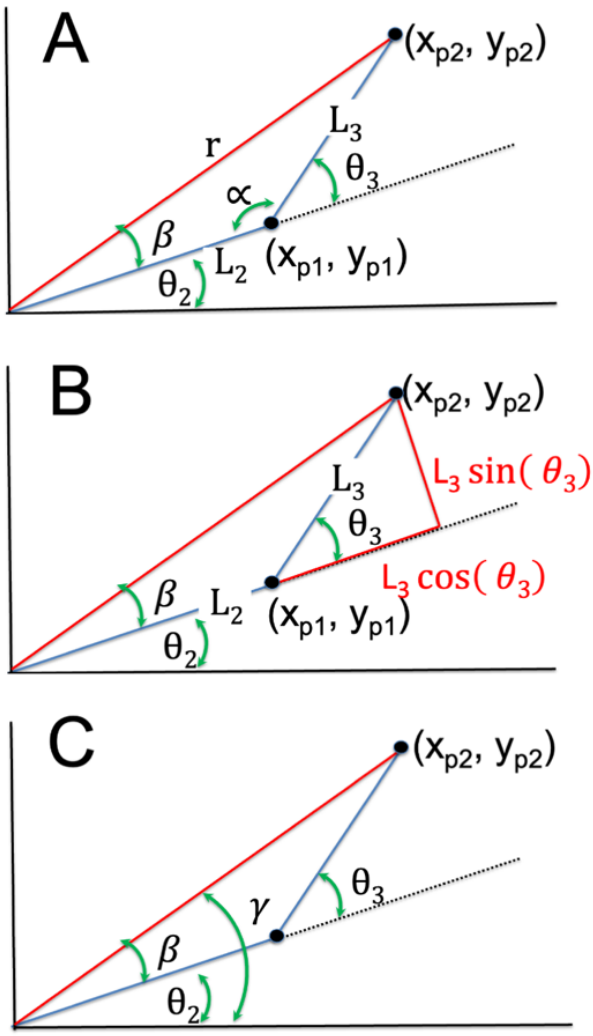


Figure 13: Schematic representation of the upper 2D parallelogram mechanism in the revolute arm: Useful for deriving inverse kinematics, the diagram illustrates the key positions, angles, and lengths.

Finally, we need to substitute the end effector location in as specified in the base coordinate frame, as defined by (x_{end}, y_{end}, z_{end}), and account for the height of the first vertical axis:

$$\begin{bmatrix} x_{p2} \\ y_{p2} \end{bmatrix} = \begin{bmatrix} \sqrt{x_{end}^2 + y_{end}^2} - L_4 \\ z_{end} - L_1 \end{bmatrix} \quad (32)$$

We substitute these values for (x_{p2}, y_{p2}) into Eqn. (30)

$$\Rightarrow \theta_2 = \operatorname{atan}\left[\frac{y_{end} - L_1}{\sqrt{x_{end}^2 + y_{end}^2} - L_4}\right] - \operatorname{atan}\left[\frac{L_3 \sin(\theta_3)}{L_3 \cos(\theta_3) + L_2}\right] \quad (33)$$

and into Eqn. (25)

$$\Rightarrow \theta_3 = \pm \operatorname{acos}\left[\frac{\left(\sqrt{x_{end}^2 + y_{end}^2} - L_4\right)^2 + (z_{end} - L_1)^2 - L_2^2 - L_3^2}{2L_2 L_3}\right] \quad (34)$$

We observe that Eqn. (34) yields two solutions and employs the value of θ_3 to compute θ_2 utilizing Eqn. (33). We opt for the negative solution for θ_3 as it aligns with the configuration illustrated in Figure 2. Lastly, we determine θ_4 using Eqn. (6):

$$\theta_4 = \theta_{end} - \theta_2 - \theta_3 \quad (35)$$

We note that this relationship was also used to calculate the minimum and maximum values of θ_4 shown in Table 1 on the basis of the minimum and maximum values of θ_2 and θ_3.

7. Control electronics

7.1. Microcontroller operation

A microcontroller system, built upon the Arduino Mega platform, was designed to operate the arm and power the worm-drive motors. This microcontroller offers an ample number of digital I/O ports with interrupt capability, facilitating operation with three incremental encoders located on the back of the motors through interrupt pins. This setup allowed for effective determination of the motor positions before the gear reduction via their worm drive mechanisms. To control the robot, a USB serial connection with the Arduino Mega was used.

7.2. Motor speed regulation using H-Bridges

To operate the DC motors used in the arm, we utilized two L298N Dual H-Bridges. These components allowed precise control over both the speed and direction of all three motors. Motor voltage is set using pulse-width modulation (PWM) activation of the on-time of the H-bridges. Drive power is provided from a 24V regulated power supply.

7.3. DIN rail controller implementation

All controller components were mounted on DIN rails, providing a simple and robust construction method. Please refer to Figure 14 to see its physical implementation. The 5V supply required to power the H-Bridge was provided by the Arduino Mega. In addition, an LM2596S DC-DC Buck converter module was attached to the DIN rail, directly driven from the 24V regulated power supply, to operate accessories such as the RC servo-operated gripper mechanism.

