

**IX INTERNATIONAL SCIENTIFIC CONFERENCE
SUMMER SESSION**

INDUSTRY 4.0

26 JUNE - 29 JUNE 2024, VARNA, BULGARIA

ISSN (Print) - 2535-0153

ISSN (Online) - 2535-0161



PROCEEDINGS



ORGANIZER

SCIENTIFIC-TECHNICAL UNION of MECHANICAL ENGINEERING "INDUSTRY 4.0"



INTERNATIONAL SCIENTIFIC CONFERENCE
INDUSTRY 4.0
www.industry-4.eu

Year IX

Volume 1/20

JUNE 2024

ISSN (Print) - 2535-0153
ISSN (Online) - 2535-0161

SESSIONS:

TECHNOLOGICAL BASIS OF “INDUSTRY 4.0”
DOMINANT TECHNOLOGIES IN “INDUSTRY 4.0”
BUSINESS & “INDUSTRY 4.0”
SOCIETY & „INDUSTRY 4.0”

ORGANIZER

**SCIENTIFIC-TECHNICAL UNION OF MECHANICAL
ENGINEERING
“INDUSTRY 4.0”**

108 Rakovski str., 1000 Sofia
e-mail: office@industry-4.eu
www.industry-4.eu

INTERNATIONAL EDITORIAL BOARD

Co-Chairs:		
Prof. D.Sc. Georgi Popov, DHC, Technical University of Sofia, BG		
Members:		
Acad. Igor Bychkov	Institute for System Dynamics and Control Theory SB RAS	RU
Cor. member Alexey Bely	National Academy of Sciences of Belarus	BY
Prof. Alexander Afanasyev	Institute for Information Transmission Problems	RU
Prof. Alexander Guts	Omsk State University	RU
Prof. Bobek Shuklev	Ss. Cyril and Methodius University of Skopje	NM
Prof. Boris Gordon	Tallinn University of Technology	EE
Prof. Branko Sirok	University of Ljubljana	SI
Prof. Claudio Melchiorri	University of Bologna	IT
Prof. Cveta Martinovska	Goce Delchev University, Stip	NM
Prof. Dale Dzemydiene	Mykolas Romeris University, Vilnius	LT
Prof. Dimitrios Vlachos	Aristotle University of Thessaloniki	GR
Prof. Dobre Runchev	Ss. Cyril and Methodius University of Skopje	NM
Prof. Dragan Perakovic	University of Zagreb	HR
Prof. Galina Nikolcheva	Technical University of Sofia	BG
Prof. Galina Zhavoronkova	National Aviation University	UA
Prof. Gerard Lyons	National University of Ireland, Galway	IE
Dr. Giovanni Pappaletta	Politecnico di Bari	IT
Prof. Henrik Carlsen	Technical University of Denmark	DK
Prof. Idilia Bachkova	University of Chemical Technology and Metallurgy	BG
Prof. Idit Avrahami	Ariel University	IL
Prof. Inocentiu Maniu	Politehnica University of Timisoara	RO
Prof. Iurii Bazhal	National University of Kyiv-Mohyla Academy	UA
Prof. Dr. Ivan Kuric	University of Zilina	SK
Prof. Jürgen Köbler	University of Offenburg	DE
Prof. Jiri Maryska	Technical University of Liberec	CZ
Prof. Katia Vutova	Institute of electronics, Bulgarian Academy of Sciences	BG
Prof. Lappalainen Kauko	University of Oulu	FI
Dr. Liviu Jalba	SEEC Manufuture Program	RO
Prof. Luigi del Re	Johannes Kepler University, Linz	AT
Prof. Majid Zamani	Technical University of Munich	DE
Prof. Martin Eigner	Technical University of Kaiserslautern	DE
Prof. Michael Valasek	Czech Technical University in Prague	CZ
Prof. Milija Suknovic	University of Belgrade	RS
Prof. Miodrag Dashic	University of Belgrade	RS
Prof. Murat Alanyali	TOBB University of Economics and Technology	TR
Prof. Nadežda Jankelová	University of Economics in Bratislava	SK
Prof. Nafisa Yusupova	Ufa State Aviation Technical University	RU
Prof. Nina Bijedic	Dzemal Bijedic University of Mostar	BA
Prof. Ninoslav Marina	University of Information Science and Technology – Ohrid	NM
Prof. Olga Zaborovskaia	State Inst. of Econom., Finance, Law and Technologies	RU
Prof. Pavel Kovach	University of Novi Sad	RS
Prof. Petar Kolev	University of Transport Sofia	BG
Prof. Peter Korondi	Budapest University of Technology and Economics	HU
Prof. Peter Sincak	Technical University of Košice	SK
Prof. Petra Bittrich	Berlin University of Applied Sciences	DE
Prof. Predrag Dasic	High Technical Mechanical School, Trstenik	RS
Prof. Radu Dogaru	University Politehnica of Bucharest	RO
Prof. Raul Turmanidze	Georgian Technical University	GE
Prof. René Beigang	Technical University of Kaiserslautern	DE
Prof. Rozeta Miho	Polytechnic University of Tirana	AL
Prof. Salloom A. Al-Juboori	Mutah University	JO
Prof. Seniye Ümit Oktay Firat	Marmara University, Istanbul	TR
Prof. Sreten Savicevic	University of Montenegro	ME
Prof. Sveto Svetkovski	St. Cyril and St. Methodius University of Skopje	NM
Prof. Tomislav Šarić	University of Osijek	HR
Prof. Vasile Cartofeanu	Technical University of Moldova	MD
Prof. Vjaceslavs Bobrovs	Riga Technical University	LV
Prof. Vladyslav Aliksieiev	Lviv Polytechnic National University	UA

CONTENTS

TECHNOLOGICAL BASIS OF “INDUSTRY 4.0”

Fabrication and application of internet of things device using sigfox network and at commands Masaya Watanabe, Hideaki Kanazawa, Satoru Tada, Masayuki Yamauchi, Koji Kakugawa, Takeshi Tanaka, and Katia Vutova	5
A single-board computer based distributed system for internal and external sound monitoring. Piotr Książek, Karol Buchman, Adam Pilch	9
Application of Neural Networks in Underwater Drone Control Jan Przepióra, Krzysztof Mamcarz	12
On some aspects of "open data" Nikolay Netov, Kamen Spassov	16
Preparation and Deployment of Secure Over-The-Air Updates for Embedded Devices: Challenges and Solutions Pavel Palurik, Michal Mikulasek, Lukas Jabloncik	20
Trends in non-linear MIMO Objects Control in the Era of Industry 4.0: The Use of Artificial Neural Networks Patryk Bałazy, Urszula Jachymczyk	25
Application of convolutional networks to detect the operating phases of energy systems using a biomass boiler as an example Szymon Podlasek, Piotr Książek	28
Review of feature selection methods for Predictive Maintenance Systems Urszula Jachymczyk, Paweł Knap	32
Deep Eye - Intelligent Device for Creating Digital Models of Underwater Topography Zuzanna Schab, Marta Żelasko, Krzysztof Lalik, Krzysztof Mamcarz	36
Unified Framework for PdM Algorithm Development: The pdm-tools Architecture Paweł Knap, Urszula Jachymczyk	40
Simulation experiment for the follow-up controller of the MIMO system Kamil Pieprzycki, Urszula Jachymczyk	44
Expanding the Capabilities of Data Transmission Technology to Improve the Efficiency of Multi-Channel Communication Systems Tatiana Sestacova, Gherman Sorochin	48
Enhancing OCR Accuracy for ID-1 Documents with Security Features through Machine Learning-driven Image Optimization Alba Haveriku, Besjana Muraku, Brunela Karamani, Hakik Paci	52
Using artificial neural networks to model climate data to adapt transport infrastructure to climate change Stoyanka Ivanova	56
Modeling solar data using artificial neural networks for solar applications in transport infrastructure Stoyanka Ivanova	61
Technologies between industry 4.0 and 5.0 Личева, Теодора	65
Cybersecurity of the autonomous vehicles as cyber-physical systems Magdalena Puncheva Marina, Ninoslav Marina	68

DOMINANT TECHNOLOGIES IN “INDUSTRY 4.0”

Investigation of the stress-strain state in the implementation of a new innovative method for obtaining hardened screw fittings Сергей Лежнев, Абдрахман Найзабеков, Евгений Панин, Елена Ширяева, Андрей Толкушкин, Дмитрий Куис, Александр Арбуз	70
Design of a lifting machine for container technology on landfills Sergey Kuzmin, Andrey Turbit, Rustem Isakov	75
Microstructure evolution during rolling in relief rolls with an asymmetry factor according to various technological schemes Evgeniy Panin, Alexandr Arbuz, Almas Yerzhanov, Dmitry Kuis, Aibol Esbolat	78
Prototype and characteristic measurement of led-lit pet bottle handmade lantern with baby milk dropped into water Mayumi Tanaka, Keiko Ishihara, Akira Okamoto, Yuko Iwahori, Satoru Tada, Masayuki Yamauchi, Koji Kakugawa, Takeshi Tanaka, and Katia Vutova	80

Extracted Handal seed oil (Citrullus Seeds) used as biodiesel fuel in internal combustion engines Ahmad O. Hasan	84
From 3D scanning, reverse engineering to real part production of personalized car parts Erald Piperi, Ilo Bodi, Artan Hoxha	86
Exergy analysis of a two-cylinder steam turbine from combined cycle power plant at three operating regimes Mrzljak Vedran, Anđelić Nikola, Baressi Šegota Sandi, Prpić-Oršić Jasna	89
Pilot Project for an Integrated Complex for clean soil-based production, harvest and delivery (logistic operations) of Iceberg lettuce Ivan Gramatikov, Teodor Georgiev, Krasimir Kumchev, Kostadin Kostadinov, Marin Midilev	93
Implementing predictive analysis using self-learning digital twins and image analysis with GPT-4 turbo with vision for inspection and repair of construction Mihail Mateev	96
Designing the spinal column of a quadruped robot by using additive manufacturing: A comprehensive approach Ilo Bodi, Erald Piperi, Pascal Doumalin, Mariel Hoxha	100
Photoluminescence properties of self-cleaning coatings on glasses containing Eu2O3 Stancho Yordanov	105
Statistical processing of data from the operation of a photovoltaic power plant Svilen Rachev, Raul Turmanidze, Vazha Shilakadze	108
Device and technology for the production of foam granules of "breathable" composite material based on heat-treated rice husks Lyuben Lakov, Bojidar Jivov, Marieta Gacheva, Krasimira Toncheva	112
Experimental self-cleaning glass coatings for photovoltaic systems placed in a real environment Stancho Yordanov, Tatyana Ivanova, Mihaela Aleksandrova, Vladimir Petkov	116
Small-Scale Photovoltaic Power Systems: Identifying Vulnerabilities and Recommending Security Protocols Michal Mikulášek, Pavel Paluřík, Lukáš Jablončík	119
Advances in Sensory Systems for Unmanned Underwater Vehicles: A Preliminary Review Krzysztof Mamcarz	122
Investigation of composite building elements in a real environment Krasimira Toncheva	125
Utilization of waste chromium electrolyte by addition of Al2O3 nanoparticles to obtain thin chromium coatings Vladimir Petkov, Mihaela Aleksandrova, Stancho Yordanov, Shaban Uzun	130
Bio-synthesis of ZnO nanoparticles as potential corrosion inhibitors Daniela Stoyanova, Irina Stambolova, Maria Shipochka, Ognian Dimitrov, Pavel Markov, Kamelia Kamburova, Tssetska Radeva	134
 <u>BUSINESS & “INDUSTRY 4.0”</u>	
Corporate culture and teamwork in the digital era Nadežda Jankelova	138
Impact of brain thinking preference on engineering project performance Saja Al-Btoosh, Salloom A. Al-Juboori	143
 <u>SOCIETY & „INDUSTRY 4.0”</u>	
A simple e-voting system ensuring identification, privacy and verifiability Mariam Haroutunian, Karen Mastoyan, Arsen Margaryan	150
Contrastive Analysis of Using Japanese Fillers and the Relationship between Fillers and Politeness:Japanese, Chinese, Korean and Vietnamese Chie Yamane-Yoshinaga	154
The promises and practices of digitization in bulgarian public schools Hristova M., Netov N., Netova D.	158
Energy efficiency measures to reduce the carbon footprint: A case study of Tirana airport. Artan Hoxha, Ardit Gjeta, Majlinda Alcani, Elona Caushi, Albert Shira	162

Fabrication and application of internet of things device using sigfox network and at commands

Masaya Watanabe¹, Hideaki Kanazawa², Satoru Tada³, Masayuki Yamauchi³, Koji Kakugawa³,
Takeshi Tanaka^{3,*}, and Katia Vutova⁴
HatanoElectric Co Ltd., Hiroshima, Japan¹
Kanazawa Transportation Co., Ltd., Tokyo, Japan²
Hiroshima Institute of Technology, Hiroshima, Japan³
Institute of Electronics, Bulgarian Academy of Sciences, Sofia, Bulgaria⁴
tanaka@cc.it-hiroshima.ac.jp

Abstract: In this study, the Sigfox network and the Internet were employed to develop an Internet of Things device equipped with temperature, humidity, barometric pressure, and other sensors to measure different physical data. In addition, we attempted to control Sigfox communication by employing AT commands using the Arduino UNOR4 Sigfox module, and social implementation of the proposed system.
Keywords: SIGFOX, ARDUINO, WIFI, MAC ADDRESS

1. Introduction

Knowledge and information have not yet been exchanged in the information society (Society 4.0), which has resulted in issues due to a lack of interdisciplinary cooperation. Society 5.0 is expected to connect all people and things through the Internet of Things (IoT), which will enable the sharing of various types of knowledge and information.

Against From this perspective, using the Sigfox network and the Internet, we have developed an IoT device equipped with temperature, humidity, barometric pressure, and other sensors to measure different physical data [1–5].

In the current study, we attempted to control Sigfox communication using AT commands using the Arduino UNOR4 Sigfox module. In addition, we attempted social implementation using the proposed system.

2. Experimental

2-1 Sigfox module

A Sigfox breakout board (SBB) was used as the Sigfox module (Fig. 1).



Fig. 1 SBB 【BRKLSM100A】 [6]

Key Features - Compact module 14 × 15 × 2.8 mm. (Typ.) - LoRaTM modulation technology. - Sigfox modulation technology. - Sensitivity down to -136dBm@LoRa (BW = 125 KHz, SF = 12) -124dBm@Sigfox (0.6 Kbps) - UART interface. - Low-Power Long Range Transceiver operating in the sub-1 GHz ISM band - Supply voltage range from 1.8 to 3.6 V. - RF interface optimized to 50 Ω. - Output Power Level up to +15 dBm - STM32WLE5CC

A program must be written to run the Sigfox module (Fig. 1) on a microcontroller, and a library for the microcontroller is generally required. To reduce the program's dependence on the

microcontroller, we created the program using AT commands. In the future, we would like to work on creating a microcontroller and dedicated CPU that can run AT commands.

2-2 Arduino UNO R4 Wifi

We selected a highly versatile Arduino microcontroller to run the Sigfox module. In addition, it is necessary to perform Wi-Fi communication to obtain the router's MAC address; thus, we employed an Arduino UNO R4 WiFi device in this study.



Fig. 2 Arduino UNO R4 WiFi device [7]

The Arduino® UNO R4 WiFi is the first UNO board to feature a 32-bit microcontroller and an ESP32-S3 Wi-Fi® module (ESP32-S3-MINI-1-N8). It features a RA4M1 series microcontroller from Renesas (R7FA4M1AB3CFM#AA0), based on a 48-MHz Arm® Cortex®-M4 microprocessor. The UNO R4 WiFi memory is larger than its predecessors, with 256 kB flash memory, 32 kB SRAM, and 8 kB EEPROM. The RA4M1's operating voltage is fixed at 5 V, whereas the ESP32-S3 module is 3.3 V. Communication between these two MCUs is performed via a logic-level translator (TXB0108DQSR).

Features

The R7FA4M1AB3CFM#AA0, often referred to as RA4M1 in this datasheet, is the main MCU on the UNO R4 WiFi, connected to all pin headers on the board as well as all communication buses.

Overview

48 MHz Arm® Cortex®-M4 microprocessor with a floating point unit (FPU)

5 V operating voltage

Real-time Clock

- Memory Protection Unit
- Digital-to-analog Converter (DAC)
- Memory
 - 256 kB Flash Memory
 - 32 kB SRAM
 - 8 kB Data Memory (EEPROM)
- Peripherals
 - Capacitive Touch Sensing Unit (CTSU)
 - USB 2.0 Full-Speed Module (USBFS)
 - 14-bit ADC
 - Up to 12-bit DAC
 - Operational Amplifier
- Power
 - Operating voltage for RA4M1 is 5 V
 - Recommended input voltage (VIN) is 6-24 V
 - Barrel jack connected to VIN pin (6-24 V)
 - Power via USB-C® at 5 V
- Communication
 - UART (pin D0, D1)
 - SPI (pin D10-D13, ICSP header)
 - I2C (pin A4, A5, SDA, SCL)
 - CAN (pin D4, D5, external transceiver is required)

2-3 Assembling Sigfox module and Arduino

The SBB was placed on a custom manufactured printed circuit board (PCB). Figure 3 shows the IoT device configured by connecting a PCB to an Arduino UNO R4 WiFi device.

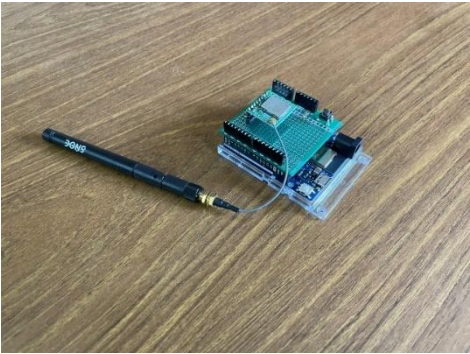


Fig. 3 Sigfox module and Arduino device

2-4 Scketch of Arduino UNO R4 Wifi

Using the IoT device with the Sigfox communication function (Fig. 3), we created a sketch to obtain the MAC address of the Wi-Fi antenna and transfer that value using Sigfox (Fig. 4).



Fig.4 Sketch program using Arduino UNO R4 Wifi

2-5 Sigfox Cloud

Hiroshima Institute of Technology is a Sigfox partner of KYOCERA Communication Systems Co., Ltd., so it can use the Sigfox Cloud[8].

2-5-1 End product certificate and prototype registration[9]

When registering a device in the Sigfox Cloud, an end product certificate must be input. Sigfox compatible commercial devices must obtain device certification, which is referred to as the Sigfox Ready Certification (formerly P2 certification).

Note that the settings for the end product certificate differ depending on whether you are registering a device that has obtained Sigfox Ready Certification or a device that has not been certified.

Device - New

Device information

Identifier (hex!)	AB12CD
Name	Test Device #00AB12CD
PAC	xxxae2ec3c878975c896ca
End product certificate	P_0006_DAD8_01

Where can I find the end product certificate?

Fig. 5 Registering a Sigfox device to the Sigfox Cloud

2-5-2 Callback

The advantage of using Sigfox Cloud is that the data obtained from IoT devices can be processed and transferred to the cloud. Here, we show an example where the data obtained from IoT devices are transferred to email via the Sigfox Cloud.

3. Results and Discussion

Communication is performed after setting up the IoT device (Fig. 3) with the Sigfox communication function. The corresponding results are shown in Fig. 6.