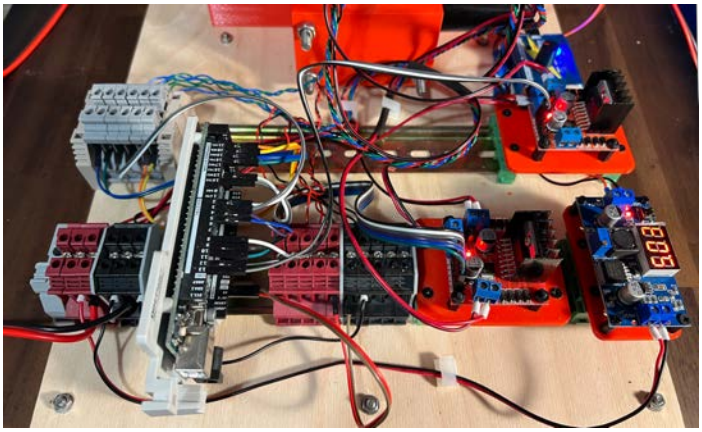


Figure 14: Controller electronics organized using DIN rail construction. The Arduino-based controller, terminal blocks, and H-bridges for the three worm drive motors can be seen mounted on the two DIN rails. A Buck controller is also shown and provides a 5V drive for the RC servo operated gripper mechanism.

7.4. Microcontroller program structure

Code was developed on the Arduino Mega to implement the kinematic control of the arm. The operation of the code consists of three distinct stages:

1. Initialization and Definitions: In this stage, variables and functions utilized by the algorithm are defined and initialized. Class objects are also constructed.
2. Setup: In this phase, operating variables are initialized and communications with the host PC are established.
3. Poll Loop: The poll loop is an endless processing loop that is used to service the program's needs. As well as implemented arm control, it must also service incoming messages from the host PC via a USB connection.

7.5. Microcontroller program menu

Within the poll loop, incoming commands from the serial interface are read and decoded. The available commands include:

- Manually set positive and negative rotation of individual drive motors.
- Reset encoders. This action needs to be carried out after the arm is positioned into its standard configuration, leading to calibration of the joint angles.
- Activate movement tracking along a predetermined trajectory.
- Activate point mode so can manually increment or decrement the robot's end-point along either the x, y or z axis in extrinsic space.
- Stop all movement.
- Display help menu and view parameters.

7.6. PID controller to track target trajectories

The arm is preliminary intended to carry out point-to-point movements for pick-and-place operations. For such tasks, the arm is typically given positional targets, or trajectories, to which it is required to reach. Inverse kinematics are then used to calculate the corresponding target motor angles. To accomplish this trajectory tracking task, a position control mode was implemented. When this mode is active, motor encoders are read to estimate the arm's joint angles, taking into considering the mechanical advantage of the worm gearing and also the drive pulley system. PID control is applied to ensure they match up with the requested joint control angles. In total there are three separate PID controllers, one for each motor. Output from the PID controllers constitutes velocity commands which make use of the H-Bridges to drive the motors.

Although the parallelogram joints J_1 and J_2 can be directly driven by the arm control angles Φ_1, Φ_2 , the third parallelogram joint J_3 must be driven by the sum of Φ_2 and Φ_3 to achieve the desired configuration of the parallelogram.

The Proportional-Integral-Derivative (PID) gains were determined empirically through experimental trials conducted on the actual robot. Notably, the integral gain was set to zero. In the next section we present the pseudo-code for the primary controller loop.

7.7. Main controller loop pseudocode

Pseudocode for the main poll loop running trajectory tracking is shown below. This makes use of function calls to calculate inverse kinematics and implement PID control and these functions are also described with pseudocode in subsequent sections.

Main Poll loop

Result: Moves arm along target trajectory
Perform Initialization of variables and trajectory
while *want tracking* **do**

Call the menu object for input commands and act accordingly

Calculate time interval 'h' since the last update

Read the desired target endpoint location (x, y, z)

Call Inverse Kinematics function to get target joint angles θ_1, θ_2 and θ_3

Compute corresponding arm control angles Φ_1, Φ_2 and Φ_3

Read the three motor actuator angles M_1, M_2 and M_3

Call PID function with control target angle Φ_1 and motor angle M_1 as input motor which returns motor velocity control U_1

Call PID function with control target angle Φ_2 and motor angle M_2 as input motor which returns motor velocity control U_2

Call PID function with control target angle $(\Phi_2 + \Phi_3)$ and motor angle M_3 as input motor which returns motor velocity control U_3

Generate PWM control velocities using Arduino for U_1, U_2 & U_3 and using H-bridges apply drive to the respective worm-drive motors

end

7.8. Inverse kinematics pseudocode

An essential aspect of the tracking process involves mapping the desired extrinsic position targets to the corresponding arm joint angles using inverse kinematics. We employ an implementation of inverse kinematics based on Eqns. (17, 33, 34), which is presented here as pseudocode in C++ style. The actual inverse kinematics code is written in C++ and makes use of a constructor to set up the link lengths L_1, L_2, L_3. L_4 and the endpoint angle θ_{end} in member variables. The inverse kinematics pseudocode is as follows:

Inverse Kinematics

Input: Desired target end point location (x_{end}, y_{end}, z_{end})
Link lengths: L_1, L_2, L_3, L_4
Endpoint angle: θ_{end}
Result: Target joint angles θ_1, θ_2 , θ_3, θ_4
Return: IK success or fail flag