Device 30D1222 - Information

Name: brkwls100a001

Protocol: V1

Activable state: ☒ ⓘ

Message modulo: 4096

Sequence number: 43 (2024-05-22 09:23:38)

Trash sequence number: N/A (N/A)

Last seen: 2024-05-22 09:23:38

PAC: AC673C00E944C220

Product certificate: P_0028_4DAF_01

Latitude: 0.000 (degrees)

Longitude: 0.000 (degrees)

Device type: brkwls100a-001

State: OK

Link Quality Indicator:  ⓘ

Communication status: 

Contract: From BUY for 1 devices

Activation date: 2024-02-19 19:27:38

Token validity: 2025-02-19

Fig. 6 Information about device 30D1222

Note that the Link Indicator (Fig. 6) was displaying the strength, indicating that it was receiving data from the IoT device. The corresponding data message is shown in Fig. 7.

Device 30D1222 - Messages					Base station reception attributes			LQI	Callbacks	Location
Time	Delay (s)	Seq Num	Data / Decoding		Station	RSSI (dBm)	Freq (MHz)			
2024-05-22 09:23:38	< 1	43	d80d17276aac001132916143		D9FC	-127.00	923.2502			
					6969	-121.00	923.2500			
					9A63	-126.00	923.2498			

(a)

Time	Delay (s)	Seq Num	Data / Decoding
2024-05-22 09:23:38	< 1	43	d80d17276aac001132916143

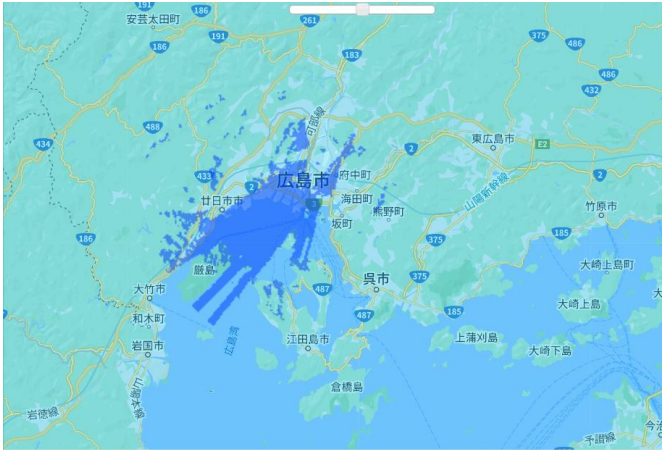
(b)

Base station reception attributes		
Station	RSSI (dBm)	Freq (MHz)
D9FC	-127.00	923.2502
6969	-121.00	923.2500
9A63	-126.00	923.2498

(c)

LQI	Callbacks	Location
		

(d)



(e)

Fig. 7 Data message from device 30D1222. (a) The entire message. (b) Expanded message time, delay, serial number, and data. (c) Base station number, strength, and frequency when receiving the signal. (d) Link quality indicator, callback, and device location. (e) Device location is highlighted in blue on the map. Its width is approximately 20 km.

Figure 8 shows the arrangement of the router and IoT devices. Figure 10 shows the email that was forwarded using this arrangement using the MAC address obtained by the IoT device employing the callback function in the Sigfox Cloud (Figure 9).

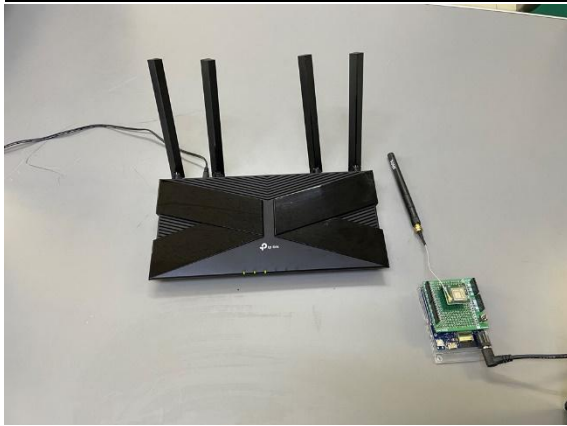


Fig. 8 Callback used to process messages

Device type brkwlsm100a-001 - Callback edition

Callbacks

Type: **DATA** **UPLINK**

Channel: **EMAIL**

Custom payload config: **sls:int:48 big-endian slg:int:48 big-endian**

Recipient: **tanaka@cc.it-hiroshima.ac.jp**

Multiple emails allowed separated by comma,semicolon or new line

Subject syntax: **Subject with device (device)**
 Message syntax: **Message containing time (time), key1 (var1), key2 (var2)...**
 Available variables: **device, time, station, data, rssi, seqNumber, deviceTypeid**
 Custom variables: **customData#sls, customData#slg**

Subject: **kanazawa_test**

Message:

```
{
  "device": "{device}",
  "time": "{time}",
  "seqNumber": "{seqNumber}",
  "data": "{data}"
}
```

Ok Cancel

Fig. 9 Callback function to process messages

```
{
  "device": "30D1222",
  "time": 1716970277,
  "seqNumber": 51,
  "data": "d807b6c70e7b0810868199fc"
}
```

Fig. 10 Email forwarded using Sigfox

Note that the first half of the data (d807b6c70e7b2c1a056815c3), i.e., d807b6c70e7b, is the router's MAC address.

4. Application to society

A router with a known MAC address can be placed at an evacuation center during a disaster, and IoT devices with Sigfox communication functionality can be provided to the evacuees in the area. Then, who is there and when did they get here? Such information is accumulated on the network.

Using such a system, it would be possible to determine who has moved to evacuation centers without human intervention. Thus, it would be possible to obtain useful information about the damage situation, which could be used to support disaster response efforts.

5. Conclusion

Knowledge and information have not yet been exchanged in the information society (Society 4.0), leading to issues due to a lack of interdisciplinary cooperation. Society 5.0 connects all people and things with the IoT to share various types of knowledge and information.

Against this background, using the Sigfox network and the Internet, we developed IoT devices equipped with temperature, humidity, barometric pressure, and other sensors to measure different physical data [1–5].

In this research, we attempted to control Sigfox communication using AT commands using the Sigfox module Arduino UNOR4. We also attempted social implementation using this system.

Acknowledgments

We thank Enago (www.enago.jp) for the English language review.

References

- 1) R. Matsuda, H. Suizu, K. Hiyama, K. Kai, I. Yamasaki, K. Oda, N. Okamitsu, K. Eto, T. Tanaka, Industry 4.0, 3(1), 2018, 11-14
- 2) T. Tanaka, N. Okamitsu, T. Hamada, K. Nishino, Electrotechnica & Electronica (E+E), 55(3-4), 2020, 35-40, ISSN: 0861-4717 (Print), 2603-5421 (Online)
- 3) T. Tanaka, N. Okamitsu, K. Nishino, Electrotechnica & Electronica (E+E), 55(5-8), 2020, 94-98, ISSN: 0861-4717 (Print), 2603-5421 (Online) 304 YEAR VIII, ISSUE 6, P.P. 302-305 (2023) INTERNATIONAL SCIENTIFIC JOURNAL "INDUSTRY 4.0" WEB ISSN 2534-997X; PRINT ISSN 2534-8582
- 4) N. Okamitsu, K. Miyazono, T. Tanaka, Electrotechnica & Electronica (E+E), 55(9-12), 2020, 99-104, ISSN: 0861-4717 (Print), 2603-5421 (Online)
- 5) T. Tanaka, M. Yamauchi, K. Kakugawa, T. Uchida, S. Ishihara, K. Ishihara, Electrotechnica & Electronica (E+E), 57(1-2), 2022, 24-30, ISSN: 0861-4717 (Print), 2603-5421 (Online)
- 6) KYOCERA. Sigfox Breakout Board [BRKLSM100]. <https://www.kccs.co.jp/sigfox/solution/product/device71/> (in Japanese)
- 7) Arduino Cloud. Arduino® UNO R4 WiFi <https://store-usa.arduino.cc/products/uno-r4-wifi?selectedStore=us>
- 8) KYOCERA. Sigfox partner list. <https://www.kccs.co.jp/sigfox/service/pdf/pdf-sigfox-partnerslist-jp-01.pdf>
- 9) KYOCERA. Sigfox: End product certificate and prototype registration. 2018. <https://www.kccs.co.jp/sigfox/blog/technical/2018/1114/>

A single-board computer based distributed system for internal and external sound monitoring.

Piotr Książek¹, Karol Buchman¹, Adam Pilch¹

¹AGH university of Krakow, Poland

piotr.ksiazek250@gmail.com

Abstract: This work showcases a distributed noise monitoring system used for measuring noise levels both inside and outside of buildings in urban environments. The system utilises single board computers as main operating units which handle the measurement of parameters and communication. The noise monitoring stations can measure instantaneous and long-term acoustic noise levels, as well as atmospheric parameters such as atmospheric pressure, air temperature and humidity. Measurement stations are connected to an external server via the MQTT protocol. The external server allows for the recording of data into a secure database, as well as for providing end users with historical and instantaneous data on request. A test run of the measurement system was performed and has shown that the measurement stations can function in their design capacity. The system shows great promise in use in internal monitoring of industrial plants both as a precautionary measure for worker health and safety and as an easily modifiable sound monitoring platform in the industry 4.0 standard.

Keywords: ACOUSTIC MONITORING, INTERNET OF THINGS, INDUSTRIAL NOISE, SOUND MEASUREMENT

1. Introduction

Noise in both urban and rural areas has increased in the past century to the point where its impact on human health and well-being is measurable on both a societal [1] and personal scale [2]. The problem is exacerbated by the transportation industry and noise produced in industrial plants, as those activities, which are necessary to our way of life, are omnipresent in urban spaces.

Monitoring the pollution of environments people live in is both mandated by law in many European countries [3] and is gaining popularity as pollution monitoring services make acquiring data pertaining to air quality both easy and convenient for those who are concerned about the impact of pollution on their health. An example of such a system is Airly [4], which provides real-time and historical data from a network of air quality sensors forming a dense network across a wide area. The sensors used by the Airly network are low-cost, easy to install and maintain measurement stations [5], which allows the system to be as widespread as possible.

Acoustic monitoring can be done in an analogous way, as solutions exist for long-term monitoring of sound pressure level. Those solutions, however, are mostly professional class 1 or 2 conforming measurement devices [6], which are costly to deploy en masse and are non-modifiable in their functionality. Previously such systems have been implemented, however most of those attempts do not focus on the integration aspect of such systems [7]. To develop a distributed monitoring system, a new open, and modular device ought to be designed. The main goal of this work was to develop a modular, easily modifiable measurement device capable of monitoring sound pressure levels and environmental parameters. The device was to implement distributed measurement system methodology [8], with all of the signal processing and data manipulation happening on-device. The design steps of the device are explained in the next section, with the software architecture, networking configuration of the system and measurement validation following.

2. Measurement device hardware

The design principles of the measurement station were mainly focused on modularity. It was imperative to design a measurement station that used as many off the shelf parts as possible with minimal custom part design required. As such, when designing the station itself, four main modules were decided to be included. The modules are as follows.

- Main computational module.
- Microphone module.
- Environmental sensing module.
- Power delivery module.

Within the bounds of the modular system two modules were decided to be connected within a single sensory assembly. Those were the microphone and environmental sensing modules. Otherwise, all of the station's systems are completely interchangeable with equivalent types of hardware.

To ensure the modifiability and to allow for sufficient computational capacity of the base measurement device, a Raspberry Pi 4B single-board computer was chosen as the base computing module of the measurement station. This allows the use of the I²S and I²C protocols for communication with both the microphone and the environment probe. The single board computer also provides power conversion for low-voltage modules and an operating system for the software to execute within.

Many types of microphones were under consideration when designing the measurement station. Finally, the microphones used were consumer-grade MEMS microphones mounted on a prototyping PCB board. The microphones were SPH0645LM4H produced by Knowles and integrated by Adafruit into the Adafruit 3421 breakout board. The microphones declare a sensitivity of -26 dBFS and a signal to noise ratio of 65 dB(A). The microphone was connected to the measurement system via the I²S protocol.

Two main environmental probes were considered when designing the environmental sensing module. The contenders were the DHT 22 temperature and humidity sensor, and the BME 280 temperature, humidity, and atmospheric pressure sensor. The DHT 22 module proved to be too large to reliably fit within the enclosure chosen for the project, as such the BME 280 sensor was chosen. It interfaces with the main compute unit of the station via the I²C protocol.

A final choice was made when it came to the power backup of the measurement station. A design constraint was established, that the measurement station is required to withstand a 120-minute power outage event without stopping the monitoring. It was quickly established that the most resource efficient way to do this was to utilise a UPS hat available for the Raspberry Pi 4B and modify it to use a larger energy cell. The Waveshare Li-Ion Battery HAT was used and modified to accommodate the larger size of 18650 battery cells. The final battery cells used were Sony US18650VTC4 2100 mAh Li-Ion cells, which would provide adequate two-hour battery backup for an expected maximum load of 1 A.

Many assembly parts needed to be 3D printed, and were designed for ease of print, with emphasis on the ease of assembly of the measurement station. An important part of the assembly is the microphone windshield which is a fully custom-made assembly consisting of two parts, between which an audio-grade microphone protective foam is placed before screwing the part into the main assembly.

The whole hardware of the measurement device can be seen in Figure 1. Assembly of a single measurement station was measured

to be around 90 minutes using simple tools. Printing of the 3D printed parts was not included in the assembly time.

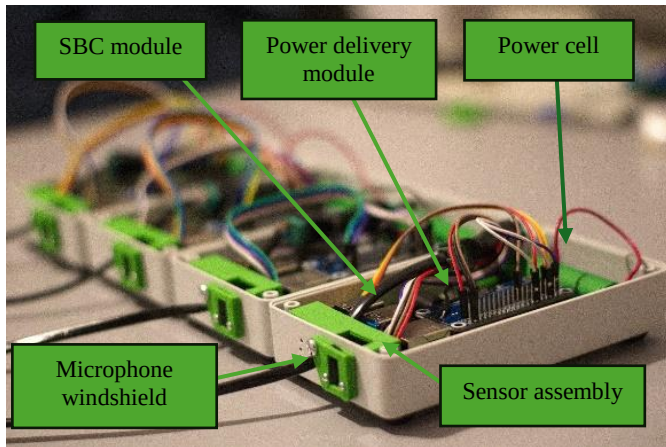


Figure 1. All four measurement stations made for the project with modules shown.

3. Software solutions and networking

The measurement station base operating system is the Raspberry Pi Bullseye OS with a kernel version of 6.1. This legacy version of the system was chosen due to compatibility concerns with drivers for older devices. The operating system was set up in text-only mode. Within the operating system, an Adafruit-provided driver was installed to allow for the interfacing of the MEMS I²C microphone with ALSA. The driver provides a simplified audio device software interface, that is nonetheless enough to use the higher-level API provided by the PortAudio package to interface with measurement scripts.

The software solution used within the measurement station is as modular as the hardware it runs on. It is split into multiple modules, which are independent from each other by design and interface only via a central overseer thread that launches appropriate module threads at appropriate points in time. It should be noted that due to the structure of the software it is easy to implement different monitoring methods, and even new measurement threads. This allows the software implementation to act as a base for future modifiable projects.

The overseer thread launches a measurement thread at the start of every full minute, which then opens a device handle and performs acoustic measurements for a full minute, after the measurement is completed, the results are appended to a file stored in a configurable location. The threads overlap in their execution in such a way that continuous monitoring is performed. The overseer launches the communication and parameter calculation thread every minute at exactly half a minute after the measurement thread begins and reads previously saved results from a file on disk. The thread then calculates the desired parameters and publishes them both to another file on disk with a configurable location and to a central server via the MQTT protocol.

All file operations are safeguarded from race conditions and conflicts in file handling. The system overall is designed in a fail-safe manner, meaning that even if a thread was to come across an unhandled exception and stop unexpectedly, the event would be logged, and another thread will start on the next measurement period. Many of those features are user-configurable and do not require a change in code to modify.

To ensure a great compatibility and accessibility to collected data across many platforms, a REST API was created. Its main function is to be a connector for measurement stations, the database and applications that use the data – mostly done via multiple endpoints. The API was created in .NET 6.0 and uses frameworks like ASP.NET Core, Entity Framework. Due to the selected approach of a code architecture, it is easy to expand and add new functionalities.

The currently selected way of utilizing the API is through a website made in Blazor, which also uses some open-source components, mainly for the user interface. The website provides features to monitor and analyse the data. The main two are a map and a variety of charts. It allows a user to watch real-time levels of noise measured by stations.

4. Measurement accuracy validation

The main purpose of the base station is to provide accurate sound pressure level measurements. To this end, the module was calibrated using the anechoic chamber at AGH UST Technical Acoustics Laboratory. Measurements were performed according to the reference microphone method of measuring the frequency characteristics of the measurement set-up. The method used was inspired by the IEC 61094-3:2016 standard [9].

First, a recording was made using the reference microphone with a known, validated frequency response. The microphone used was a Gras 46AE measurement microphone and pre-amplifier set. The microphone was then connected to the Gras 12AX signal conditioner, and then finally to the USI U24XL audio interface. The speaker used in the measurement session was a Genelec 8040BPM studio monitor. The measurement set-up is shown in Figure 2. The distance between the speaker and the microphone was 2 m.

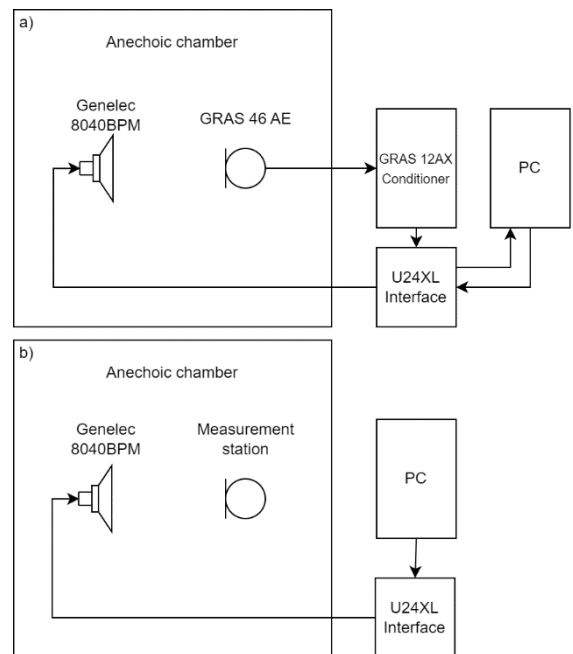


Figure 2. Measurement set up for determining free-field microphone characteristics, a) Reference microphone b) Test scenario with the measurement station.

After performing measurements with the reference microphone, the measurement station was placed in the same spot, as can be seen in Figure 3. The measurement station was positioned so that the microphone of the station was exactly at the same distance from the speaker as the reference microphone capsule was. Measurements were performed by both recording the raw audio data and the sound power levels reported by the measurement station.



Figure 3. Measurement station positioned within the anechoic chamber.

The resulting frequency response curve of the monitoring station can be seen in Figure 4. It can be seen that the frequency response of the measurement set-up is not flat, and shows around 6 dB of gain around 2000 Hz, which then stays consistent up to 20 kHz. This is to be expected of a low-cost microphone and was accounted for in the sound pressure level processing. When looking at total A-weighted levels, the absolute error between the reference microphone and the measurement station sound pressure level was established to be -2.3 dB. The calculated error value is the calibration coefficient of this microphone set-up and was put into the configuration file for the station. The noise floor of the microphone was also measured by leaving the station within the anechoic chamber and performing measurements. The measured noise floor of the microphone is 26.1 dB(A).

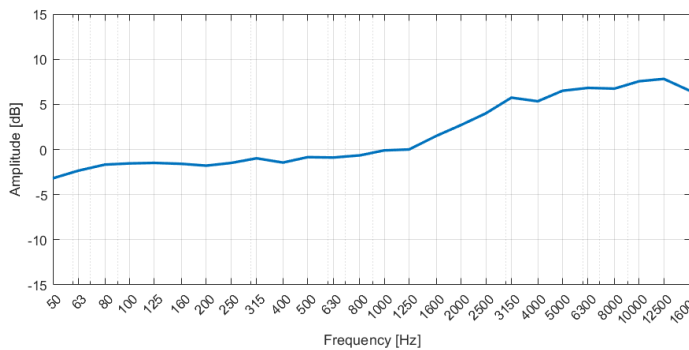


Figure 4. Frequency response curve of the tested device. A sharp rise in the response can be seen around 2000 Hz.

5. Test rollout

After calibrating measurement stations, a test rollout of the system was performed. In total, three stations were placed in the environment around Krakow. One was located at AGH University D1 building, outside the laboratory window. The other two were situated outside the windows of private apartments and connected to the wi-fi networks available at the measurement site. Test locations of microphones can be seen in Figure 5.



Figure 5. Station mounted outside a private apartment (top). Station mounted at AGH D1 building (bottom).

The test was carried out through the weekend, with the stations performing continuous measurements and sending the results to the MQTT broker, which passed data onto the central server being tested at the same time. After the weekend-long stress test, a long-term test was carried out on the station at AGH. The test was concluded after two months, due to the extreme cold conditions, which caused the power delivery chip to malfunction. This, despite the hardware failure is a positive result of the test, as the main goal of the long-term test was to establish the effective operating conditions of the station. The temperature at the moment of failure recorded by the station was around -10°.

6. Summary and future potential

The goal of this work was to develop an open, easily modifiable in both hardware and software measurement station that could monitor sound pressure levels both within rooms and in the environment. The system was to utilise off the shelf parts with minimal 3D printing or machining required and was to provide a codebase that would be modular and easy to expand upon.

Future work with this project may also include a fuller exploration of the Raspberry Pi's computing capabilities. The stations may be used, for example, as a base to develop a deep learning based distributed early fault detection system for large-scale industrial plant monitoring, the need for which is evidenced by the impact such systems make both in industry and science [10]. With the reduced cost per station and relatively high compute ability there are many possibilities to explore in the future.

References

- [1] Stansfeld Stephen, Haines Mary, and Brown Bernadette, "Noise and Health in the Urban Environment," vol. 15, no. 1–2, pp. 43–82, 2000, doi: doi:10.1515/REVEH.2000.15.1-2.43.
- [2] J. Hegewald *et al.*, "Traffic Noise and Mental Health: A Systematic Review and Meta-Analysis," *Int J Environ Res Public Health*, vol. 17, no. 17, 2020, doi: 10.3390/ijerph17176175.
- [3] S. Kephalopoulos, M. Paviotti, F. Anfosso-Lédée, D. Van Maercke, S. Shilton, and N. Jones, "Advances in the development of common noise assessment methods in Europe: The CNOSSOS-EU framework for strategic environmental noise mapping," *Science of The Total Environment*, vol. 482–483, pp. 400–410, 2014, doi: <https://doi.org/10.1016/j.scitotenv.2014.02.031>.
- [4] "Airly network webpage," <https://airly.org/en/>.
- [5] M. Banach, R. Długosz, J. Pauk, and T. Talaška, "Hardware Efficient Solutions for Wireless Air Pollution Sensors Dedicated to Dense Urban Areas," *Remote Sens (Basel)*, vol. 12, no. 5, 2020, doi: 10.3390/rs12050776.
- [6] "IEC 61672-1:2013 Electroacoustics - Sound level meters - Part 1: Specifications," *International Electrotechnical Commission*. 2013.
- [7] C. Mydlarz, J. Salamon, and J. P. Bello, "The implementation of low-cost urban acoustic monitoring devices," *Applied Acoustics*, vol. 117, pp. 207–218, 2017, doi: <https://doi.org/10.1016/j.apacoust.2016.06.010>.
- [8] D. Grimaldi and M. Marinov, "Distributed measurement systems," *Measurement*, vol. 30, no. 4, pp. 279–287, 2001, doi: [https://doi.org/10.1016/S0263-2241\(01\)00019-7](https://doi.org/10.1016/S0263-2241(01)00019-7).
- [9] "IEC 61094-8:2012 Measurement microphones - Part 8: Methods for determining the free-field sensitivity of working standard microphones by comparison," *International Electrotechnical Commission*, 2012.
- [10] A. Qurthobi, R. Maskeliūnas, and R. Damaševičius, "Detection of Mechanical Failures in Industrial Machines Using Overlapping Acoustic Anomalies: A Systematic Literature Review," *Sensors*, vol. 22, no. 10, p. 3888, May 2022, doi: 10.3390/s22103888.

Application of Neural Networks in Underwater Drone Control

Jan Przepióra^{1,*}, Krzysztof Mamcarz¹
AGH University of Kraków¹
przepiora@agh.edu.pl

Abstract: The integration of neural networks has revolutionised various technological domains, including the control mechanisms of underwater drones, especially in the context of Industry 4.0. This review explores the application of neural network architectures to improve the navigation, stability and overall performance of underwater drones. A systematic analysis of current methods is provided, focusing on their effectiveness in addressing challenges such as dynamic underwater environments, sensor noise, and real-time decision making. Key advances in neural network-based control strategies are discussed. Furthermore, the integration of these networks with traditional control systems to achieve robust and adaptive control frameworks is highlighted. Through the examination of case studies and experimental results, this review identifies potential areas for future research and development to further the advancement of autonomous underwater vehicles through intelligent control systems.

Keywords: UNDERWATER DRONE, REINFORCEMENT LEARNING, NEURAL CONTROL, ARTIFICIAL INTELIGENCE, UUV

1. Introduction

Over 70% of the Earth's surface is covered by water, including vast oceans, seas and other aquatic environments. Despite their immense size, these underwater realms remain far less explored than the landmasses. The depth and vastness of these aquatic areas present significant challenges and make them some of the least understood parts of our planet.

There are many difficulties in exploring underwater environments. Human divers face many limitations, such as the dangers associated with high pressure, limited dive times and the complexities of maintaining life support systems. In addition, the sheer size and often challenging conditions of underwater areas make direct human exploration particularly problematic. These factors contribute to the limited understanding we have of the underwater environment [1].

Underwater drones, also known as autonomous underwater vehicles (AUVs) and remotely operated vehicles (ROVs), offer a promising solution to these challenges. These drones can reach depths and stay underwater for periods of time that are impossible for human divers. Equipped with cameras, sensors and navigation systems, they can collect detailed images and data from the underwater world. This capability enables continuous monitoring and exploration of underwater environments without the risks associated with human diving.

2. Underwater Drone Applications

Unmanned underwater vehicles (UUVs) have a wide range of applications in the military, scientific and industrial sectors, each of which benefits from the unique capabilities of these advanced machines[1], as shown in Fig.1.

In the military domain, UUVs play a key role in enhancing national security and defence operations. They are used extensively for surveillance, providing real-time monitoring of vast ocean areas and detecting potential threats[2]. Anti-submarine warfare is another critical application where UUVs detect and track enemy submarines, significantly enhancing naval capabilities. In addition, UUVs are being used in mine countermeasures, the identification and neutralisation of underwater mines to ensure safe passage for military and commercial vessels[3].

On the scientific side, drones have revolutionised marine research and exploration. In marine biology, they enable the study of marine ecosystems, allowing researchers to observe and analyse underwater life in its natural habitat without human intervention. Geological surveys benefit from UUVs through detailed mapping of the seafloor, revealing underwater geological formations and processes[4]. Archaeological surveys also use such vehicles to locate and inspect underwater wreckage, providing valuable insights into historic shipwrecks and submerged settlements[4], [5]. In addition, UUVs are essential in tracking oil spills and gas leaks, helping to mitigate environmental disasters and monitor their impact on the marine environment.

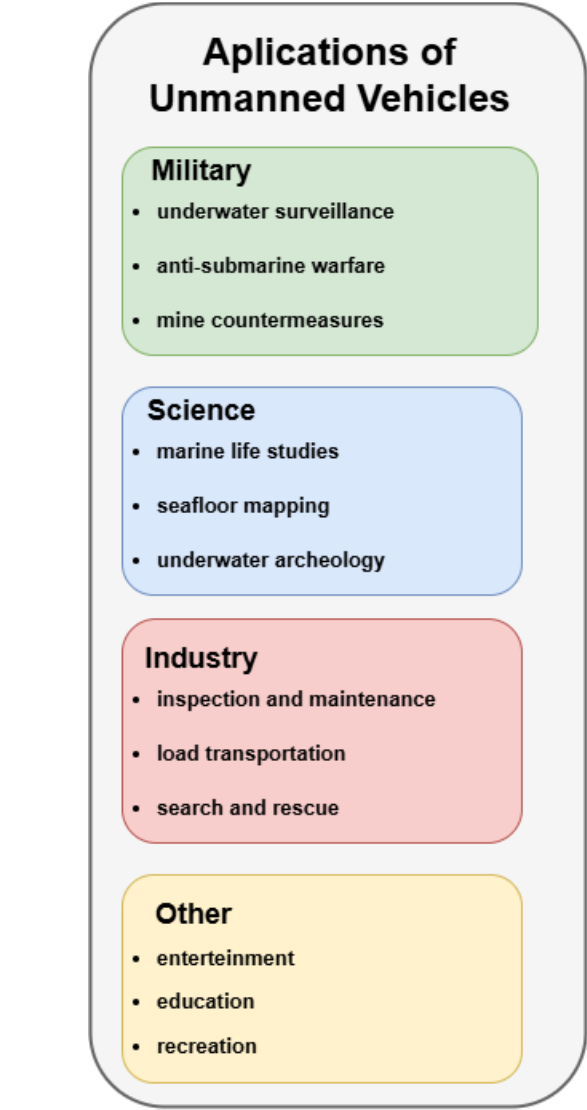


Fig. 1 UUVs applications.

In the industrial sector, underwater drones are essential tools for maintaining and inspecting underwater infrastructure[6], [7]. They are used to inspect, repair and maintain underwater structures such as pipelines, offshore platforms and cables, ensuring their integrity and functionality[8]. UUVs and ROVs facilitate the delivery of payloads to the seabed, supporting various subsea construction and installation projects. They also play a vital role in search and rescue operations, helping to quickly locate and recover objects lost at sea. UUVs are also used to collect underwater video footage, providing high-quality images for scientific research, environmental monitoring and media production.

In addition to the above, unmanned underwater vehicles are expanding their impact in areas such as tourism, education and recreation. In the tourism industry, UUVs provide unique underwater experiences for tourists, such as guided tours of coral reefs and shipwrecks, adding a new dimension to underwater exploration without the need for diving. Educational institutions are using UUVs to engage students in marine science through interactive learning experiences, virtual field trips, and hands-on training in oceanography and robotics. On the recreational side, hobbyists and enthusiasts are using smaller, consumer-grade UUVs for activities such as underwater photography, videography, and personal exploration of lakes and coastal waters.

3. Challenges for Underwater Vehicle Control

The control of unmanned underwater vehicles is inherently complex due to their multi-dimensional manoeuvrability and the challenging marine environment. The UUV itself is a relatively complex object to control as it has six degrees of freedom - three translational (pitch, roll and heave) and three rotational (roll, pitch and yaw). These degrees of freedom make the dynamics of UUVs inherently non-linear and coupled, requiring advanced control algorithms to effectively manage their behaviour. In addition, the marine environment presents other challenges that complicate control efforts[1], [9].

First of all, external ocean disturbances represented in Fig.2, such as waves and currents (F_d), induce nonlinear damping and perturbations of dynamic parameters. These disturbances are usually highly unpredictable and can vary significantly in time and space, making it difficult to maintain stable control of the UUV.

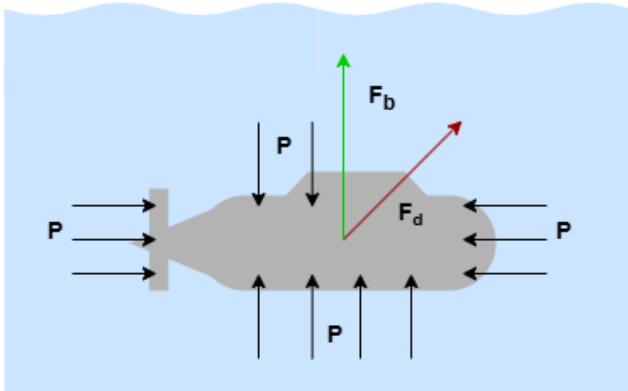


Fig. 2 Sample external disturbances for UUV.

Secondly, structural changes to the vehicle, whether due to modifications, wear or load variations, can alter its hydrodynamic parameters. The motion of the UUV itself also affects these parameters, adding another layer of uncertainty to the dynamic model. Variations in buoyancy (F_b), which can occur due to load changes or internal ballast adjustments, have a significant impact on the stability and control of the vehicle. In addition, increasing water pressures (P) at greater depths affect the structure and buoyancy of the vehicle, as well as the performance of its sensors and actuators. These factors require adaptive control strategies that can respond to such variations in real time, ensuring that the UUV can maintain its intended trajectory and operational effectiveness.

Thirdly, obtaining accurate sensor signals is problematic due to several factors. Communication delays and ocean noise can result in missing or corrupted velocity data, making it difficult for the controller to make timely and accurate adjustments. In addition, due to the underwater operating environment, such vehicles are often unable to acquire GPS data, which is advantageous for determining precise position and speed in classic autonomous vehicle navigation. This lack of reliable GPS information forces reliance on alternative, often less accurate, navigation methods. The harsh marine environment, characterised by high pressure, low

temperatures and the possible presence of biofouling, places additional constraints on sensor performance. Low-cost sensors, often used to keep the overall cost of UUVs manageable, are particularly susceptible to these challenges, making it difficult to filter out ocean noise and obtain reliable measurements.

In the context of Industry 4.0 standards and concepts, overcoming these challenges is critical to advancing the autonomy and efficiency of UUVs. The integration of smart sensors and advanced data analytics, key components of Industry 4.0, can improve the performance and reliability of UUVs. By using real-time data processing, machine learning algorithms and robust communication networks, UUVs can achieve better situational awareness and adaptive control. These technologies can help mitigate the effects of sensor inaccuracies and environmental noise, ensuring more accurate speed and position estimates. The adoption of digital twin technology, another Industry 4.0 concept, can further improve the design and operation of UUVs by enabling virtual simulations and predictive maintenance, ultimately leading to more resilient and efficient underwater operations.

4. Reinforcement Learning Integration

Given the challenges of traditional control approaches for complex and uncertain environments, such as those faced by unmanned underwater vehicles, reinforcement learning (RL) emerges as a promising solution[10], [11], [12], [13]. RL, a branch of machine learning, addresses decision problems in dynamic and uncertain environments by learning policies to determine actions in a given state to achieve specific goals. In particular, RL offers several advantages in system control. It enables adaptation to changing environments, resilience to external disturbances, and control of systems with uncertain behaviour. RL works through stages of exploration, learning and active learning, allowing the agent to continuously improve its policies based on the rewards it receives.

In a conventional unmanned underwater vehicle control system, a feedback controller can be used to observe the state of the UUV and use pre-determined gains to generate a feedback signal that is then used to adjust the behaviour of the UUV. This controller receives feedback from the UUV's sensors, measures its response to the actions that are applied, and refines its behaviour over time in response.

In a reinforcement learning system, however, the traditional controller is replaced by an agent. This agent interacts directly with the UUV, acting as a controller, and receives a reward signal from the environment. The agent uses this reward signal to learn which actions yield the maximum reward, effectively controlling the UUV. A schematic of such a structure is depicted in Fig. 3.

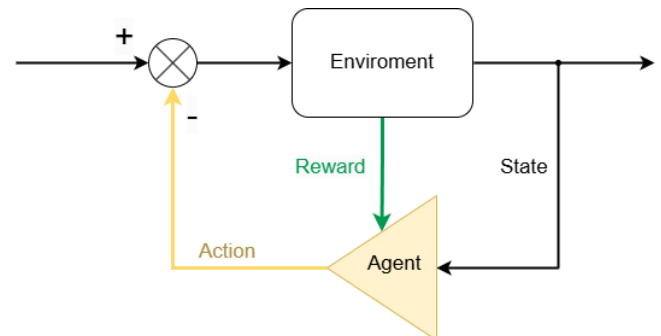


Fig. 3 RL control system

The process of control within this RL based control system unfolds in the following way: Firstly, the agent observes the state of the UUV. It then takes an action that generates feedback gains. Next, the UUV, which is the object and therefore part of the environment, receives a control signal calculated as the difference between the setpoint and the action taken by the agent. The environment then provides the agent with a reward signal based on

the UUV's response to the action. The agent then updates its policy based on this reward signal, and the process repeats. A key difference between a traditional feedback control system and an RL-based system is that in the latter operation there is no model of the UUV applied. Instead, the agent learns to control the UUV through direct interaction, with the reward signal serving as the primary source of information for learning and adaptation. This characteristic makes RL-based control systems particularly well-suited to dealing with the complexity and uncertainty inherent in controlling UUVs in dynamic and unpredictable marine environments.

RL operates through exploration, learning, and active learning stages, allowing the agent to continually improve its policies based on received rewards.

Exploration is the initial phase of neural network training where the agent interacts with its environment by experimenting with different actions. Through this process, the agent attempts to gain insight into the dynamics of the environment and to identify patterns regarding which actions lead to favourable outcomes or rewards. Exploration serves as the basis for the agent's understanding of the complexity of the environment and helps it to develop effective decision-making strategies.

The exploration phase is followed by the learning phase, the iterative process by which the agent enhances its understanding of the environment and refines its decision policies. Using the data collected during exploration, the agent uses complex algorithms to identify underlying patterns and correlations between actions and rewards. Through continuous iteration and adaptation, the agent incrementally improves its decision-making capabilities and gradually refines its policies to maximise cumulative rewards over time.

The final stage of the RL process is active learning, in which the agent focuses its attention on states and actions that offer the greatest potential for reward. Based on the knowledge gained during exploration and learning, the agent strategically allocates its resources towards actions that yield the most beneficial outcomes.

Reinforcement Learning (RL) offers several notable advantages in the field of system control. Firstly, RL facilitates adaptation, enabling agents to adapt seamlessly to a dynamic environment. With the ability to update policies in response to environmental changes, RL ensures that satisfactory levels of performance are consistently maintained. Secondly RL exhibits resilience in the face of perturbations. Even in the presence of environmental disturbances, RL-enabled agents can continue to perform optimally, demonstrating a robustness that increases system reliability. In addition, RL proves invaluable in dealing with uncertainty, particularly in systems where complete knowledge of behaviour is lacking. By learning policies that provide satisfactory levels of performance despite incomplete understanding of the system, RL enables effective control in uncertain environments.