Function call

Calculate $\theta_1 = \text{atan2}(y_{end}, x_{end})$

Calculate x-axis of planar arm: $x_{p3} = \text{sqrt}\,(x_{end}^2 + y_{end}^2)$

Account for end link: $x_{p2} = x_{p3} - L_4$

Account for base height of planar arm: $y_{p2} = z_{end} - L_1$

Calculate $\cos(\theta_3)$: $c_3 = (x_{p2}^2 + y_{p2}^2 - L_2^2 - L_3^2) / (2 * L_2 * L_3)$

Calculate $\sin(\theta_3)$: $s_3 = \text{sqrt}(1 - c_3^2))$

Check if can reach target: if $((1 - c_3^2) < 0)$ return **false**

Calculate $\theta_3 = -\text{atan2}(s_3, c_3)$

Calculate variable $k_1 = L_2 + L_3 * c_3$

Calculate variable $k_2 = -L_3 * s_3$

Calculate $\theta_2 = \text{atan2}(y_{p2}, x_{p2}) - \text{atan2}(k_2, k_1)$

Calculate $\theta_4 = \theta_{end} - \theta_2 - \theta_3$

Succeeded reaching target so: return **true**

end

7.9. PID block pseudocode

The PID function control code is written in C++. This code includes a constructor for setting up the gain parameters as member variables, a reset function which is needed to initialize the last time and last error member variables, and to reset the error integral. It also includes and a PID error block that calculates the control signal based on the target and measured joint angles. Pseudocode is for the computation of the control U is shown below.

PID Control

Input: Desired control target angle Φ
Measured motor angle θ_M
PID gains: k_p, k_I, k_d
Last error: e_{last}
Integral of error: e_i
Current and last times $T_{current}$, T_{last}
Result: Motor velocity control **U**

Function call

Calculate proportional error: $e_p = \theta_M - \Phi$

Calculate time step: $h = T_{current} - T_{last}$

Update last time: $T_{last} = T_{current}$

Calculate error derivative: $\frac{de}{dt} = (e_p - e_p)/h$

Calculate error integral: $e_i = e_p + e_i$

Calculate PID control: $U = e_p * k_p + e_i * k_i + \frac{de}{dt} * k_d$

end

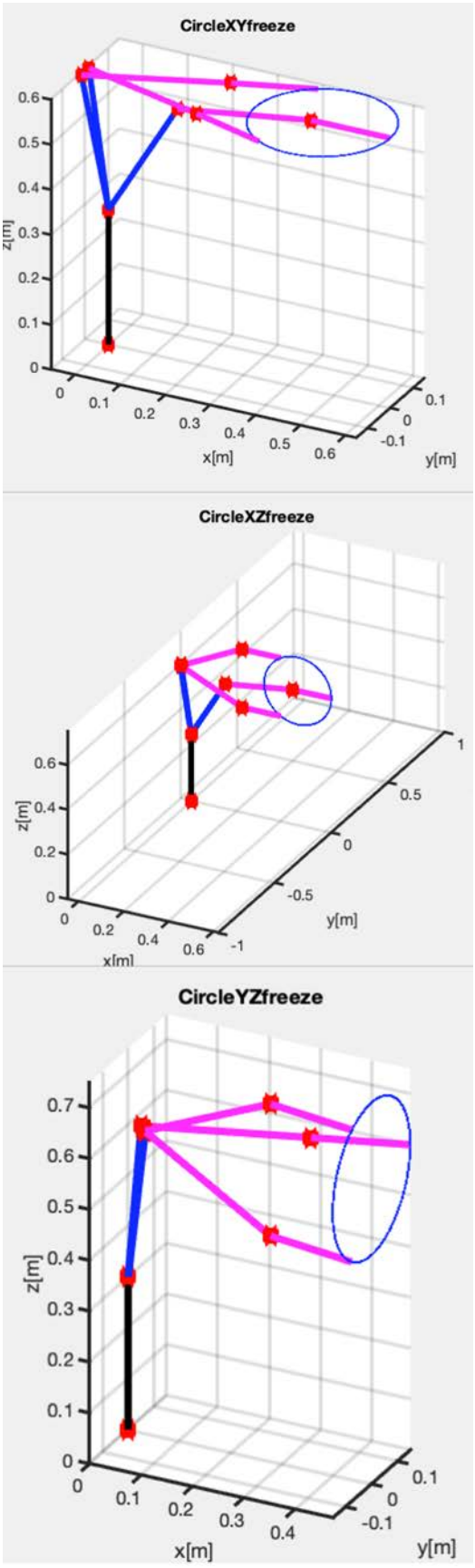


Figure 15. Arm configurations to selected targets. This visualization demonstrates how the arm accesses three distinct points on target circles within the XY, XZ, and YZ planes. The arm joint angles are computed via inverse kinematics, while the arm joint positions are computed through the application of forward kinematics

8. Arm operation

8.1. MATLAB Simulation

To test and showcase the inverse kinematics, we implemented a set of simple tracking tasks in MATLAB. We generated three distinct trajectories. Each of these trajectories is composed of 500 points, uniformly distributed around the perimeters of circles in the XY, XZ, and YZ planes. These circles, with their centers positioned at the (x, y, z) coordinates of (0.475, 0.0, 0.60) meters and a radius of 0.15 meters, comfortably fit within the robotic arm's operational workspace. Utilizing inverse kinematics, we computed the necessary arm angles to target specific points on these circles. Figure 15 visually presents both the desired circles and the requisite arm postures to access three chosen points from each circle.