5. Artificial Intelligence Applications in UUVs

Alternative approaches to the control of underwater vehicles are also proposed, focusing on the implementation of a neural network-based self-tuning PID controller[14], [15], [16]. Unlike traditional PID controllers, which require manual tuning and can be difficult to adapt to changing conditions, this method uses neural networks to automatically adjust PID gains in real time. This is particularly important for underwater vehicles, where parameters such as weight, buoyancy and added mass can vary significantly due to factors such as attached tools or environmental conditions such as ocean currents. By using neural networks, the controller can continuously estimate and optimise the PID gains to ensure stability and accurate position tracking even in dynamic and uncertain environments.

Neural networks can be found in navigation systems, demonstrate their versatility by addressing the challenge of accurately modelling thrust in practical trajectory control

scenarios[17]. Instead of relying on accurate thrust models, some innovative approaches directly design the thruster control voltage as a system input. Using Taylor's polynomial, the trajectory tracking error system can be transformed into a nonlinear system format with the control voltage as the input. An innovative method uses an adaptive sliding mode control approach incorporating a local recurrent neural network to estimate the unknown components of the system. In order to overcome the chattering problem associated with sliding mode control, a switching gain adjustment method based on an exponential function is proposed. The effectiveness of this approach is underlined by Lyapunov theory, which confirms that the trajectory tracking error remains uniformly bounded.

Another application of artificial neural networks in underwater navigation is the development of an adaptive path planning system[18], [19]. Based on real-time environmental data and potential obstacles, such a system uses ANNs to dynamically adjust the AUV's route. By integrating data from sensors such as sonar, Doppler velocity logs and inertial measurement units, the ANN can predict and respond to changes in the underwater environment, optimising the AUV's trajectory for efficiency and safety. This adaptive path planning system enhances the AUV's ability to navigate complex underwater terrain and maintain high performance in various conditions, such as strong currents or low visibility. The implementation of this ANN-based system significantly improves the operational autonomy and reliability of the AUV, making it an invaluable tool for underwater exploration, scientific research and environmental monitoring.

Enhancing the self-localisation capabilities of autonomous underwater vehicles is another notable application of artificial neural networks in underwater navigation[20], [21], [22]. Given the limitations of traditional global positioning systems and the challenging nature of the underwater environment, ANNs can be used to improve the accuracy and efficiency of localisation based on on-board sensors. By integrating data from multiple sensors, such as inertial measurement units, depth sensors and sonar, ANNs can effectively process and filter noisy and dynamic sensory information. This approach addresses the shortcomings of conventional multi-sensor fusion algorithms, which struggle with non-Gaussian noise and computational inefficiencies. ANNs can learn complex patterns and relationships within the sensor data, enabling more robust and real-time localisation. This advancement not only improves autonomous navigation of underwater robots, but also reduces computational costs, making it a critical development for underwater exploration and surveillance missions.

The use of neural networks to identify models for underwater vehicles represents a significant advance in dealing with uncertainty in dynamic systems[23], [24]. Instead of using neural networks in parallel with known models to broadly account for unmodeled phenomena, approaches are proposed that use domain-specific knowledge to target the most uncertain parts of the model. For example, neural networks are used to accurately identify the damping characteristics of underwater vehicles, which are notoriously difficult to model due to their complex interactions with the fluid environment. The effectiveness of this neural network based identification is validated through simulations where the neural networks are integrated into a feedforward controller. These simulations highlight the advantages of neural networks in online learning and demonstrate their robustness in scenarios with noisy measurements and varying dynamics, such as those caused by changes in the vehicle's toolskid. This method enhances the accuracy and adaptability of the model, which is crucial for the reliable operation of underwater vehicles in diverse and challenging environments.

A major step forward in improving the autonomous capabilities of unmanned underwater vehicles is the integration of neural networks into their obstacle avoidance planning[25], [26]. There are proposed approaches that use a recurrent neural network with convolution to improve obstacle avoidance planning. The method reduces the number of parameters while improving feature

extraction capabilities by incorporating convolution instead of full connection in the standard recurrent neural network. By generating training and test data sets and integrating multibeam forward-looking sonar data, the proposed learning system can autonomously perform obstacle avoidance for UUVs in real time.

6. References

- [1] A. Sahoo, S. K. Dwivedy, and P. S. Robi, "Advancements in the field of autonomous underwater vehicle," *Ocean Engineering*, vol. 181, 2019, doi: 10.1016/j.oceaneng.2019.04.011.
- [2] S. N. Shome *et al.*, "Development of Modular Shallow Water AUV: Issues & Trial Results," *Journal of The Institution of Engineers (India): Series C*, vol. 93, no. 3, 2012, doi: 10.1007/s40032-012-0026-0.
- [3] D. Munteanu, D. Moina, C. G. Zamfir, Ștefan M. Petrea, D. S. Cristea, and N. Munteanu, "Sea Mine Detection Framework Using YOLO, SSD and EfficientDet Deep Learning Models," *Sensors*, vol. 22, no. 23, 2022, doi: 10.3390/s22239536.
- [4] G. Mattei *et al.*, "Multi-Method Technics and Deep Neural Networks Tools on Board ARGO USV for the Geoarchaeological and Geomorphological Mapping of Coastal Areas: The Case of Puteoli Roman Harbour," *Sensors*, vol. 24, no. 4, 2024, doi: 10.3390/s24041090.
- [5] G. Bucci, "Hydrographic Drone Non-Invasive Underwater Investigations: New Archaeological Discoveries in Valle Fossa di Porto (Comacchio—FE, Italy)," *Heritage*, vol. 6, no. 2, 2023, doi: 10.3390/heritage6020096.
- [6] A. Birk *et al.*, "Dexterous Underwater Manipulation from Onshore Locations: Streamlining Efficiencies for Remotely Operated Underwater Vehicles," *IEEE Robot Autom Mag*, vol. 25, no. 4, 2018, doi: 10.1109/MRA.2018.2869523.
- [7] N. Palomeras *et al.*, "I-AUV docking and intervention in a subsea panel," in *IEEE International Conference on Intelligent Robots and Systems*, 2014, doi: 10.1109/IROS.2014.6942870.
- [8] D. Youakim, P. Ridao, N. Palomeras, F. Spadafora, D. Ribas, and M. Muzzupappa, "MoveIt!: Autonomous Underwater Free-Floating Manipulation," *IEEE Robot Autom Mag*, vol. 24, no. 3, 2017, doi: 10.1109/MRA.2016.2636369.
- [9] J. Guo, J. Wang, and Y. Bo, "An Observer-Based Adaptive Neural Network Finite-Time Tracking Control for Autonomous Underwater Vehicles via Command Filters," *Drones*, vol. 7, no. 10, 2023, doi: 10.3390/drones7100604.
- [10] P. Bałazy, P. Gut, and P. Knap, "Positioning algorithm for AGV autonomous driving platform based on artificial neural networks," *Robotic Systems and Applications*, vol. 1, no. 2, 2021, doi: 10.21595/rsa.2021.22109.
- [11] F. Gul, W. Rahiman, and S. S. Nazli Alhady, "A comprehensive study for robot navigation techniques," *Cogent Engineering*, vol. 6, no. 1, 2019, doi: 10.1080/23311916.2019.1632046.
- [12] L. Buşoniu, T. de Bruin, D. Tolić, J. Kober, and I. Palunko, "Reinforcement learning for control: Performance, stability, and deep approximators," *Annual Reviews in Control*, vol. 46, 2018, doi: 10.1016/j.arcontrol.2018.09.005.
- [13] O. Elhaki, K. Shojaei, and P. Mehrmohammadi, "Reinforcement learning-based saturated adaptive robust neural-network control of underactuated autonomous underwater vehicles," *Expert Syst Appl*, vol. 197, 2022, doi: 10.1016/j.eswa.2022.116714.
- [14] R. Hernandez-Alvarado, L. G. Garcia-Valdovinos, T. Salgado-Jimenez, A. Gomez-Espinosa, and F. F. Navarro, "Self-tuned PID control based on backpropagation Neural Networks for underwater vehicles," in *OCEANS 2016 MTS/IEEE Monterey, OCE 2016*, 2016, doi: 10.1109/OCEANS.2016.7761412.
- [15] E. Dong, S. Guo, X. Lin, X. Li, and Y. Wang, "A neural network-based self-tuning PID controller of an autonomous underwater vehicle," in *2012 IEEE International Conference on Mechatronics and Automation, ICMA 2012*, 2012, doi: 10.1109/ICMA.2012.6283262.
- [16] R. Hernández-Alvarado, L. G. García-Valdovinos, T. Salgado-Jiménez, A. Gómez-Espinosa, and F. Fonseca-Navarro, "Neural network-based self-tuning PID control for underwater vehicles," *Sensors (Switzerland)*, vol. 16, no. 9, 2016, doi: 10.3390/s16091429.
- [17] M. J. Zhang and Z. Z. Chu, "Adaptive sliding mode control based on local recurrent neural networks for underwater robot," *Ocean Engineering*, vol. 45, 2012, doi: 10.1016/j.oceaneng.2012.02.004.
- [18] F. García-Córdova and A. Guerrero-González, "Intelligent navigation for a solar powered unmanned underwater vehicle," *Int J Adv Robot Syst*, vol. 10, 2013, doi: 10.5772/56029.
- [19] A. Guerrero-González, F. García-Córdova, and J. Gilabert, "A biologically inspired neural network for navigation with obstacle avoidance in autonomous underwater and surface vehicles," in *OCEANS 2011 IEEE - Spain*, 2011, doi: 10.1109/Oceans-Spain.2011.6003432.
- [20] H. Xing *et al.*, "A Multi-Sensor Fusion Self-Localization System of a Miniature Underwater Robot in Structured and GPS-Denied Environments," *IEEE Sens J*, vol. 21, no. 23, 2021, doi: 10.1109/JSEN.2021.3120663.
- [21] Y. Wu, Y. Wei, and D. An, "Research on Multi-sensor Information Fusion Method of Underwater Robot Based on Elman Neural Network," in *2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design, CSCWD 2022*, 2022, doi: 10.1109/CSCWD54268.2022.9776294.
- [22] U. Ali, W. Muhammad, M. J. Irshad, and S. Manzoor, "Multi-sensor fusion for underwater robot self-localization using PC/BC-DIM neural network," *Sensor Review*, vol. 41, no. 5, 2021, doi: 10.1108/SR-03-2021-0104.
- [23] P. W. J. van de Ven, T. A. Johansen, A. J. Sørensen, C. Flanagan, and D. Toal, "Neural network augmented identification of underwater vehicle models," *Control Eng Pract*, vol. 15, no. 6, 2007, doi: 10.1016/j.conengprac.2005.11.004.
- [24] P. J. G. Lisboa, V. S. Kodogiannis, and J. Lucas, "Neural network identification and control of an underwater vehicle," *Transactions of the Institute of Measurement & Control*, vol. 19, no. 4, 1997, doi: 10.1177/014233129701900405.
- [25] C. Lin, H. Wang, J. Yuan, D. Yu, and C. Li, "An improved recurrent neural network for unmanned underwater vehicle online obstacle avoidance," *Ocean Engineering*, vol. 189, 2019, doi: 10.1016/j.oceaneng.2019.106327.
- [26] C. Lin, H. Wang, M. Fu, J. Yuan, and D. Yu, "A convolution neural network based obstacle avoiding method for unmanned underwater vehicle," *ICIC Express Letters*, vol. 13, no. 11, pp. 1079–1086, 2019, doi: 10.24507/icicel.13.11.1079.

Върху някои аспекти на „отворените данни“

Николай Нетов¹ Камен Спасов²
 Стопански факултет
 Софийски университет „Св. Кл. Охридски“, София, България^{1,2}
 nnetoff@feb.uni-sofia.bg¹, kspassov@feb.uni-sofia.bg²

Резюме: Големият интерес, който предизвика ChatGPT, постави акцент не само върху големите езикови модели /Large language models (LLM)/. Всеки един LLM трябва да бъде обучен върху голям обем данни и най-често това са различни масиви от данни с отворен достъп. Ускореното въвеждане в практиката на предварително обучени LLM фокусира значителен интерес и по темата за „отворените данни“. Нашето изследване описва състоянието и възможностите за достъп до национални „отворени данни“ в България. Използвахме редица изследвания, като тези на Световната банка, PARIS21 и Open Data Watch - Open Data Inventory (ODIN), за да анализираме доколко „пълни“ са наличните национални „отворени данни“ в България и дали те отговарят на международните стандарти за откритост. Резултатите за покритие се основават на наличието на ключови типове и групи от данни, за които са достъпни и подходящи разбивки както във времето, така и по отношение на различни географски подразделения. Анализът за отвореност се основава на това дали и доколко данните са достъпни за изтегляне, използване и повторно използване в машинно четими и непатентовани формати, дали са налични в необходимия обем и качество свързаните с тях метаданни, дали съществуват опции за изтегляне като групово изтегляне или API и доколко са приложими за тях отворени условия за използване или лиценз за данни. В заключение са направени и някои предложения за подобряване и разириряване на достъпа и свързването на „отворени данни“ за изследователски и бизнес цели.

Ключови думи: отворени данни, отворени данни в България, изкуствен интелект, индикатори за отворени данни, прозрачност

On some aspects of "open data"

Nikolay Netov¹, Kamen Spassov²
 FEBA, Sofia University "St. Kl. Ohridski", Sofia, Bulgaria^{1,2}
 nnetoff@feb.uni-sofia.bg¹, kspassov@feb.uni-sofia.bg²

Abstract: The great interest caused by ChatGPT put emphasis not only on Large Language Models (LLM). Every single LLM must be trained on a large amount of data, and most often it is a variety of open access data sets. The accelerated deployment of pre-trained LLMs has also focused considerable interest on the topic of 'open data'. Our research describes the state and possibilities of access to national "open data" in Bulgaria. We used data from The World Bank, PARIS21 and Open Data Watch - Open Data Inventory (ODIN) to analyze how "complete" is the available national "open data" in Bulgaria and whether it meets international standards for openness. Coverage results are based on the availability of key types and groups of data for which appropriate breakdowns are available, both over time and across different geographic divisions. An openness analysis is based on whether and to what extent the data are available for download, use and reuse in machine-readable and non-proprietary formats, whether the associated metadata is available in the required volume and quality, whether download options such as bulk downloads exist or APIs, and to what extent open terms of use or data license are applicable to them. In conclusion, some proposals are also made to improve and expand access to and linking of "open data" for research and business purposes.

Keywords: open data, Bulgaria open data, artificial intelligence, open data indicators, transparency

1. Въведение

Протичащите през последните години ускорени тенденции за дигитализация и дигитална трансформация във всички елементи на икономиката и политическото управление демонстрират по впечатляващ начин потенциала на управляваните от данни приложения за машинно обучение и технологиите с елементи на изкуствен интелект (ИИ). Актуалните през последните години технологии като анализ и моделиране на големи данни, отворени правителствени данни и ИИ (Big Data and Open Government Data Analytics and Artificial intelligence - BD&OGDA&AI) и тяхната интеграция в организационните практики и процедури имат значително въздействието върху икономиката, обществения живот, прозрачността на политическото управление и ефективността на административните и бизнес процеси [1-6].

Интегрирането на отворени инфраструктури за данни ще даде възможност на AI да се използва в почти всяка област на икономическия и обществен живот. Прилагането на методологии за задълбочено обучение на ИИ и споделянето на код и данни все още са в ранен етап и изискват разработването на повече стандарти. Много от моделите все още се обучават от нулата с помощта на вътрешни сървъри и графични процесори, което отнема време и ограничава процеса [7].

Днес прозрачността се счита за ключова характеристика на доброто управление. Почти безспорно концепцията е придобила важност както в академичните среди, така и в средите на практиките поради силната си присъща и/или инструментална стойност. През последните години се наблюдават засилващи се позитивни тенденции по отношение на глобалния, регионалния и националния капацитет на държавите да създават, споделят, публикуват и преизползват статистически и „отворени данни“, както и нарастващото им въздействие върху правителствените политики [8].

В свое актуално изследване Krah, R. D. Y., & Mertens, G. установяват забележителен ръст в броя и качеството на изследванията, посветени на темата за прозрачността на местното управление, като тя е най-силно застъпена по отношение на местните власти в Европа и Северна Америка. Авторите на това изследване установяват, че социалните, икономическите, политическите и институционалните фактори обясняват тази тенденция. Наблюдавана е добра практика на местните власти с използването на информационните технологии да бъде свързвано и повишаване на прозрачността на политическото управление. Натискът местните власти да бъдат прозрачни до голяма степен се свързва с приемането на закони за отворени данни [9].

2. Методика за анализ на състоянието

Съществува обективна необходимост всяка отделна държава да може безпристрастно да проследява напредъка си по отношение на възможностите на бизнеса и държавните административни органи да създават, споделят, публикуват и преизползват „отворени данни“ по удобен и достъпен начин. Важно е да се постигне консенсус и около общовъзприети референтни показатели, с помощта на които да могат да се оценяват и анализират обхватът и отвореността на официалните статистически данни.

Всички официални данни за дадена държава, публикувани директно на специализираните държавни платформи за „отворени данни“ или на платформите за достъп до официална статистическа информация на съответните национални статистически институти и държавни административни органи (включително местни власти), трябва да бъдат оценени като част от този процес на оценка.

Много добър пример в това отношение е инициативата PARIS21 (<https://paris21.org/>), която стартира през април 2019 г. и има за цел да опише тенденциите в глобалния, регионалния и националния статистически капацитет. Инициативата PARIS21 наблюдава и публикува информация за около 100 индикатора, а интуитивният интерактивен интерфейс прави географското и секторно сравняване и анализ бързи и лесни. Резултатите от мониторинга се класифицират в пет основни групи:

- Планиране на процесите по осигуряване на пълни, точни и актуални данни;
- Производство: процесът на конструиране на масивите от данни;
- Разпространение: начините за достъп до и за широкото разпространение на наличните данни;
- Употреба: грамотност на потребителите на данни;
- Инвестиции: отделяне на финансови и други ресурси за поддръжка и развитие на съществуващите платформи.

Голяма част от показателите са с годишна периодичност. Потребителите могат да визуализират данните и да кръстосват различни видове индикатори. Изследването обхваща 242 държави.

Друг добър пример в това отношение е инициативата за инвентаризацията на отворени данни (ODIN) на Open Data Watch (<https://opendatawatch.com/>), която измерва доколко изчерпателни са наличните в дадена държава платформи за достъп до „отворени данни“ и дали те отговарят на международните стандарти за отвореност. ODIN оценява обхвата и отвореността на официалните статистически данни, за да идентифицира пропуски, да насърчи политики за отворени данни, да подобри достъпа и да улесни диалога между националните статистически служби и потребителите на данни. Резултатите за покритие се определят въз основа на наличието на ключови индикатори с подходящи разбивки във времето и за географски подразделения. Резултатите за отвореност измерват дали данните могат лесно да бъдат изтеглени в машинно четими и непатентовани формати, поддържат ли се с метаданни, имат множество опции за изтегляне като групово изтегляне и избор на потребител или API и притежават отворени условия за използване или лицензи за данни. Актуалното изследване за периода 2022/23 г. обхваща 195 държави.

Световната банка (The World Bank) също събира, анализира и публикува актуална информация за тенденциите в глобалния, регионалния и националния капацитет на държавите да създават, споделят, публикуват и преизползват статистически и „отворени данни“. Методическата рамка на Световната банка (<https://www.worldbank.org/en/programs/statistical-performance-indicators/Framework>) оценява зрелостта и ефективността на

националните статистически системи в пет ключови области, наречени стълбове. Петте стълба са: (1) Използване на данни, (2) Услуги за данни, (3) Продукти с данни, (4) Източници на данни и (5) Инфраструктура за данни. Всеки от тези стълбове се поддържа от четири или пет измерения и използва дефинирани и налични като „отворени данни“ индикатори. Актуалното изследване за 2022 г. обхваща 217 държави.

3. Анализ на състоянието

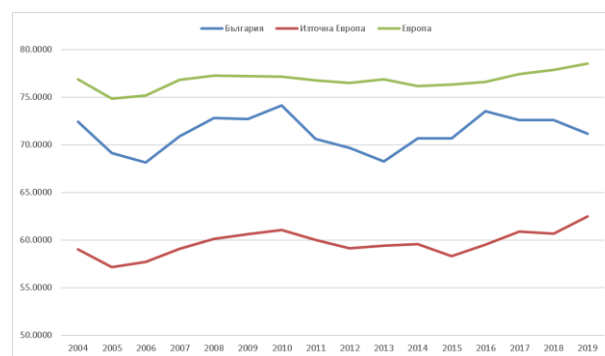
PARIS21 Statistical Capacity Monitor, базиран на WGI, е изготвен от Daniel Kaufmann (Институт за управление на природните ресурси и Brookings Institution) и Aart Kraay (Изследователска група за развитие на Световната банка).

PARIS21 индикаторът за регулаторно качество оценява способността на отделните правителства да формулират и прилагат стабилни политики и разпоредби, които позволяват и насърчават развитието на частния сектор. В табл. 1 са представени данните за България, Източна Европа и Европа като цяло в периода от 2004 до 2019 г.

Таблица 1: Индикатор за регулаторно качество

Година	България	Източна Европа	Европа
2004	72.4138	59.0148	76.8966
2005	69.1177	57.1569	74.8565
2006	68.1373	57.6961	75.1674
2007	70.8738	59.0777	76.8056
2008	72.8155	60.1456	77.2792
2009	72.7273	60.6220	77.2319
2010	74.1627	61.0526	77.1735
2011	70.6161	60.0000	76.7657
2012	69.6682	59.1469	76.4883
2013	68.2464	59.3839	76.8928
2014	70.6731	59.5673	76.1609
2015	70.6731	58.3173	76.3602
2016	73.5577	59.5192	76.5830
2017	72.5962	60.9135	77.4508
2018	72.5962	60.6731	77.8494
2019	71.1539	62.5000	78.5413

Фигура 1 илюстрира динамиката на индикатора за регулаторно качество в периода от 2004 до 2019 г.



Фиг. 1. Индикатор за регулаторно качество

Това, което можем да отбележим по отношение на индикатора за регулаторно качество в България, е липсата на ясно изразена тенденция. Двете му най-ниски нива са през 2006 г. (68.1373) и през 2013 г. (68.2464). Нивата на този индикатор за страните от Източна Европа остават трайно по-ниски, с ясно изразена тенденция за покачване през последните години. Същата тенденция наблюдаваме и по отношение на данните за Европа като цяло, където и стойностите на индикатора са по-високи от тези за България през целия наблюдаван период.

ODIN индикаторът за отвореност на данните се базира на Open Data Watch - Open Data Inventory. Този индикатор се базира на следните пет показателя за отвореност на данните:

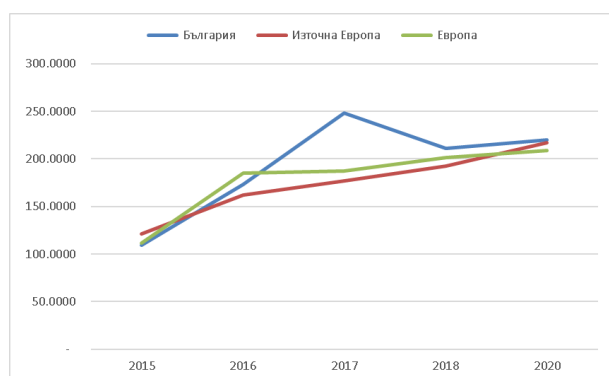
(1) машинна четимост, (2) непатентован формат; (3) опции за изтегляне, (4) наличност на метаданни и (5) безплатни и неограничени условия за ползване.

В табл. 2 са представени данните за България, Източна Европа и Европа като цяло в периода от 2015 до 2020 г.

Таблица 2: ODIN Индикатор за отвореност на данните

Година	България	Източна Европа	Европа
2015	108.9000	121.4571	111.1250
2016	172.8000	161.8636	184.8450
2017	248.4000	176.9727	187.6071
2018	210.9000	192.7200	201.0805
2020	219.6000	217.3200	208.8214

Фигура 2 илюстрира динамиката на индикатора за отвореност на данните в периода от 2015 до 2020 г.



Фиг. 2. ODIN Индикатор за отвореност на данните

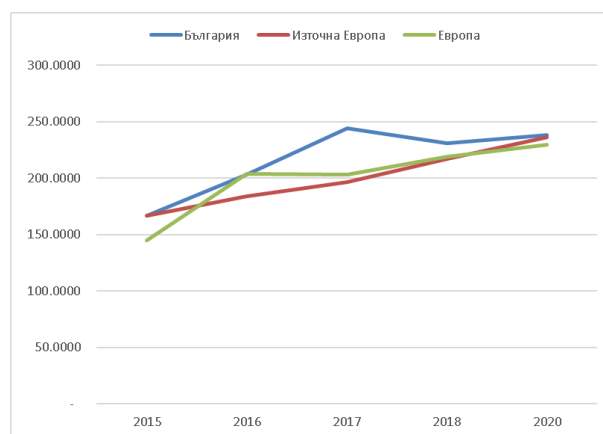
По отношение на индикатора за отвореност на данните за България наблюдаваме аналогична липса на ясно изразена тенденция. Докато данните за Източна Европа и за Европа като цяло показват ясно изразена тенденция за нарастване, в България този показател е на най-ниското си ниво в началото на наблюдавания период и изостава както спрямо Източна Европа, така и спрямо Европа като цяло. В периода от 2015 до 2017 г. за България се наблюдава изпреварващата тенденция за нарастване по отношение на този показател и той достига значително по-високи от средноевропейските стойности (248.4000). През 2018 г. имаме бързо връщане към средноевропейските нива и последваща тенденция за бавно покачване, като нивото на индикатора се доближава до средните европейски нива.

PARIS21 индикаторът за икономическа статистика се базира на Open Data Watch - Open Data Inventory. Този индикатор се базира на следните шест елемента на откритост: (1) Национални сметки, (2) Статистика на труда, (3) Индекси на цените, (4) Държавни финанси, (5) Пари и банково дело, (5) Международна търговия, (6) Платежен баланс. В табл. 3 са представени данните за България, Източна Европа и Европа като цяло в периода от 2015 до 2020 г.

Таблица 3: Индикатор за икономическа статистика

Година	България	Източна Европа	Европа
2015	166.8000	166.8857	145.1250
2016	203.1000	183.8182	203.6250
2017	244.5000	196.8273	203.2000
2018	231.0000	216.8700	219.3732
2020	238.5000	236.3700	229.5357

Фигура 3 илюстрира динамиката на индикатора за икономическа статистика в периода от 2015 до 2020 г. Както се вижда на фиг. 3, индикаторът за икономическа статистика следва същата динамика, каквато наблюдавахме по отношение на индикатора за отвореност на данните.



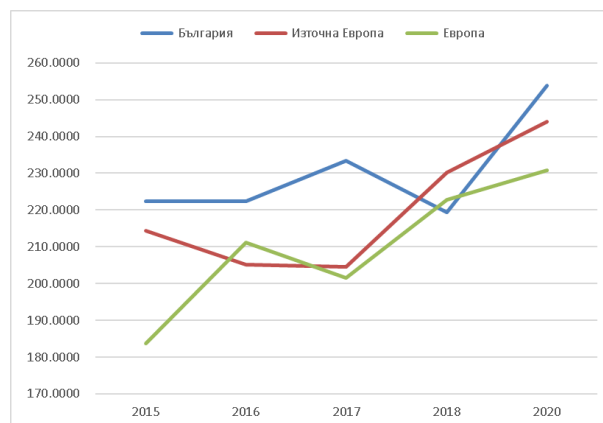
Фиг. 3. Индикатор за икономическа статистика

Индикаторът за обхвата на данните от икономическата статистика се базира на Open Data Watch - Open Data Inventory. В табл. 4 са представени данните за България, Източна Европа и Европа като цяло в периода от 2015 до 2020 г.

Таблица 4: Индикатор за обхвата на данните от икономическата статистика

Година	България	Източна Европа	Европа
2015	222.3000	214.3286	183.8000
2016	222.3000	205.0636	211.2000
2017	233.4000	204.5727	201.5571
2018	219.3000	230.1900	222.8049
2020	253.8000	243.9900	230.8000

Фигура 4 илюстрира динамиката на индикатора за обхвата на данните от икономическата статистика в периода от 2015 до 2020 г.



Фиг. 4. Индикатор за обхвата на данните от икономическата статистика

Както се вижда на фиг. 4, динамиката по отношение на индикатора за обхвата на данните от икономическата статистика е доста по-разнообразна. По отношение на този индикатор България трайно има по-добро представяне от средните за Източна Европа и за Европа като цяло нива с едно малко изключение през 2018 г., когато се наблюдава и най-ниската му стойност (219.3) за целия анализиран период. За Източна Европа наблюдаваме намаляващи стойности на индикатора до 2017 г., последвани от бърза тенденция за нарастване през следващите години и най-висока стойност през 2020 г. Стойностите на индикатора за обхвата на данните от икономическата статистика за Европа като цяло стартират от

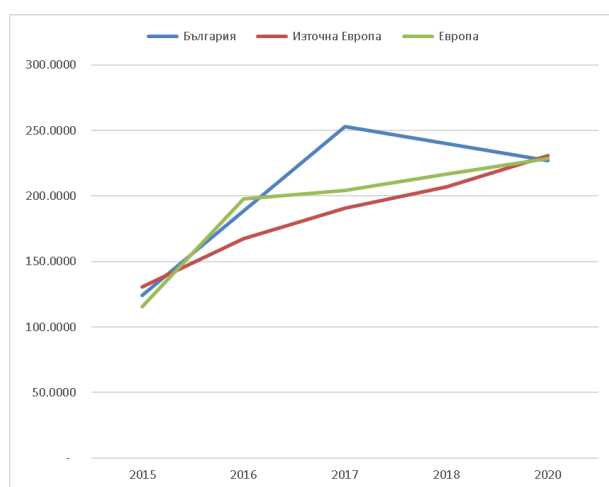
много ниско ниво (183.8) през 2015 г. и се наблюдава трайна тенденция на бързо нарастване до ниво от 230.8 през 2020 г. с едно малко изключение през 2017 г. Като цяло по отношение на обхвата на данните от икономическата статистика Източна Европа и България в частност показват по-добро представяне от средното за Европа.

Индикаторът за отвореност на икономическите данни се базира на Open Data Watch - Open Data Inventory. В табл. 5 са представени данните за България, Източна Европа и Европа като цяло в периода от 2015 до 2020 г.

Таблица 5: Индикатор за отвореност на икономическите данни

Година	България	Източна Европа	Европа
2015	124.2000	130.4143	115.3750
2016	188.7000	167.5909	197.8125
2017	252.9000	190.8818	204.4714
2018	240.0000	206.9700	216.7902
2020	227.1000	230.5500	228.6714

Фигура 5 илюстрира динамиката на индикатора за отвореност на икономическите данни в периода от 2015 до 2020 г.



Фиг. 5. Индикатор за отвореност на икономическите данни

По отношение на индикатора за отвореност на икономическите данни за България отново наблюдаваме липса на ясно изразена тенденция. Докато данните за Източна Европа и за Европа като цяло показват ясно изразена тенденция за нарастване, в България този показател е на най-ниското си ниво в началото на наблюдавания период и изостава по отношение на Източна Европа. В периода от 2015 до 2017 г. за България се наблюдава изпреварващата тенденция за нарастване по отношение на този показател и през 2017 г. той достига значително по-висока стойност от средноевропейските (252.9000). След 2018 г. имаме бързо връщане към средноевропейските нива, като през 2020 г. стойността на индикатора (227.1) вече е по-ниска от тази за Европа като цяло (228.7) и за Източна Европа в частност (230.6).

4. Дискусия

Технологиите с елементи на ИИ привличат все повече внимание поради техния потенциал да подобрят ефективността на административните и бизнес процеси. Превръщането на технологичните перспективи в осезаема обществена стойност за бизнеса, за държавните и местните администрации е изправено пред съществени предизвикателства [10]. Основните пречки за успешното използване на технологиите с елементи на ИИ, независимо от характера на различните организации, са свързани не само с липсата на достатъчно вътрешноорганизационни технологични умения, но и с юридически и нормативни ограничения, с ограничени възможности за реинженеринг на установените вече

административни и бизнес процеси, с ограничения по отношение на бързия и ефективен достъп на системите с ИИ до големи данни и отворени данни.

5. Изводи и заключения

Организациите трябва да се съсредоточат не само върху изграждане на капацитет за създаване, обучение и внедряване на системи с елементи на ИИ, но и върху изграждане на такъв за работа с големи данни и най-вече с отворени данни. Наличието на национални отворени данни с широк обхват, високо качество и отвореност в машинно четим формат е предпоставка за бързото и ефективно въвеждане на технологии с ИИ в практиката. Анализът на състоянието на отворените данни в България, и конкретно на икономическите данни, показва, че имаме по-високо ниво на покритие в сравнение със средното за Европа и в момента се наблюдава положителна тенденция в това отношение. За съжаление, по отношение на показателя за отвореност на икономическите данни за България се наблюдава точно обратна тенденция.

Благодарности

Публикацията е във връзка с резултати от работата ни по Проект № 80-10-154 / 17.4.2024 г.

Библиография

- [1] Hossain, M. A., Dwivedi, Y. K., & Rana, N. P., „State-of-the-art in open data research: Insights from existing literature and a research agenda, “*Journal of organizational computing and electronic commerce*, Vol. 26, № 1-2, pp. 14-40, 2016.
- [2] Afful-Dadzie, Eric, and Anthony Afful-Dadzie, „Liberation of public data: Exploring central themes in open government data and freedom of information research.,“ *International Journal of Information Management*, Vol. 37, № 6, pp. 664-672, 2017.
- [3] Altman, M., Wood, A., O'Brien, D. R., Vadhan, S., & Gasser, U., „Towards a modern approach to privacy-aware government data releases, “*Berkeley Technology Law Journal*, Vol. 30, № 3, pp. 1967-2072, 2015.
- [4] A. Cerrillo- i- Martinez, „The Reuse of Public Sector Information in Europe and Its Impact on Transparency.,“ *European Law Journal*, Vol. 18, № 6, pp. 770-792, 2012.
- [5] Dawes, S. S., Vidasova, L., & Parkhimovich, O., „Planning and designing open government data programs: An ecosystem approach, “*Government Information Quarterly*, Vol. 33, № 1, pp. 15-27, 2016.
- [6] Khurshid, M. M., Zakaria, N. H., Rashid, A., Ahmad, M. N., Arfeen, M. I., & Faisal Shehzad, H. M., „Modeling of open government data for public sector organizations using the potential theories and determinants—a systematic review“, *Informatics*, Vol. 7, № 3, p. 24, 2020.
- [7] Henning Otto Brinkhaus, Kohulan Rajan, Jonas Schaub, Achim Zielesny, Christoph Steinbeck, „Open data and algorithms for open science in AI-driven molecular informatics, “*Current Opinion in Structural Biology*, Vol. 79, p. 102542, 2023.
- [8] Cahlikova, T., & Mabillard, V., „Open data and transparency: Opportunities and challenges in the Swiss context, “*Public Performance & Management Review*, Vol. 43, № 3, pp. 662-686, 2020.
- [9] Krah, R. D. Y., & Mertens, G., „Transparency in local governments: patterns and practices of twenty-first century, “*State and Local Government Review*, Vol. 52, № 3, pp. 200-213, 2020.
- [10] Colin van Noordt, Luca Tangi, „The dynamics of AI capability and its influence on public value creation of AI within public administration, “*Government Information Quarterly*, Vol. 40, № 4, p. 101860, 2023.

Preparation and Deployment of Secure Over-The-Air Updates for Embedded Devices: Challenges and Solutions

Pavel Palurik*, Michal Mikulasek, Lukas Jabloncik
Brno University of Technology, Brno, Czech Republic
xpalur01@vutbr.cz

Abstract: This paper provides a comprehensive overview of the preparation and deployment of secure Over-The-Air (OTA) updates for embedded devices. In the era of intelligent devices, the need for remotely updating software and firmware has become increasingly prevalent. This paper discusses the principles of OTA updates, focusing on the various types used in the field of embedded devices and the complications that need to be resolved for a successful process. Furthermore, it delves into the theoretical description of OTA updates, security mechanisms, and practical implementation. The paper also highlights the main benefits of OTA updates, such as convenience, security, scalability, reduced maintenance costs, and improved performance. Additionally, it addresses the challenges and issues associated with OTA updates, including update reliability, security risks, compatibility, network propagation, and integrity and warranty implications. Overall, this paper serves as a valuable resource for understanding the complexities and importance of OTA updates in the realm of embedded devices.

Keywords: APPLICATION, ESP32, OTA, SECURITY, UPDATE, WIRELESS

1. Introduction

In today's era of intelligent devices, ranging from simple sensors to complex communication units, it is typical for new functions not originally included in the device to be enhanced and added during use. Therefore, suppliers need to address the issue of remotely updating the software (SW) and/or firmware (FW) of the device so that it is always a safe operation for the user without the possibility of damaging or destroying the device. This process is commonly referred to as Over the Air (OTA) or Remote Update. The fundamental principle of this approach is that the manufacturer, when manufacturing and initially installing the device software, prepares the device to support future software updates via commonly used communication technologies (typically IEEE 802.11 Wi-Fi or IEEE 802.15.1 Bluetooth). Ideally, this makes the update process transparent to the end user, who may not even be aware that their smart devices are receiving new software, including new features, according to the manufacturer's capabilities and policy.

A key requirement for a successful process is that the manufacturer must always provide the correct version of the device update. At the same time, the entire update process must be managed, from the preparation of the update itself to the delivery of the update to the device. This is to ensure that at no point during the process can the device or the process enter a state that could irreversibly damage or destroy the device itself.

During the above steps, there is a risk of cyber-attack, typically by competitors or persons interested in gaining the manufacturer's know-how and access to the manufacturer's system. As a result, much emphasis is now being placed not only on the update itself, but also on the entire OTA update process.

To clarify the picture, the reader will be introduced to the principle of OTA updates, supplemented by a description of the most common types used in the field of embedded devices. At the same time, the complications that need to be resolved during OTA in order to complete the whole process successfully are mentioned.

The second section focuses on the theoretical description of OTA updates. The third section focuses on the description of the security mechanisms, followed by the fourth section which discusses the practical implementation. A summary is given in the fifth section.

2. Over The Air updates

Over The Air updates are software or firmware processes in embedded devices that allow remote and wireless updates without the need to physically connect the device to a computer or special programming tool. This technology is a key feature on the Internet of Things (IoT) and Industry 4.0, as it allows devices to be remotely managed, maintained and enhanced in functionality [1, 2].