We made use of Eqn. (15), to calculate motor shafts angular velocities corresponding to the endpoint of the arm moving around the circle every 2 seconds. Figure 16 illustrates this, revealing it corresponds to an endpoint speed of 0.47 m/s. The peak joint speed was identified to be 1.64 Rad/s, which is notably beneath the maximum speed the motors are capable of achieving, 1.77 Rad/s, as denoted by Eqn. (1). It's noteworthy that pursuing the target circle with the arm's endpoint through inverse kinematics and calculating velocities via finite differences deliver consistent results.

Using Eqn. (16) and discounting the impact of the arm's inherent dynamics (a simplification that holds true primarily at low movement velocities), we computed the torques at the motor shafts. These torques are required to sustain a realistic payload of 1kg while applying a force of 1Kgf in the negative x-direction. This is for endpoint positions encircling the target circle in the YZ plane. Figure 17 reveals a peak joint torque of approximately 8.56Nm. Notably, this value marginally surpasses the designated joint torque the motors can produce, which is 8.33Nm as referenced in Eqn. (2). However, this is attainable during intermittent operation.

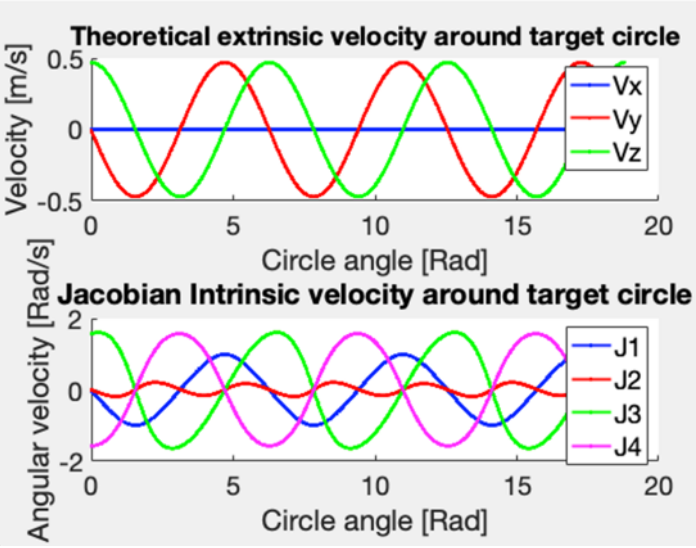


Figure 16. Arm velocity relationships: The arm's endpoint navigates around a target circle with a 0.15 m radius, executing three full rotations within a 10-second time frame. The center of the circle coincides with the arm's endpoint in its canonical position. Panel A: Extrinsic endpoint velocities of the arm. Panel B: Angular velocities of the arm joints

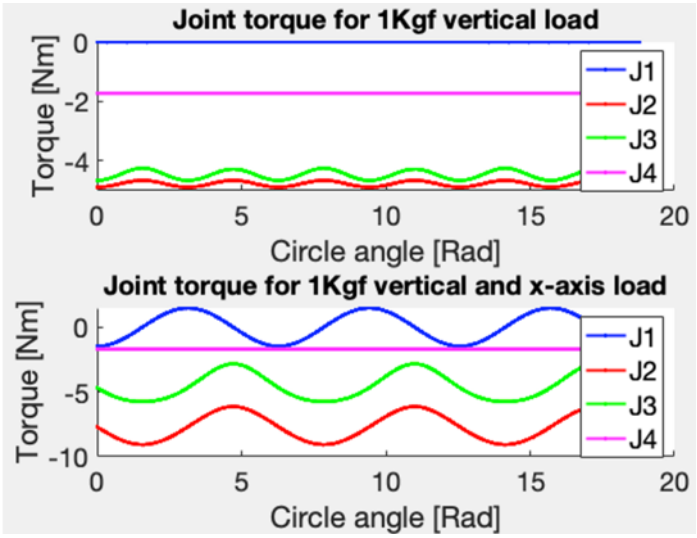


Figure 17. Simulated joint torques under combined loading conditions: Effect of simultaneous 1 Kgf horizontal and 1 Kgf vertical loads. The scenario involves the endpoint moving around a target circle with a 0.15 m radius, the center of which aligns with the endpoint of the arm in its canonical position.

8.2. Trajectory tracking using the physical robotic arm

Demonstrations of tracking around XY, XZ and YZ circles were performed using the actual physical robotic arm.

To implement tracking around square and circular trajectories, firstly target 3D trajectories were generated offline using the microcontroller and held in an array. During the trajectory tracking operation, target points were read out from the array at the appropriate time-index, and their values converted to target joint angles using the microcontroller's implementation of the arm's inverse kinematics.

A Proportional-Derivative (PD, i.e., PID with a zero integral gain component) controller implemented on the microcontroller was then used to ensure the motors were appropriately driven so that the target joint angles were tracked.

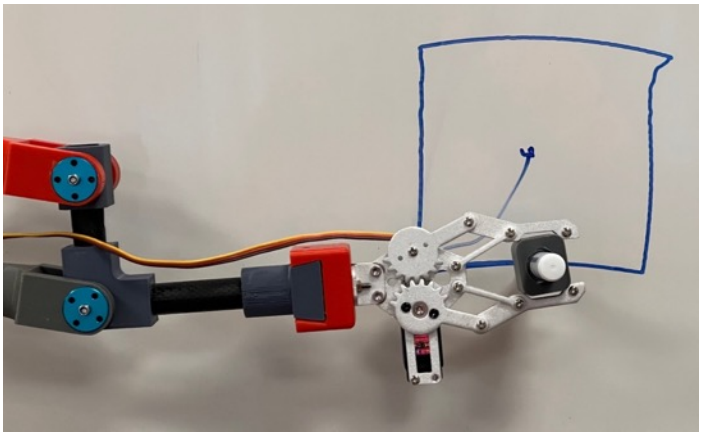


Figure 18: Arm using a low-cost commercially available gripper to hold a felt-tip marker pen and draw on a whiteboard.

8.3. Tracking using the physical robotic arm

For demonstration purposes, a simple commercial gripper was affixed to the arm's endpoint, enabling it to grasp and subsequently release objects. Tracking around a square while holding a pen and drawing on a whiteboard is depicted in the photograph shown in

Figure 18 (refer also to section 8.5 for online YouTube videos). It's worth noting there is some overshoot in the upper right-hand corner due to friction caused by the arm pressing against the whiteboard before suddenly slipping. Future improvements to the controller should address this issue.

8.4. Pick and place operation using the physical robotic arm

We also demonstrate the pick-and-place teleoperation of the arm through a berry-picking task. This task entailed the remote control of the arm to harvest artificial berries from an artificial bush. Both the berries and the bush were designed for a study that evaluated gripper mechanisms under direct human operator control [33].

The teleoperation is implemented as a variant of the previous trajectory tracking task, utilizing inverse kinematics and PID control to appropriately drive the actuating worm-gear motors. However, in this instance, rather than tracking a pre-defined trajectory, keyboard commands were used in real time to increment or decrement a stationary target position in extrinsic space. In this manner, an operator could direct the endpoint of the arm to the desired location in space. Opening and closing the gripper was implemented using an RC servo tester unit.

8.5. YouTube demonstrations of arm simulations and operation

Video demonstrations of these real robot tasks are available on our YouTube channel, "Robotics, Control, and Machine Learning," within the playlist “Design and prototyping of a 3DOF worm-drive robot arm.” The playlist is accessible directly via the following link:

https://youtube.com/playlist?list=PLjKvJX8cBCKWBZDUbaPWgSSMs_msS1MA0.

9. Discussion

9.1. Summary

In this project, we engineered a lightweight robotic arm mechanism, capable of operating in a three-dimensional workspace, primarily designed with the agricultural industry in mind. However, it could also find applications in various other industries, as well as providing a platform suitable for education in robotics.

The integration of a two-dimensional planar mechanism, affixed to a vertical rotary axis, achieves comprehensive 3D spatial coverage. The parallelogram design takes inspiration from the vBOT robotic manipulandum, a research instrument utilized to explore human sensorimotor control in arm movements [37].

We opted for worm-drive actuation to ensure that static postures could be sustained without necessitating an active motor drive, thereby mitigating power consumption. By leveraging off-the-shelf and light-weight structural components, the design not only becomes cost-effective but due to its minimal moving mass, also remains inherently safe particularly when operated in close proximity to individuals. Overall, this culminates in a design that is easy to fabricate, low-cost, lightweight, and is power-efficient.

9.2. Extension of previous work

The current work extended a previous prototype design in several important ways. An additional vertical rotary axis is added, enabling the original planar arm mechanism to rotate and thus operate in a 3D workspace. The kinematic analysis is expanded, and forward, inverse, and differential kinematics for the new 3D arm structure are derived. Furthermore, the former planar arm support structure has been fully redesigned, and we now incorporate 3D-printed pulleys on the arm itself, replacing the previous cumbersome aluminum pulleys. Finally, in the results section, we present evidence of the arm's point-to-point operation, including pick-and-place tasks, demonstrating its ability to perform practical real-world activities.

9.3. Novel contribution of the work

The novel aspects of the design arise from combining the use of 3D printing, carbon fiber sections for arm fabrication, a parallelogram planar upper arm structure that passively maintains endpoint orientation, and the use of worm gear motors.

The key features of the arm, as well as its similarities and differences with the previous prototype arm presented in [1], are summarized in Table 2.

9.4. Bill of materials and cost of the arm

The arm design incorporates numerous low-cost components, resulting in a construction that can be assembled relatively inexpensively. Table 3 presents a bill of the main materials for the project, indicating the costs of individual components as listed by online suppliers for small quantities.

The total cost of constructing a single unit (in 2023) is approximately £480, although discounts would likely apply for larger quantities. While assembly costs have not been included in this calculation, the arm, can easily be put together within a day by a single person. This assumes that the 3D printed custom parts have already been printed and are available for assembly.

Clearly, a construction cost of less than £500 makes the design competitive against many commercial arms, as well as those targeted more towards the hobby market. The former typically start at several thousand pounds per unit, while the latter are still on the order of £1000.