The main benefits of OTA updates are:

- **Convenience and flexibility:** Allows manufacturers to update software features or fix bugs in devices without requiring physical user intervention,
- **Security:** OTA updates can fix security flaws and vulnerabilities, helping to keep devices secure and protected against attacks,
- **Scalability:** Manufacturers can easily distribute updates to a large number of devices simultaneously, which is key for IoT ecosystems,
- **Reduced maintenance costs:** Allows you to minimize costs associated with physical maintenance and device updates,
- **Improve performance:** Updates can include optimization and performance improvements to devices.

From the list above, it is clear that OTA updates are a key tool in maintaining extensive systems with large numbers of devices. However, as the intensity of use increases, so does the number of security incidents that can lead to security breaches or complete device downtime. As a result, the following issues are increasingly being addressed:

- **Update reliability:** OTA updates must be reliable to avoid unwanted outages or errors during the process,
- **Security risks:** Security risks associated with OTA updates include the possibility of attacks on devices during the update process or the distribution of compromised updates,
- **Compatibility:** Updates should be designed to be compatible with different versions of hardware and software on different devices,
- **Network propagation:** Distributing updates over wireless networks can be challenging, especially in areas with poor coverage,
- **Integrity and warranty:** Updates can affect the integrity of the device and may have legal implications for warranty terms.

2.1 Principles

A remote update is a series of interdependent actions that collectively determine whether the process is successful or not. The result of each action is evaluated, and if it is deemed to have failed, the entire process is considered a failure, and the previous version of the software or firmware is reinstalled on the device.

2.1.1 Reception of updates

In order to implement the update on the end device, it is necessary to receive data for the subsequent update. The most commonly used application communication protocol is HTTP (Hypertext Transfer Protocol) in the secure variant of HTTPS in cooperation with TLS (Transport Layer Security). Through this secure tunnel, the end devices receive information about the available update and, if necessary, the update data itself (all depends

on the specific implementation of the device manufacturer). Given that end devices typically lack oversized operation memory, the received data is transmitted in smaller packets to ensure the devices can safely and reliably receive and store the data in their persistent memory before the actual update application begins. Once the device has received all the update data packets, a series of verification steps are performed to ensure the integrity of the data, to confirm the source of the data, and to subsequently decrypt the received update data (encrypted data are exclusively transmitted in production environments).

2.1.2 Application of updates

Once the device has received all the necessary update data and completed the authentication and decryption sequence, the actual firmware/software update process can commence. The manner in which the update application process is conducted is contingent upon the policy established by the device manufacturer. The two principal categories are full and differential updates [1, 3].

A less frequently employed approach in the context of small embedded systems is the differential update, which is predicated on the assumption that the alterations between FW/SW versions are comparatively minor in comparison to the entirety of the system. Consequently, it is deemed more efficient to transmit solely the differences between the two versions. For these updates, the manufacturer must calculate and generate a change file for each new update. This file contains records of what and where the new FW version differs from the original. In order for devices to be able to handle this information, the entire calculation is conducted at the level of the firmware/software binary files of the device. The most significant advantage of this update approach is the reduction in bandwidth resources due to the need to send only changes between versions. This is a highly advantageous feature for communication networks with limited bandwidth, such as LoRaWAN. Conversely, a significant drawback of this update system is the necessity to calculate changes between multiple different versions in order to ensure that devices with different versions of the current system can update under the most favorable conditions. An alternative approach would be to transfer change updates from the current version of the device to the most recent version. Consequently, this approach may be less efficient than transferring the entire update in a single file [3]. In summary, this approach to updates is suitable under the conditions that updates are always performed across the board, and it is not possible for any device to be out of date by more than two updates. This is because of the inefficiency of transfer and the potential for network overload.

The second most prevalent type of update is the so-called "full" OTA update. The fundamental principle of this update is to completely overwrite the firmware/software of the end device with the new version. The indisputable advantage of this approach is that the update distributor is not required to address the potential for collisions between different system versions, and the same file can be distributed to all endpoints. Conversely, a significant disadvantage is the higher demand on network transmission resources through the need to deliver the entire system update file to each device. Furthermore, in the event of a full update, there are a number of approaches that can be employed on the endpoint side with respect to the memory allocation of the device in order to deal with the update itself. The available options are as follows:

- Updates that are not reversible to the original system version (only one preceding functional version of the system is always available), see Figure 1 with limited OTA support,
- Updates with the ability to revert to the default system version (memory space is partitioned so that space is allocated for the originally loaded system and then several other updates), see Figure 1 with extended OTA support.

The primary distinction between the two options is that the first essentially represents a rotation of system versions in a linear sequence (assuming the availability of two memory locations for system images).

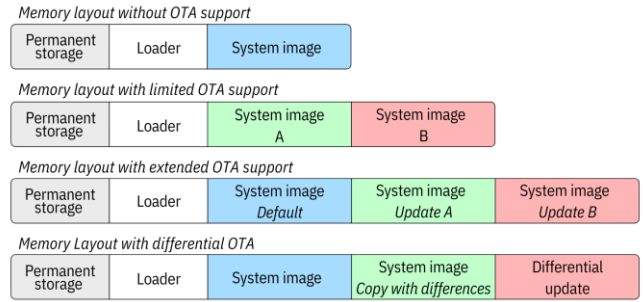


Fig. 1 Display of embedded device memory layout for different OTA update schemes.

In contrast, the second variant enables the user to revert to the default system version that was installed on the device at its manufacturing stage, in addition to the features of the first option. This approach is highly effective in ensuring that the endpoint device always has a secure point in case of failed remote updates. The fundamental differences between the two approaches are also evident in the memory requirements, as the version without the original system image has the potential to save the space of a single system image [1, 3]. This feature is frequently employed in small, simple embedded systems, where the acquisition of smaller memory units through mass production results in significant cost savings.

The process of booting a device with the application of different system images is illustrated in Figure 2. This figure considers the case of an OTA update with limited capabilities, which is most often applied through memory resource reduction.

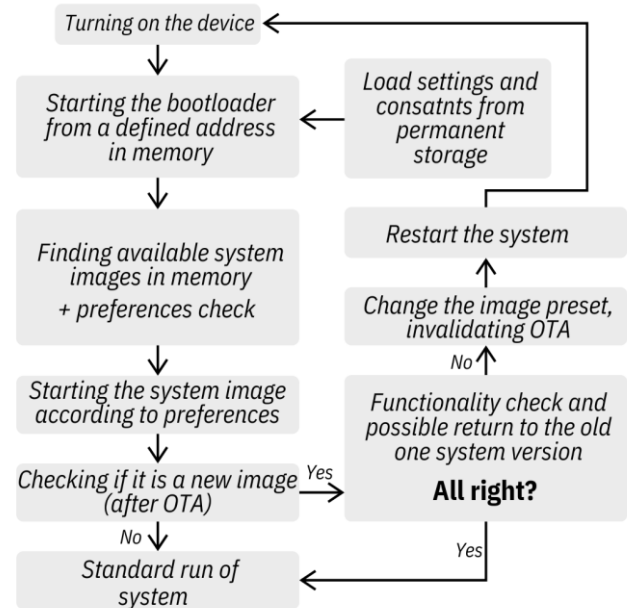


Fig. 2 Simplified state diagram of embedded system startup with OTA update [4, 5, 6].

2.2 Associated risks

OTA updates are prone to various interruptions and failures that can compromise device functionality and user experience, necessitating effective crisis handling mechanisms to ensure devices can recuperate seamlessly from such occurrences. One common issue in OTA updates is the interruption of downloads, often due to network instability or power loss. To mitigate this risk, devices can implement resumable download protocols, allowing the update process to resume from the point of interruption rather than restarting entirely. This is typically achieved by maintaining a persistent state of the download progress in non-volatile memory.

Additionally, malfunctions in the OTA update process, such as corruption or incorrect installation of the downloaded package, require robust verification mechanisms like cryptographic hash

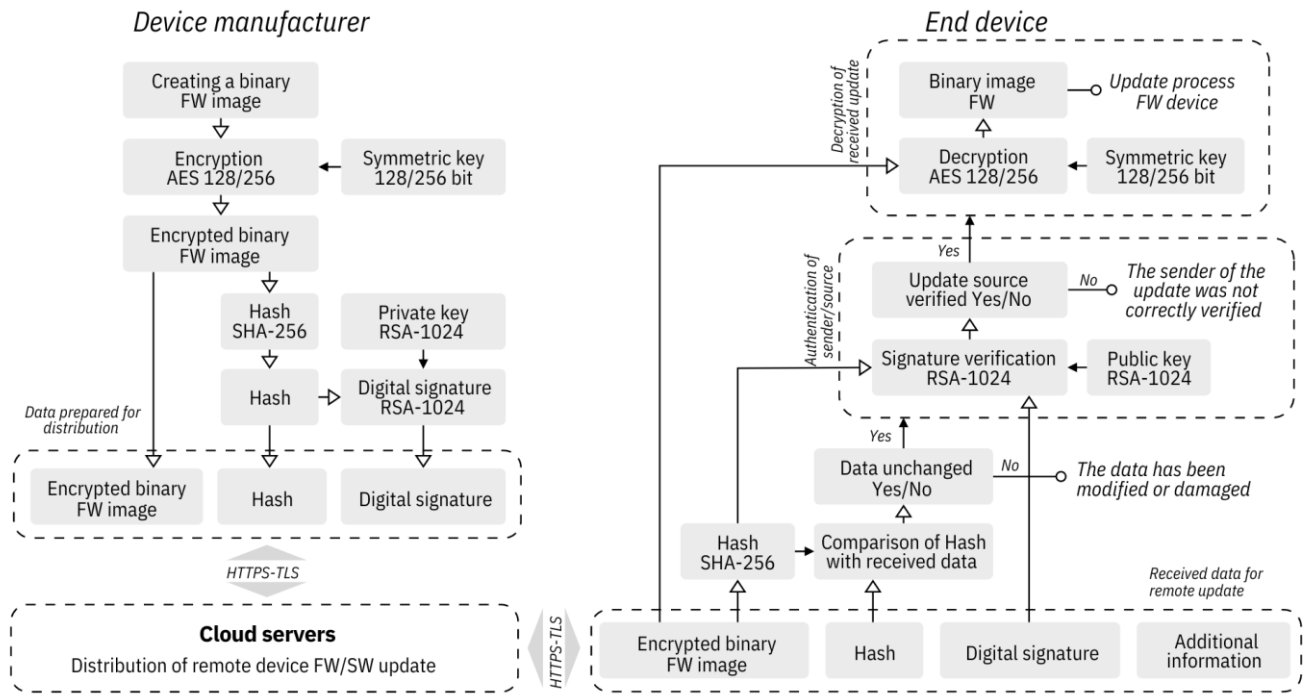


Fig. 3 Simplified view of the remote update data transfer to the embedded device [7, 4, 5, 6].

checks and digital signatures to validate the integrity and authenticity of the update package before installation. The implementation of atomic updates, where the previous software version is replaced only once the new version has been fully validated and is ready to run, can prevent the system from being left in an inconsistent state.

In scenarios where an OTA update introduces critical issues or fails to complete successfully, the ability to revert to a previous software version is imperative. This capability, often facilitated by dual-image systems (see Figure 1), ensures that if the new update fails, the device can revert to the fallback version, maintaining continued functionality. This approach requires careful management of available memory to accommodate both images and ensure the fallback version remains intact during the update process.

Efficient memory management is crucial for the successful application of OTA updates. Given the limited memory resources in devices, the process of updating firmware or software must be carefully orchestrated to avoid memory overflows and ensure data integrity. Devices often use dedicated partitions for the update process, dividing the device's memory into segments such as active, inactive, and update partitions. The update partition temporarily stores the new software image during download and verification, and once the update is confirmed to be intact, the device can switch from the old active partition to the new one. This structured approach ensures a seamless and reliable OTA update process.

3. Security of Over The Air updates

OTA update security is a key factor without which no device in use today can be deployed in a production environment. In the case of remote update of embedded devices, security is addressed on several levels. Typically, these include:

- Security of the FW/SW update binary image,
- Secure transfer and storage of the binary image to a distribution server from the device manufacturer,
- Secure communication from the distribution server to the end device,
- Security of the end-device.

The security of the binary image on the manufacturer side is managed by a number of processes. The complete procedure is

illustrated in Figure 3. The initial process is the encryption of the binary image using symmetric cryptography, specifically AES-128 (Advanced Encryption Standard) or AES-256, which is the most commonly employed variant. This allows for hardware acceleration on the majority of embedded processors and microcontrollers currently available. Subsequently, a hash is extracted from the encrypted image (typically SHA-256, Secure Hash Algorithm), which serves two purposes. Firstly, it provides a small, unidirectional value that can be used by the recipient to verify the integrity of the data against alteration or error during transmission. Secondly, it is used in a digital signature that is then used by the recipient to verify the authenticity of the data source. The most common method for obtaining the digital signature is asymmetric cryptography using RSA (Rivest, Shamir, Adleman) in the RSA-1024 variant, given the limited computing resources. In general, embedded devices are not well-suited for extensive use of asymmetric cryptography. When circumstances necessitate its use, a less robust but less demanding variant is often selected as a compromise between performance and security.

Secure transmission and storage on a distribution server is typically achieved through a secure connection to a cloud service, typically Amazon AWS (Amazon Web Services), which offers off-the-shelf solutions for all common variants.

In the case of the communication of the distribution server with the end devices, the situation is more complex due to the fact that there are several commonly used communication protocols for the Internet of Things (IoT). The most prevalent communication protocol is MQTT (Message Queuing Telemetry Transport), which offers secure communication via TLS. Other protocols in use include CoAP (Constrained Application Protocol), LWM2M (Light Weight Machine To Machine), etc. These typically implement similar measures to MQTT, which is currently the de facto standard. For the sake of simplicity, the referenced communication will be referred to as MQTT in the following. The process of passing information and transmitting the update itself is based on the functionality of the device. A continuous listen is maintained for any messages with information regarding a new update or a periodic poll is conducted to ascertain whether a more up-to-date image is available. The choice of communication protocol and possible update queries is entirely at the discretion of the device manufacturer. The optimal approach is a combination of the two,

which resolves the issues of ensuring that the device has the most up-to-date software as quickly as possible and maintaining an active connection to the server through cyclic queries, including on intermediate active network elements.

Upon receipt of information regarding a new image, the end-device initiates the download and caching process. A communication protocol may be employed to transfer the data itself; however, due to the considerable inefficiency, the HTTP protocol in the HTTPS variant with TLS is most often chosen.

When the device finishes receiving the new update data, it starts the validation process, which involves several steps. The first step is to calculate a hash similar to the one after encrypting a binary image, in order to compare the received hash and the calculated one. If they match, then it can be declared that the received data has not been altered or corrupted. If no match is found, the device should terminate the update process and inform the distribution server of the failure via a secure channel (encrypted MQTT) with the mentioned reasons. Then the digital signature verification is performed, which involves the need to keep in memory the computed Hash, the public key of the source (distribution server or manufacturer) and the digital signature itself. This data is then verified using asymmetric cryptography (RSA-1024) and if the signature matches, then it can be declared that the source of the update and data is the real source and is not a spoof. In the last step, the binary image of the update is decrypted using symmetric cryptography (AES-128 or AES-256) and a pre-shared key to obtain a binary image suitable for insertion into the prepared memory space.

These steps are followed by the actual application and verification process of the update, which is depicted in simplified form in Figure 4 [4-9].

4. Practical implementation of secure OTA in IoT

The Espressif ESP32 platform [10], was selected for the demonstration. This platform represents the most commonly used representative of a smart, wirelessly communicating IoT unit in the current era. The correct setup and application of the secure OTA update will be performed for the variant of connection via IEEE 802.11 to the public network (Internet), which is the most commonly used in this case. In order to demonstrate the baseline, the ESP IDF (Integrated Development Framework) development tool was employed [11].

The actual process of implementing remote OTA updates in this system can be divided into several steps in order to provide a higher level of clarity.

4.1 Preparation of Distribution Server Root Certificate and Encryption Keys

Prior to commencing the security implementation, it is essential to provide several security keys and a certificate as described in section 3, which will subsequently be required in the FW and without which the secure version of the update would not be possible.

The first certificate to be considered is the public certificate of the distribution server, also known as the server root certificate. This will be employed to authenticate the distribution server when the HTTP-TLS connection is established. This secure transmission channel will then be used to facilitate the secure update of data [12]. An alternative approach is to utilise the so-called "ESP X509 Certificate Bundle", which is a pre-packaged set of the most commonly used root certificates, with the intention of simplifying the implementation of the solution and relieving the manufacturer of the responsibility of verifying the certificates' validity.

The second key is a symmetric key used to decrypt the received binary image of the system in the final phase of the OTA update [13]. The ESP IDF environment lacks a framework for this aspect, rendering it wholly within the purview of the manufacturer. The manufacturer may select one of several options for key preparation

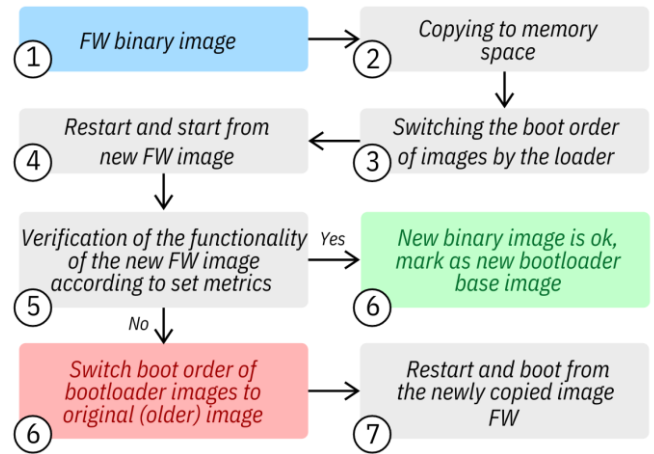


Fig. 4 Simplified view of the process of applying a remote update to an embedded device [4, 5, 6].

and sharing, contingent upon the change policy of the key in question. The first option is to prepare the key at the time of device creation and insert it into the secure memory of the device. This option has the advantage of being faster (the key is ready for immediate use), but it also carries a significant security risk.

An attacker may be able to break the protection of the secure memory, thereby obtaining a symmetric key and potentially being able to decrypt the binary image of the system. The indisputable advantage is that the manufacturer is able to require the key to be changed over time and insert a new key into the system via a remote update, thereby overwriting the old, potentially invalid and insecure key. A more complex alternative is to negotiate a symmetric key using asymmetric DH cryptography (Diffie-Hellman algorithm). The ESP IDF environment supports a standard variant of this DH algorithm as well as modifications in the form of Elliptic Curve Diffie-Hellman (ECDH).

4.2 Process of update

To perform a remote update, the ESP IDF environment includes a prebuilt ESP OTA library [14], which offers the option of using it without security for testing purposes outside the production environment. This involves omitting the source/server authentication certificate and optionally the unencrypted binary image. In order to utilize the library in a production environment with recommended security, it is necessary to specify the server root certificate in PEM (Privacy-enhanced Electronic Mail) format.

Once the certificate has been prepared, the actual update procedure can be initiated. This can be accomplished in two distinct manners, whether the manufacturer is interested in monitoring the entire process of connection establishment, data download, and update point-by-point implementation, or alternatively, be satisfied to simply "perform the update."

The automated process is designed for the fundamental form of update, which assumes a Uniform Resource Locator (URL) containing a binary image and a root certificate. The entire process of establishing a connection through the HTTPS server, requesting the binary data and then saving the data is fully automatic, and the user application is not required to deal with additional handler code. The sole exception to this is when the binary image is encrypted. In such a case, the "CONFIG_ESP_HTTPS_OTA_DECRYPT_CB" menu must be enabled, which serves to enable the code that is responsible for supporting the decryption of the binary data received from the server. The decryption may be initiated during the download process, as one of the packets is downloaded, or at the end when of the entire binary image is downloaded. The decryption process is not specified in any way in the case of ESP libraries. Consequently, the form of encryption implemented by the device manufacturer is entirely at their discretion. This option allows for the dynamic switching between a simplified encryption method for

battery-powered devices and a more robust form of encryption for plug-in devices.

The second option is to utilize existing functions from the "esp_https_ota" library, which must then be called and served in the correct function within the user application, that is to say, the application code. The sole advantage of this solution is that the user can monitor the status of the update and potentially react in real time (e.g. decryption error, inappropriate binary image, etc.).

According to the available documentation, it is generally recommended to use the automatic form, which contains the main steps of the update, including automatic saving, changing the boot priority of the system image, and automatic restart.

5. Conclusion

The theoretical introduction provided a synopsis of the fundamental principles associated with remote OTA updates, including an analysis of the key advantages and disadvantages. Moreover, the text delineates the fundamental tenets that must be observed during remote updating to guarantee security.

To provide a comprehensive understanding of the subject matter, the practical component then proceeded to illustrate the implementation of a secure remote update for an ESP 32 microcontroller, which is the most commonly utilized for the majority of smart home applications globally.

The demonstration illustrates that each manufacturer may have a distinct implementation of remote update within their respective development environments, regardless of the prescribed standards. Consequently, it is imperative to ascertain the current status prior to commencing the implementation of remote updates, after which the actual work may begin.

References

- [1] K. G. Crowther, R. Upadrashta, and G. Ramachandra, "Securing over-the-air firmware updates (fota) for industrial internet of things (iiot) devices," in 2022 IEEE International Symposium on Technologies for Homeland Security (HST), 2022, pp. 1–8.
- [2] J. Bauwens, P. Ruckebusch, S. Giannoulis, I. Moerman, and E. D. Poorter, "Over-the-air software updates in the internet of things: An overview of key principles," IEEE Communications Magazine, vol. 58, no. 2, pp. 35–41, 2020.
- [3] V. Nikic, D. Bortnik, M. Lukic, D. Danilovic, and I. Mezei, "Comparisons of firmware delta updates over the air using wlan and lpwan technologies," in 2022 30th Telecommunications Forum (TELFOR), 2022, pp. 1–4.
- [4] I. Benjamin Bucklin Brown Analog Devices, "Over-the-air (ota) updates in embedded microcontroller applications: Design trade-offs and lessons learned," online. [Online]. Available: <https://www.analog.com/en/analog-dialogue/articles/over-the-air-ota-updates-in-embedded-microcontroller-applications.html>
- [5] S. Laboratories, "Kba bt 0804: Secure ota dfu," online. [Online]. Available: https://community.silabs.com/s/article/kba-bt-0804-secure-ota-dfu?language=en_US
- [6] S. Schmidt, "Secure firmware updates in the iot," online, University of Applied Sciences Campus Vienna, Tech. Rep., 2023. [Online]. Available: https://sec4dev.io/assets/uploads/slides/Secure-Firmware-Updates-OTA-in-the-IoT_2.pdf
- [7] N. Lethaby, "A more secure and reliable ota update architecture for iot devices," online, Texas Instruments, Tech. Rep., 2018. [Online]. Available: <https://www.ti.com/lit/wp/sway021/sway021.pdf>
- [8] T. Authors, "Documentation: Over-the-air firmware and software updates," online. [Online]. Available: <https://thingsboard.io/docs/user-guide/ota-updates>
- [9] M. T. Inc., "Over-the-air updates," online. [Online]. Available: <https://www.microchip.com/en-us/products/wireless-connectivity/over-the-air-updates>
- [10] E. S. Company, "Esp32," online. [Online]. Available: <https://www.espressif.com/en/products/socs/esp32>
- [11] "Esp-idf programming guide," online. [Online]. Available: <https://docs.espressif.com/projects/esp-idf/en/latest/esp32/>
- [12] "Esp-tls," online. [Online]. Available: https://docs.espressif.com/projects/esp-idf/en/latest/esp32/api-reference/protocols/esp_tls.html
- [13] "Esp-https-ota," online. [Online]. Available: https://docs.espressif.com/projects/esp-idf/en/latest/esp32/api-reference/system/esp_https_ota.html
- [14] "Over the air updates (ota)," online. [Online]. Available: <https://docs.espressif.com/projects/esp-idf/en/latest/esp32/api-reference/system/ota.html>

Trends in non-linear MIMO Objects Control in the Era of Industry 4.0: The Use of Artificial Neural Networks

Patryk Bałazy^{1,*}, Urszula Jachymczyk¹
AGH University of Krakow, Poland¹
balazy@agh.edu.pl

Abstract: The Industry 4.0 revolution has significantly influenced the control of non-linear Multiple Input Multiple Output (MIMO) systems, particularly through the application of artificial neural networks (ANNs). This paper explores current trends in the control of non-linear MIMO objects, emphasizing the role of ANNs in enhancing performance and efficiency. Key developments, methodologies, and case studies are reviewed to illustrate the impact of ANNs on non-linear MIMO control

Keywords: NON-LINEAR MIMO SYSTEMS, ARTIFICIAL NEURAL NETWORKS, CONTROL SYSTEMS, AUTOMATION, INTELLIGENT CONTROL, SMART MANUFACTURING, SYSTEM OPTIMIZATION.

1. Introduction

The advent of Industry 4.0 has ushered in an era of smart manufacturing and intelligent automation [1], driving significant advancements in control systems for non-linear Multiple-Input, Multiple-Output (MIMO) objects. This new industrial revolution is characterized by the integration of advanced digital technologies, such as the Internet of Things (IoT) [2], big data analytics, artificial intelligence (AI), and cloud computing, into traditional manufacturing processes. These innovations are fundamentally transforming how industries operate, leading to more efficient, flexible, and responsive production systems. Traditional control methods often struggle with the complexity and interdependencies inherent in non-linear MIMO systems. These systems, which are prevalent in various industrial applications [3-8] such as robotics, manufacturing, and automotive systems, pose significant challenges due to their multiple interacting inputs and outputs, as well as their non-linear behaviors. As a result, conventional control strategies may fail to deliver the required performance, robustness, and adaptability needed in dynamic and complex industrial environments [9]. This paper examines how Artificial Neural Networks (ANNs), a core component of machine learning, are being leveraged to address these challenges. ANNs are computational models capable of learning and adapting through experience. By utilizing ANNs, control systems can develop adaptive and robust solutions that are better suited to handle the intricate dynamics of non-linear MIMO systems. ANNs can approximate complex functions and model non-linear relationships between inputs and outputs, making them ideal for controlling systems with significant non-linearities and interactions.

2. Non-linearities in crane control

Non-linearities in crane control systems arise from various sources [10-12] such as friction, gear backlash, load sway, and the non-linear dynamics of the crane structure. These non-linearities complicate the control process, making it difficult to achieve precise positioning and stability. Identifying and understanding these non-linearities is crucial for developing effective control strategies.

For instance, friction in the crane's moving parts can introduce hysteresis effects, where the system's response depends on its previous states. This hysteresis effect complicates the control of crane movements, as the frictional forces vary with the direction and magnitude of the motion, making it difficult to predict and compensate for them accurately.

Gear backlash, another source of non-linearity, refers to the play or looseness between mating gears. This backlash can lead to a loss of motion precision, as the crane's control inputs might not result in immediate or proportionate movements. The delay and slack caused by backlash can hinder the crane's ability to perform smooth and precise operations, especially in tasks requiring high accuracy.

Load sway is a common issue in crane operations, arising from the pendulum-like motion of the suspended load. This sway can cause significant instability, making it challenging to accurately

position loads. The oscillatory nature of load sway requires the control system to continuously adjust and counteract the motion, which is particularly difficult given the non-linear and dynamic characteristics of the swinging load.

The non-linear dynamics of the crane structure itself, including flexible components, changing sling length and varying load distributions, further add to the complexity. Cranes often have large, flexible arms or booms that can bend and twist under different loads and movements. These deformations introduce additional non-linearities, as the crane's structural response changes with varying load positions and weights. Additionally, the distribution of the load can shift during operations, altering the center of gravity and affecting the crane's stability and control.

Traditional linear control methods often fall short in addressing these issues, as they are typically based on simplifying assumptions that do not hold in the presence of significant non-linearities. These methods may provide acceptable performance under certain conditions but can fail when the system operates outside the assumed linear range. As a result, there is a need for more sophisticated approaches that can handle the complexities of non-linear crane dynamics. One such approach is the adoption of Artificial Neural Networks (ANNs). ANNs are capable of learning and modeling complex non-linear relationships through training on data. By leveraging ANNs, control systems can develop adaptive strategies that better account for the non-linear behaviors of friction, backlash, load sway, and structural dynamics. ANNs can continuously learn and update their control policies based on real-time data, providing more robust and precise control for crane operations. This capability makes ANNs particularly suited for applications where traditional linear control methods are insufficient, offering a path towards improved performance and reliability in crane control systems.

3. Applied regulation solutions using ANNs

Traditional control methods often struggle to address these complexities effectively. Artificial Neural Networks (ANNs) offer a powerful tool for addressing these challenges and improving crane operation in the era of Industry 4.0. This chapter explores how ANNs are being utilized to enhance crane control in several key areas.

3.1. Sensorless Anti-Swing Control

Traditional anti-swing control relies on sensors to detect the payload swing angle. However, sensor placement can be cumbersome and expensive, particularly for cranes with complex hoisting mechanisms. Additionally, these sensors can be susceptible to noise and require regular maintenance. Wahyudi and Yusof [13] propose a sensorless anti-swing control system using ANNs. Their approach utilizes a multilayer feedforward network trained with a backpropagation learning algorithm. This network acts as a "soft sensor," estimating the swing angle based on readily available data from other crane sensors (e.g., motor positions, accelerometers). This approach offers a cost-effective and reliable method for mitigating payload sway.

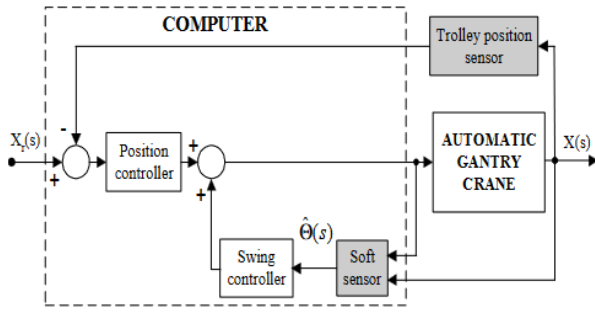


Fig. 1 Sensorless anti-swing system. - source [12]

The benefits include:

- **Reduced Cost:** Eliminates the need for dedicated swing angle sensors.
- **Simplified Installation:** No need for complex sensor placement and wiring.
- **Increased Reliability:** Less susceptible to sensor failures and maintenance issues.

3.2. Adaptive Control for Uncertainties

The dynamics of a crane system can be significantly affected by uncertainties such as external disturbances, where wind, uneven floors, or building vibrations introduce unpredictable forces; unmodeled dynamics, where complexities in the crane's mechanical structure are difficult to capture perfectly in mathematical models; and varying environmental conditions, where temperature fluctuations affect friction coefficients and other parameters. The paper [14] discusses a novel adaptive hierarchical sliding mode control scheme proposed for 3D overhead cranes. This scheme utilizes Radial Basis Function Networks (RBFNs) to estimate uncertain parameters within the control system. The RBFN acts as a powerful function approximator, dynamically adapting to changing conditions.

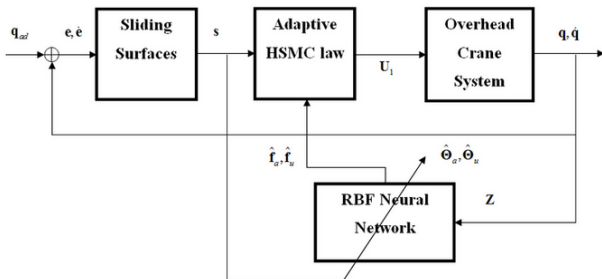


Fig. 2 Adaptive control for uncertainties system. - source [13]

The benefits include:

- **Robustness to Uncertainties:** The control system can adjust to unforeseen disturbances and model inaccuracies.
- **Improved Performance:** The crane maintains optimal performance even under varying operating conditions.
- **Enhanced Safety:** Reduced risk of control errors due to unexpected dynamics.

3.3. Inverse Dynamic Control with ANNs

Crane operation often suffers from vibrations due to the system's non-linear behaviour. These vibrations can lead to reduced positioning accuracy, where the load may not reach its intended destination precisely; increased stress on machinery, as vibrations can accelerate wear and tear on crane components; and operator discomfort, since rough crane movements can be uncomfortable and

potentially unsafe for the operator. This chapter explores a new strategy based on Inverse Dynamic Control Simulation (IDCS) using ANNs [15]. This approach utilizes an ANN to learn the inverse dynamic model of the crane. The inverse model essentially predicts the required control inputs to achieve a desired crane motion. The learned model is then used within the IDCS framework to generate suitable control signals that minimize vibrations and ensure accurate positioning. This method eliminates the need for extra vibration sensors and reduces noise within the feedback loop.

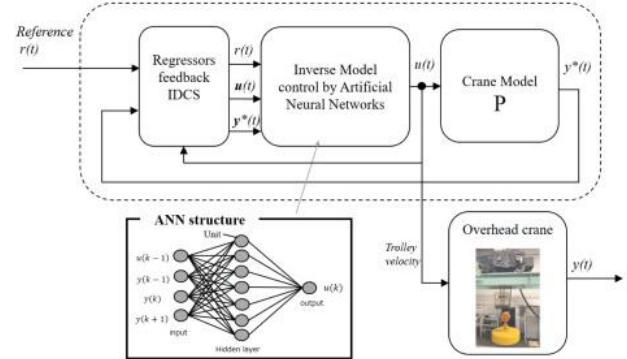


Fig. 3 Inverse Dynamic Control - source [14]

The benefits include:

- **Smoother Operation:** Reduced vibrations lead to a more precise and comfortable crane experience.
- **Improved Positioning Accuracy:** The load reaches its target location with greater accuracy.
- **Reduced Maintenance Costs:** Less wear and tear on crane components due to minimized vibrations.

The applications mentioned above highlight the potential of ANNs to revolutionize crane control in Industry 4.0. By offering sensorless solutions, adapting to uncertainties, and mitigating vibrations, ANNs pave the way for more efficient, reliable, and safer crane operation.

3.4. Reinforcement learning – based control

The authors of this paper propose their own novel crane control system based on artificial neural networks and the reinforcement learning (RL) method [16]. The system uses a neural feedback controller in which the ANN, which is called RL agent, determines feedback gains in real time. The adaptive nature of the system is due to the recalculation of the gains at each time sample, which allows it to account for varying environmental parameters. The authors are researching the stability of this type of system, developing a new control quality index dedicated to ANN-based control systems.

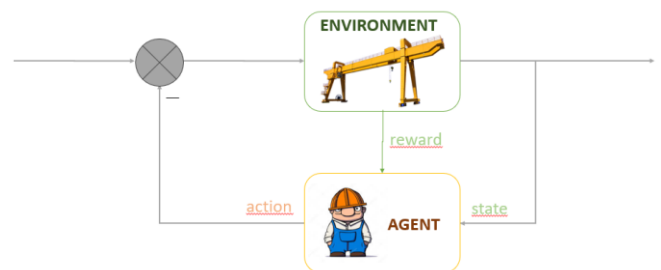


Fig. 4 RL-based system control.

The proposed control system consists of the following components:

- **Crane model** - It represents the dynamics of the crane by means of equations of motion.

- Artificial neural network - Acts as an RL agent and is responsible for determining optimal control actions based on observations of the state of the crane and the environment.
- Reinforcement function - Defines the reward or punishment for the RL agent for the actions taken.
- Learning mechanism - Controls the process of updating ANN parameters based on the RL method.

The ANN takes object's output (e.g., position, speed, acceleration) and determines the optimal control values based on this data. The gains are recalculated in each time sample to take into account the current operating conditions of the crane.

4. Results and Discussion

The application of ANNs in non-linear MIMO control has yielded promising results. Improved accuracy is one of the most significant benefits, with ANNs enhancing precision in positioning and movement control, especially in systems with high degrees of non-linearity. This improvement is attributed to the ANN's ability to model complex, non-linear relationships that traditional control methods cannot. Additionally, ANNs offer adaptive learning capabilities, allowing them to learn from new data and adapt to changing conditions in real-time. This adaptability is crucial in dynamic industrial environments where system parameters can vary frequently. The robustness of ANNs is another critical advantage. They are resilient to disturbances and uncertainties, making control systems more reliable and efficient. This robustness stems from the ANN's ability to generalize from training data and handle variations in input conditions. These benefits underscore the potential of ANNs to revolutionize non-linear MIMO control, aligning with the goals of Industry 4.0 to create smarter, more efficient manufacturing processes.

The integration of ANNs into non-linear MIMO control systems is not without challenges. One significant challenge is the computational complexity of training and deploying ANNs, which can require substantial computational resources and expertise. Despite these challenges, the benefits of ANNs in non-linear MIMO control are clear, and ongoing research and development efforts are focused on overcoming these obstacles. The future of ANNs in industrial control looks promising, with continued advancements in machine learning algorithms, computational power, and data collection techniques paving the way for more widespread adoption.

4. Conclusions

The application of Artificial Neural Networks (ANNs) offers transformative solutions for crane control in the context of Industry 4.0, addressing challenges that traditional methods struggle to overcome. Through sensorless anti-swing control, adaptive control for uncertainties, inverse dynamic control, and reinforcement learning-based control, ANNs enhance efficiency, reliability, and safety in crane operations. These advanced methods reduce costs, simplify installations, and improve robustness against disturbances and model inaccuracies, ultimately leading to smoother operations and better positioning accuracy. The promising results demonstrated by ANNs in non-linear MIMO control systems underscore their potential to revolutionize industrial control processes. While challenges such as computational complexity remain, ongoing research and technological advancements are likely to mitigate these issues, paving the way for the broader adoption of ANNs in industrial applications. The integration of ANNs aligns with the goals of Industry 4.0, promising smarter, more adaptive, and efficient manufacturing processes in the future. Future research should focus on further refining these techniques and exploring new applications to fully realize the potential of ANNs in industrial automation. As the field continues to evolve, the synergy between ANNs and Industry 4.0 technologies promises to drive significant improvements in industrial efficiency, productivity, and innovation. The authors of this article plan to develop a new control quality

indicator for control systems based on artificial neural networks. Many designers of control systems question those based on artificial neural networks due to the lack of clear proof of the stability of such systems. Hence the idea to develop a method for testing the stability of these systems.