Table 2: Features of the new arm design highlighting differences with the previous 2D prototype

ARM FEATURE	NEW ARM	PROTOTYPE ARM
SIMILARITIES IN DESIGN		
Worm-drive motor actuation	Yes	Yes
Low weight carbon fiber upper limb construction	Yes	Yes
Passive end-link horizontal orientation mechanism	Yes	Yes
Construction makes use of low-cost off-the-shelf components wherever possible	Yes	Yes
Construction technique employs 3D printing for custom parts	Yes	Yes

DIFFERENCES IN DESIGN		
Linear endpoint degrees of freedom (DOFs) of the arm in extrinsic space	3	2
Arm configuration	Revolute arm mechanism	Planar 2D mechanism
Dovetail end-effector attachment point	Yes	No
End-effector accessories	The new design integrates the use of both a commercial and a new bespoke gripper mechanism	No accessories
Demonstration of pick-and-place operation	Yes	No
Timing belt drive mechanism	Driven from custom 3D printed pillars and baseplate	Driven from baseplate
Forward kinematic analysis and implementation	FK derived and implemented for 4-link 3DOF arm	FK derived and implemented for 3-link 2DOF arm
Inverse kinematic analysis and microcontroller implementation	IK for 4-link 3DOF arm	IK for 3-link 2DOF arm

Table 3: Main arm mechanism bill of materials indicating component costs in (as of 2023)

DESCRIPTION	UNIT COST	NUMBER	TOTAL
DRIVE MECHANISM COMPONENTS			
Worm-drive motor DC 24V 28RPM with Hall encoder	£35	3	£105
50mm I/D 90mm O/D 20mm Thick deep-groove bearings	£10	1	£10
HTD 5M 24 tooth pulley	£12	3	£36
HTD 345 5M belt	£16	3	£48
ARM STRUCTURAL COMPONENTS			
CF tube 16x14 x250mm	£12	3	£36
CF tube 16x14 x100mm	£6	1	£6
CF rod 8mmx250mm	£12	2	£24
8mm 200mm stainless shaft	£5	1	£5
Aluminum flange 3mm I/D x H13 x D10	£2	14	£28
M3 x 95mm bolt	£1	8	£8
M6 x 150mm bolt	£2.50	6	£15
M6 x 170mm bolt	£2.50	4	£10
8mm I/D x 22mm O/D x 7mm thick bearing	£2	2	£2
Miniature bearings I/D 3mm x 10mm O/D x 4mm	£0.8	16	£13
Shaft collars 8mm	£1.5	2	£3
ARM CONTROLLER COMPONENTS			
Arduino Mega clone	£16	1	£16
Dual H-Bridge Motor Drive Controller Board L298N	£2.5	2	£5
Buck converter to power RC servo driven accessories	£2.5	1	£2.5
DIN rails 300mm	£3	2	£6
DIN clips	£2	6	£12
DIN rail terminals	£1	22	£22
Arduino DIN rail Mount	£8	1	£8
3D PRINTING MATERIAL			
1Kg roll PLA+	£20	3	£60
TOTAL SUM			£ 478

9.5. Simple upgrades to the current design

In our current design, the fourth end-link consistently maintains a 0° orientation angle, resulting in a translation of the overall endpoint along the horizontal x-axis. However, alternative fixed endpoint orientations could be achieved, by adapting the design of the wrist component.

We utilized parts fabricated from PLA+ through 3D printing. While this material is a suitable choice for prototyping due to its printing ease, more sturdy and environmentally resilient materials could be seamlessly substituted for actual deployment, such as in fruit-picking applications. Potential alternatives might encompass the use of PETG, ABS, nylon, or even composite materials [33, 34].

In line with the approach adopted by an earlier prototype [1], incorporating an efficient controller that estimates motor torque by monitoring motor current would be a simple task. This feature would enable the incorporation of a safety mechanism into the controller, devised to identify unforeseen collisions. Motor current could be compared with anticipated values for a specific task, enabling automatic deactivation of the control if these limits were surpassed.

9.6. Future improvements

The current arm system employs a simple controller based on the Arduino Mega and L298N Dual H-bridges. While this cost-effective combination displays reasonable performance and is suitable for the current proof-of-concept arm design, it does exhibit significant shortcomings. The Arduino Mega's default capability to generate a PWM operating frequency for controlling three

outputs is restricted, achieving only a relatively low operating frequency of 490 Hz. Additionally, the microcontroller's clock speed is quite slow (16 MHz), limiting the control loop update rate to around 100 Hz. Future developments of the arm will involve creating a higher-performance controller, leveraging faster, cost-effective microcontrollers such as the ESP32.

Current measurement of endpoint position using motor encoders cannot account for potential deformations or flexibilities inherent in the arm's structure, or the slight backlash in the worm gear motors. Integrating additional angular position sensors on the arm's joint axes, e.g., using low-cost magnetic encoders, would improve the estimation of endpoint position which could be used to improve arm operational accuracy.

Direct force sensing using force transducers would provide an effective method for the estimation of endpoint force. Although commercial industrial force transducers are prohibitively expensive, research suggests that 3D-printed sensors may provide a suitable low-cost solution to this issue [38].

Finally, it is noteworthy that in the design described here, direct-drive brushless DC motors could serve as a substitute for the current method of actuation and support force and torque control performance [39]. However, this would also incur higher motor costs and escalate power consumption, particularly during the maintenance of static arm configuration.

Conflict of Interest

The author confirms there are no conflicts of interest to disclose.

Acknowledgements

We thank Enterprise Solutions at the University of Plymouth for their financial support of this work on a project entitled, "Low-cost Compliant Robotic Arm for Fruit and Vegetable Harvesting". We also thank our industrial partner, Fieldwork Robotics Ltd., and in particular, Martin Stoelen of Western Norway University of Applied Sciences, for their invaluable support and insightful discussions.

References

[1] I.S. Howard, "Design and prototyping of a low-cost light weight fixed-endpoint orientation planar Cobot," in 2022 International Conference on System Science and Engineering (ICSSE), IEEE: 47–54, 2022, doi:10.1109/ICSSE55923.2022.9947353.

[2] T. Duckett, S. Pearson, S. Blackmore, B. Grieve, W.-H. Chen, G. Cielniak, J. Cleaversmith, J. Dai, S. Davis, C. Fox, "Agricultural Robotics: The Future of Robotic Agriculture," UK-RAS Network, **UK-RAS White Papers**, 2018, doi:https://www.ukras.org.uk/wp-content/uploads/2021/01/UKRASWP_AgriculturalRobotics2018_online.pdf.

[3] A. Bechar, C. Vigneault, "Agricultural robots for field operations: Concepts and components," Biosystems Engineering, **149**, 94–111, 2016, doi:https://doi.org/10.1016/j.biosystemseng.2016.06.014.

[4] F. Chaumette, S. Hutchinson, P. Corke, "Visual servoing," Springer Handbook of Robotics, 841–866, 2016, doi:https://doi.org/10.1007/978-3-319-32552-1_34.

[5] D.F. Kutz, A. Wölfel, T. Meindl, D. Timmann, F.P. Kolb, "Spatio-temporal human grip force analysis via sensor arrays," Sensors, **9**(8), 6330–6345, 2009, doi:10.3390/s90806330.

[6] B. Jia, A. Zhu, S.X. Yang, G.S. Mittal, "Integrated gripper and cutter in a mobile robotic system for harvesting greenhouse products," in 2009 IEEE International Conference on Robotics and Biomimetics (ROBIO), IEEE: 1778–1783, 2009, doi:10.1109/ROBIO.2009.5420430.

[7] S. Gorjian, H. Ebadi, M. Trommsdorff, H. Sharon, M. Demant, S. Schindele, "The advent of modern solar-powered electric agricultural machinery: A solution for sustainable farm operations," Journal of Cleaner Production, **292**, 126030, 2021, doi:https://doi.org/10.1016/j.jclepro.2021.126030.

[8] I.O. for Standardization, "Robots for industrial environments–safety requirements–Part 1: robot, ISO10218-1: 2006," 2006.

[9] B. Vanderborght, A. Albu-Schäffer, A. Bicchi, E. Burdet, D.G. Caldwell, R. Carloni, M. Catalano, O. Eiberger, W. Friedl, G. Ganesh, "Variable impedance actuators: A review," Robotics and Autonomous Systems, **61**(12), 1601–1614, 2013, doi:https://doi.org/10.1016/j.robot.2013.06.009.

[10] G. Grioli, S. Wolf, M. Garabini, M. Catalano, E. Burdet, D. Caldwell, R. Carloni, W. Friedl, M. Grebenstein, M. Laffranchi, "Variable stiffness actuators: The user's point of view," The International Journal of Robotics Research, **34**(6), 727–743, 2015, doi:https://doi.org/10.1177/027836491456651.

[11] M.F. Stoelen, R. de Azambuja, B. López Rodríguez, F. Bonsignorio, A. Cangelosi, "The GummiArm Project: A Replicable and Variable-Stiffness Robot Arm for Experiments on Embodied AI," Frontiers in Neurorobotics, **16**, 24, 2022, doi:https://doi.org/10.3389/fnbot.2022.836772.

[12] M.F. Stoelen, F. Bonsignorio, A. Cangelosi, "Co-exploring actuator antagonism and bio-inspired control in a printable robot arm," in From Animals to Animats 14: 14th International Conference on Simulation of Adaptive Behavior, SAB 2016, Aberystwyth, UK, August 23-26, 2016, Proceedings 14, Springer: 244–255, 2016, doi:https://doi.org/10.1007/978-3-319-43488-9_22.

[13] A. De Luca, A. Albu-Schaffer, S. Haddadin, G. Hirzinger, "Collision detection and safe reaction with the DLR-III lightweight manipulator arm," in 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE: 1623–1630, 2006, doi:10.1109/IROS.2006.282053.

[14] S. Haddadin, A. Albu-Schäffer, G. Hirzinger, "Safety evaluation of physical human-robot interaction via crash-testing," in Robotics: Science and systems, Citeseer: 217–224, 2007, doi:10.15607/RSS.2007.III.028.

[15] A. Albu-Schäffer, S. Haddadin, C. Ott, A. Stemmer, T. Wimböck, G. Hirzinger, "The DLR lightweight robot: design and control concepts for robots in human environments," Industrial Robot: An International Journal, **34**(5), 376–385, 2007, doi:https://doi.org/10.1108/01439910710774386.

[16] J.E. Colgate, W. Wannasuphoprasit, M.A. Peshkin, "Cobots: Robots for collaboration with human operators," in Proceedings of the 1996 ASME international mechanical engineering congress and exposition, 1996, doi:https://doi.org/10.1115/IMECE1996-0367.

[17] W.T. Townsend, J.K. Salisbury, "Mechanical design for whole-arm manipulation," in Robots and Biological Systems: Towards a New Bionics?, Springer: 153–164, 1993, doi:https://doi.org/10.1007/978-3-642-58069-7_9.

[18] K. Bodie, C.D. Bellicoso, M. Hutter, "ANYpulator: Design and control of a safe robotic arm," in 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE: 1119–1125, 2016, doi:10.1109/IROS.2016.7759189.

[19] S. Haddadin, S. Parusel, L. Johannsmeier, S. Golz, S. Gabl, F. Walch, M. Sabaghian, C. Jähne, L. Hausperger, S. Haddadin, "The franka emika robot: A reference platform for robotics research and education," IEEE Robotics & Automation Magazine, **29**(2), 46–64, 2022, doi:10.1109/MRA.2021.3138382.

[20] S. Mick, M. Lapeyre, P. Rouanet, C. Halgand, J. Benois-Pineau, F. Paclet, D. Cattaert, P.-Y. Oudeyer, A. de Rugy, "Reachy, a 3D-printed human-like robotic arm as a testbed for human-robot control strategies," Frontiers in Neurorobotics, **13**, 65, 2019, doi:https://doi.org/10.3389/fnbot.2019.00065.

[21] T. Lens, J. Kunz, O. Von Stryk, C. Trommer, A. Karguth, "Biorob-arm: A quickly deployable and intrinsically safe, light-weight robot arm for service robotics applications," in ISR 2010 (41st International Symposium on Robotics) and ROBOTIK 2010 (6th German Conference on Robotics), VDE: 1–6, 2010.

[22] G.A. Pratt, M.M. Williamson, "Series elastic actuators," in Proceedings 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems. Human Robot Interaction and Cooperative Robots, IEEE: 399–406, 1995, doi:10.1109/IROS.1995.525827.

[23] M. Quigley, A. Asbeck, A. Ng, "A low-cost compliant 7-DOF robotic manipulator," in 2011 IEEE International Conference on Robotics and Automation, IEEE: 6051–6058, 2011, doi:10.1109/ICRA.2011.5980332.

[24] D. V Gealy, S. McKinley, B. Yi, P. Wu, P.R. Downey, G. Balke, A. Zhao, M. Guo, R. Thomasson, A. Sinclair, "Quasi-direct drive for low-cost compliant robotic manipulation," in 2019 International Conference on Robotics and Automation (ICRA), IEEE: 437–443, 2019, doi:10.1109/ICRA.2019.8794236.

[25] M. Fernandez-Vicente, W. Calle, S. Ferrandiz, A. Conejero, "Effect of infill parameters on tensile mechanical behavior in desktop 3D printing," 3D Printing and Additive Manufacturing, **3**(3), 183–192, 2016, doi:https://doi.org/10.1089/3dp.2015.0036.

[26] V.H. Pinto, J. Gonçalves, P. Costa, "Model of a DC motor with worm gearbox," in CONTROLO 2020: Proceedings of the 14th APCA International Conference on Automatic Control and Soft Computing, July 1-3, 2020, Bragança, Portugal, Springer: 638–647, 2021, doi:https://doi.org/10.1007/978-3-030-58653-9_61.

[27] B. Zhang, Y. Xie, J. Zhou, K. Wang, Z. Zhang, "State-of-the-art robotic grippers, grasping and control strategies, as well as their applications in agricultural robots: A review," Computers and Electronics in Agriculture, **177**, 105694, 2020.

[28] J. Hernandez, M.S.H. Sunny, J. Sanjuan, I. Rulik, M.I.I. Zarif, S.I. Ahamed, H.U. Ahmed, M.H. Rahman, "Current designs of robotic arm grippers: a comprehensive systematic review," Robotics, **12**(1), 5, 2023.

[29] F. Rodríguez, J.C. Moreno, J.A. Sánchez, M. Berenguel, "Grasping in agriculture: State-of-the-art and main characteristics," Grasping in Robotics, 385–409, 2013.

[30] J.F. Elfferich, D. Dodou, C. Della Santina, "Soft robotic grippers for crop handling or harvesting: A review," IEEE Access, **10**, 75428–75443, 2022.

[31] E. Navas, R. Fernandez, D. Sepúlveda, M. Armada, P. Gonzalez-de-Santos, "Soft grippers for automatic crop harvesting: A review," Sensors, **21**(8), 2689, 2021.

[32] J. Hughes, U. Culha, F. Giardina, F. Guenther, A. Rosendo, F. Iida, "Soft manipulators and grippers: a review," Frontiers in Robotics and AI, **3**, 69, 2016.

[33] L. Alvarez-Hidalgo, I.S. Howard, Human evaluation of robotic grippers for berry picking, Springer Nature Switzerland, 2023, doi:https://doi.org/10.1007/978-3-031-43360-3.

[34] B. Siciliano, L. Sciavicco, L. Villani, G. Oriolo, "Motion control," Robotics: Modelling, Planning and Control, 303–361, 2009, doi:https://doi.org/10.1007/978-1-84628-642-1.

[35] K.M. Lynch, F.C. Park, Modern robotics, Cambridge University Press, 2017, doi:10.1017/9781316661239.

[36] M.W. Spong, S. Hutchinson, M. Vidyasagar, Robot modeling and control, Wiley New York, 2006.

[37] I.S. Howard, J.N. Ingram, D.M. Wolpert, "A modular planar robotic manipulandum with end-point torque control," J Neurosci Methods, **181**(2), 199–211, 2009, doi:10.1016/j.jneumeth.2009.05.005.

[38] N. Hendrich, F. Wasserfall, J. Zhang, "3D printed low-cost force-torque sensors," IEEE Access, **8**, 140569–140585, 2020, doi:10.1109/ACCESS.2020.3007565.

[39] H. Asada, T. Kanade, "Design of direct-drive mechanical arms," 1983, doi:https://doi.org/10.1115/1.3269106.