6. References

1. Samigulin, T. I. M. U. R., and O. L. G. A. Shirayeva. "Development of a SMART-system for a complex industrial object control based on metaheuristic algorithms of swarm intelligence." *WSEAS Transactions on Power Systems* 16 (2021): 231-240.
2. Gong, Yi, et al. "Nonlinear MIMO for industrial Internet of Things in cyber-physical systems." *IEEE Transactions on Industrial Informatics* 17.8 (2020): 5533-5541.
3. Umair, S. Muhammad, and R. Usman. "Automation of irrigation system using ANN based controller." *International Journal of Electrical & Computer Sciences IJECS-IJENS* 10.02 (2010): 41-47.
4. Przepióra, Jan, et al. "Smart IoT platform for Big Data analysis in predictive maintenance." *2023 24th International Carpathian Control Conference (ICCC)*. IEEE, 2023.
5. Knap, Paweł, Krzysztof Lalik, and Patryk Bałazy. "Boosted Convolutional Neural Network Algorithm for the Classification of the Bearing Fault Form 1-D Raw Sensor Data." *Sensors* 23.9 (2023): 4295.
6. Bałazy, Patryk, Paweł Gut, and Paweł Knap. "Positioning algorithm for AGV autonomous driving platform based on artificial neural networks." *Robotic Systems and Applications* 1.2 (2021): 41-45.
7. Lalik, Krzysztof, et al. "Self-powered wireless sensor matrix for air pollution detection with a neural predictor." *Energies* 15.6 (2022): 1962.
8. Lalik, Krzysztof, et al. "Q-learning neural controller for steam generator station in micro cogeneration systems." *Energies* 14.17 (2021): 5334.
9. Yu, Lean, Shouyang Wang, and Kin Keung Lai. "A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates." *Computers & Operations Research* 32.10 (2005): 2523-2541.
10. Yu, J., F. L. Lewis, and T. Huang. "Nonlinear feedback control of a gantry crane." *Proceedings of 1995 American Control Conference-ACC'95*. Vol. 6. IEEE, 1995.
11. Chin, C., A. H. Nayfeh, and E. Abdel-Rahman. "Nonlinear dynamics of a boom crane." *Journal of Vibration and Control* 7.2 (2001): 199-220.
12. Fang, Yongchun, et al. "Nonlinear coupling control laws for an underactuated overhead crane system." *IEEE/ASME transactions on mechatronics* 8.3 (2003): 418-423.
13. Yusof, Nor Thahirah Mohamad. "ANN-based sensorless anti-swing control of automatic gantry crane systems: Experimental result." *2008 5th International Symposium on Mechatronics and Its Applications*. IEEE, 2008.
14. Le, Viet-Anh, et al. "An efficient adaptive hierarchical sliding mode control strategy using neural networks for 3D overhead cranes." *International Journal of Automation and Computing* 16.5 (2019): 614-627.
15. Rincon, Liz, et al. "Inverse dynamic control via "simulation of feedback control" by artificial neural networks for a crane system." *Control Engineering Practice* 94 (2020): 104203.
16. Podlasek, Szymon, et al. "Application of Ann Control Algorithm for Optimizing Performance of a Hybrid Orc Power Plant." Available at SSRN 4720421.

Application of convolutional networks to detect the operating phases of energy systems using a biomass boiler as an example

Szymon Podlasek^{1*}, Piotr Książek¹
AGH University of Krakow, Poland¹
podlasek@agh.edu.pl

Abstract: The development of neural algorithms opens new perspectives for the analysis of technological processes. Particularly relevant are strongly nonlinear and complex objects, such as power plants. One of the modern solutions enabling data analysis are convolutional neural networks (CNNs). The research presents the application of CNNs to monitor and optimize combustion processes in biomass boilers. The fuel analyzed was gray straw, which is difficult to control due to the nature of combustion. The proposed technique is based on the processing of temporal data, which represent different stages of the combustion process. The work examined the effectiveness of the model in identifying key operating parameters and detecting the stages of firing from ignition initialization to nominal operation. Analysis of images of parameter curves from the time waveforms makes it possible to capture repeatable relationships that enable faster response to future changes in the conditions of the combustion process. Determining the phase of the process, based on data and trends of selected parameters, allows the control system to react faster, without operator intervention. As a result of the study, the efficiency of process stage change detection by the convolutional network, expressed by means of an error matrix, through the F1-score parameter (harmonic mean between precision and sensitivity) was achieved at a level close to 96%. The proposed solution can be effectively applied to a number of technological processes including those that are part of Industry 4.0 effectively influencing technological transformation..

Keywords: RES, COMPUTER MODELING, CONVOLUTIONAL NEURAL NETWORKS, OPTIMIZATION ALGORITHMS, PREDICTION

1. Introduction

Modern biomass boilers play a crucial role in renewable energy production, forming an essential part of strategies to reduce greenhouse gas emissions (Fig.1). Biomass, as an energy source, is characterized by low costs and wide availability, making it an attractive alternative to fossil fuels. However, the biomass combustion process is complex and requires precise management of various variables, such as fuel moisture content, chemical composition, and boiler operating conditions.

One of the main challenges in biomass combustion is ensuring an efficient and stable combustion process that minimizes the emission of pollutants, such as nitrogen oxides (NO_x) and carbon monoxide (CO) [1-3]. Traditional methods of controlling the combustion process are often not flexible enough to respond to dynamic changes in combustion conditions. Effective management of this process requires advanced technologies capable of processing large amounts of data and adjusting operational parameters in real-time.

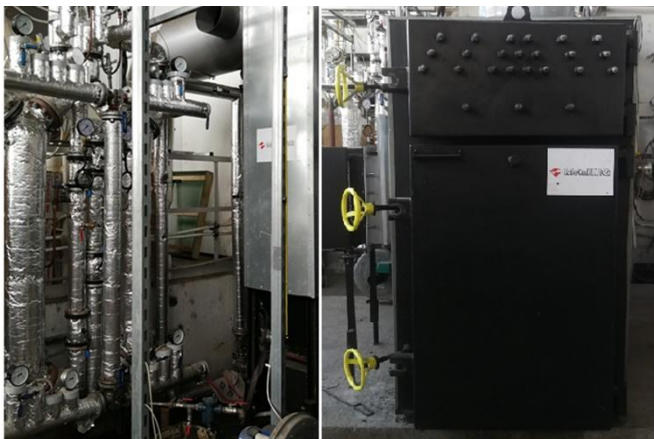


Fig. 1 Tested object - biomass boiler fueled by gray straw..

Convolutional Neural Networks (CNNs) are one such technology that is finding increasing application in the energy industry [4]. This method is increasingly used in Renewable Energy Sources (RES) to both forecast and support the operation of solar and wind installations [5, 6]. Also, research has been undertaken on the use of neural networks to reduce emissions from boilers by predicting components of the combustion process [7, 8]. CNNs, a type of deep neural network, are particularly effective in analyzing spatial data structures, such as images [9-11], making them ideal for

analyzing operational data from biomass boilers or prediction of the occurrence of an event such as air pollution [12].

This article presents the application of CNNs for analyzing and predicting combustion processes in biomass boilers. Operational data from actual boilers and numerical simulations were used to train and test the CNN model. The objective of the study was to optimize the combustion process and reduce pollutant emissions. The results show that the application of CNNs can significantly improve combustion efficiency and reduce the negative environmental impact. The approach presented can serve as a basis for future research on the application of advanced machine learning technologies in optimizing industrial processes. Additionally, this solution aligns with Industry 4.0 standards through advanced analysis of measurement data.

2. Prerequisites and means for solving the problem

The work analyzed the possibility of using neural networks to detect the combustion stage. The turbulent nature of straw combustion requires a rapid response to a change in a parameter, which calls for the study of new algorithms for controlling these processes. In the study, it was decided to analyze the time courses of boiler parameters using CNNs.

Convolutional Neural Networks are a special class of deep neural networks particularly effective in processing grid-like data structures, such as images. Convolutional networks can also be used to analyze and interpret time series data, where the dependent variable, such as the intensity of electrical energy flow, is recorded at regular intervals. Such data can be treated analogously to images, where the time axis corresponds to the width of the image, and the variable values represent pixel intensities. By training on historical or simulated data, a CNN can learn to recognize recurring patterns in variable profiles and predict planned changes or anomalies.

The confusion matrix, also known as an error matrix or confusion table, is a tool used in machine learning, such as CNNs, to evaluate the quality of classification. It is a table that allows for the visualization of the performance of a classification algorithm. The confusion matrix shows the numbers of true positive, false positive, true negative, and false negative results (Fig. 2).

These four terms are key to calculating classifier performance metrics, such as Precision, Recall, Specificity, Accuracy, and F1-Score. The formulas for these indicators are presented in equations (r1) to (r5).

$$\text{Precision} = \frac{TP}{TP + FP} \quad (r1)$$

Precision measures what percentage of positive predictions is actually positive.

$$Recall = \frac{TP}{TP + FN} \quad (r2)$$

Recall measures what percentage of actual positive cases was correctly identified.

$$Specificity = \frac{TN}{TN + FP} \quad (r3)$$

Specificity measures what percentage of actual negative cases was correctly identified.

$$Accuracy = \frac{TN + TP}{TN + TP + FP + FN} \quad (r4)$$

Accuracy measures what percentage of all classifications (positive and negative) is correct.

$$F1 - Score = \frac{Precision \cdot Recall}{Precision + Recall} = \frac{\frac{TP}{TP + FP} \cdot \frac{TP}{TP + FN}}{\frac{TP}{TP + FP} + \frac{TP}{TP + FN}} \quad (r5)$$

F1-Score is the harmonic mean of precision and recall, providing a single measure of classifier accuracy.

In the context of change detection, CNNs can be used to analyze data from sensors or time records, identifying patterns that are unusual or indicate potential problems. The confusion matrix in this context would serve to evaluate the effectiveness of models in detecting and correctly classifying these changes.

TRUE NEGATIVES ⇒ Predicted negative result ⇒ Actual negative result	FALSE POSITIVES ⇒ Predicted positive result ⇒ Actual negative result
FALSE NEGATIVES ⇒ Predicted negative result ⇒ Actual positive result	TRUE POSITIVES ⇒ Predicted positive result ⇒ Actual positive result

Fig. 2. Example representation of a confusion matrix for a case with two possible states and the assessment of their proper detection based on actual and predicted values [13].

3. Solution of the examined problem

As part of the work, experimental data from straw combustion were used as input data for analyzing combustion runs. The process was carried out under the same conditions, the only difference was the air feeding point to introduce some variability in the data. In addition, modeling and simulation work was carried out to increase the amount of information for the neural network. A full analysis of the impact of location on the process is available in article [14].

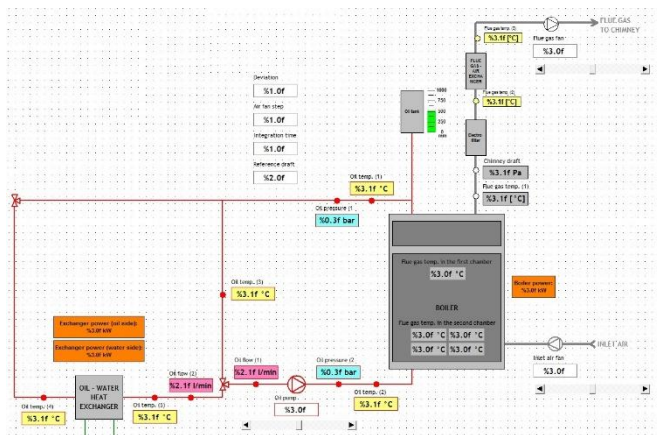


Fig. 3. Visualization of the experimental stand with measurement points [15].

The test stand was equipped with automatic data acquisition from temperature probes and an industrial exhaust gas analyzer (fig. 3). For a detailed analysis of combustion processes in batch boilers, four key operational parameters were identified as the basis for validating the simulation model. The focus was on examining the dynamics of changes in carbon monoxide (CO) and carbon dioxide (CO₂) concentrations, as well as oxygen (O₂) in the flue gases, and monitoring the temperature in the secondary combustion chamber. For components with a high response time (such as an air damper), tripping the system before the information about the need for a shortage is essential to respond in time.

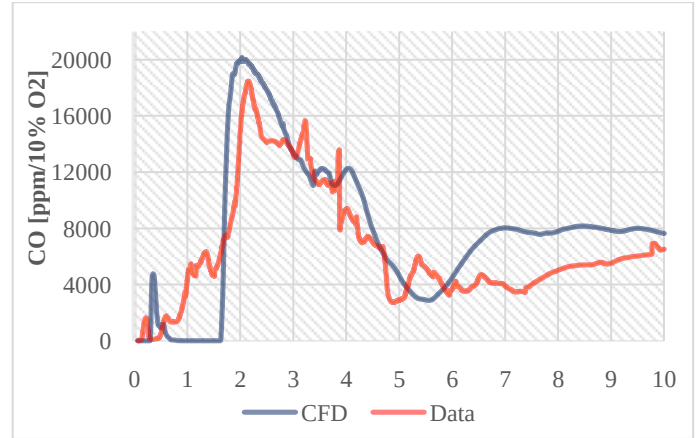


Fig. 4. Comparison of experimental data and CFD simulations, dynamics of concentration changes for selected CO parameter. In red the average measurement, in blue the CFD results.

The entire startup period of the installation – igniting the boiler – lasting 10 minutes can be represented in the form of intermediate states, where the detection of their start and end can serve as a faster signal than direct parameter measurement. The decision scheme for such an algorithm may look as follows:

Initialization: Preliminary Heating: The algorithm starts from an initial state, preparing the system for ignition by monitoring initial parameters such as temperature and chemical composition in the combustion chamber.

- Stage 1: Ignition (up to 60 seconds)

Ignition Detection: The algorithm detects an increase in CO emissions and other chemical products, as well as a temperature rise, indicating local ignition.

Parameter Adjustment: In response to detected ignition, the algorithm adjusts the fan and damper settings to ensure an optimal amount of air for the combustion process.

- Stage 2: Flame Spread (60 – 240 seconds)

Monitoring and Adaptation: During the intensive flame spread phase, the algorithm monitors and analyzes changes in emissions and temperature. It adjusts the fan and damper settings to avoid a rich mixture or suboptimal chemical reaction conditions.

- Stage 3: Stabilization (240 – 360 seconds)

Process Stabilization: When the algorithm detects a decrease in flame spread dynamics, it transitions to maintaining optimal parameters for balanced combustion – i.e., nominal operation.

- Stage 4: Nominal Combustion (360 – 600 seconds)

Maintaining Nominal Combustion: The algorithm maintains stabilized system parameters while being ready to detect and respond to potential anomalies resulting from fuel load fluctuations.

Intervention - Anomaly Response: In the event of anomaly detection, the algorithm immediately responds by adjusting the system parameters to prevent inefficient combustion or potential hazards.

Process Conclusion - Combustion Finalization: After completing the nominal combustion process, the algorithm switches

to a monitoring state, checking whether the residual fuel smoldering process is proceeding correctly.

Continuous Learning: Data Analysis and Optimization: At each stage, the collected data is analyzed, and the algorithm learns and optimizes its actions for the future, increasing efficiency and reducing emissions.

It is also essential to observe that for such complex processes, there is partial overlap of stages, illustrating the process's complexity and the need to provide multiple runs based on models and experimental work to train the state transition detection algorithm.

Analysis of the results of the experimental work makes it possible to observe certain relationships that can serve as flags for subsequent stages of combustion. It is possible to identify common trends and characteristic points that can be used to optimize a control algorithm with a predictive component (Fig. 4). Data from the initial phase of combustion is crucial for quick and efficient process startup. Subsequent phases of combustion require the algorithm to adjust and maintain optimal conditions for full and even combustion. As in the case above, it is possible to analyze a whole range of data and their time courses, such as circulating medium temperature, losses or boiler efficiency. Comprehensive data analysis indicates the need for a flexible algorithm that adapts to changing conditions in real time and predicts next steps to respond more quickly to changing demands, such as air.

4. Results and discussion

The model was trained using a dataset consisting of experimental results and numerical simulations of boiler ignition. Each measurement was labeled by an experienced user and constituted the training dataset. Supervised learning was utilized for training, with cross-entropy as the loss function and the Adam optimization algorithm used to update the network weights based on the calculated loss. To prevent overfitting, cross-validation was employed, where the data was divided into training and validation sets. After each iteration, the model's performance was assessed on the validation set, and the best model was saved and tested on a separate test set. A classic split was applied, with 80% of the input data used for training, 10% for validation, and 10% for the test set.

After the initial training, hyperparameter optimization was conducted, including the number of layers, the size of convolutional filters, and the number of neurons in fully connected layers, to achieve the best possible performance. The model was validated using a previously unseen dataset to evaluate its ability to generalize to new data.

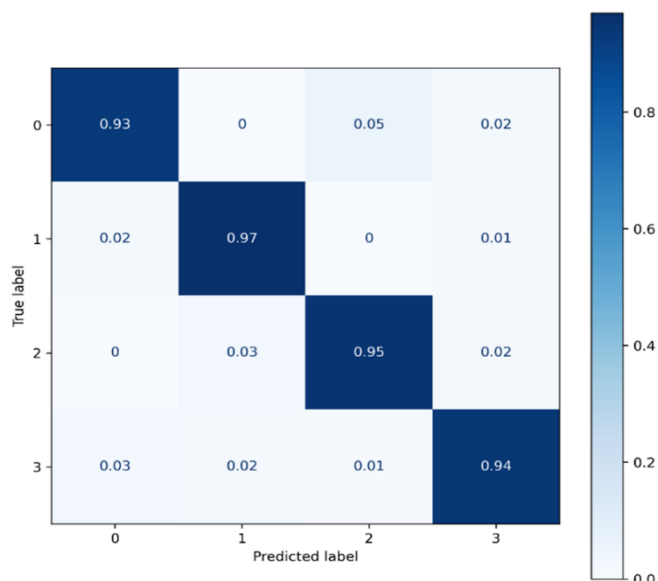


Fig. 5. Confusion matrix for convolution network for classification of combustion stages.

Based on the obtained confusion matrix (Fig. 5) and according to the formulas (r1) to (r5) presented in section 4.3, the following coefficients for the entire system were calculated: Recall: 95.8%, Accuracy: 95.8%, F1-Score: 95.79%, Precision: 95.80%. Additionally, the F1-Score for individual combustion stages was calculated as follows:

- Stage 0 (Ignition): 0.93
- Stage 1 (Flame Spread): 0.97
- Stage 2 (Process Stabilization): 0.95
- Stage 3 (Nominal Combustion): 0.94

Based on the confusion matrix (Fig. 5), it can be concluded that the neural network performs well in classifying the combustion stages. For each combustion stage, the majority of the network's predictions match the actual labels, indicating a high ability to recognize the characteristic features of each stage. For stages "0" and "3", small FP values are observed, where the network incorrectly classified the stage as "3" or "0". Similarly, for stage "2", a small number of FN values were obtained, where the network failed to recognize stage "2" and incorrectly classified it as "1". These low values indicate good specificity and sensitivity of the model.

5. Conclusion

The high F1-Score values for each stage (close to 1) indicate a good balance between precision and recall, meaning that the model performs well in identifying positive cases while avoiding false positives. The high values along the diagonal of the matrix indicate that the network has a high capability for correctly classifying each combustion stage. Based on the conducted training, it can be concluded that the model is capable of predicting the ignition stages based on measurement data, which allows for the automation of the monitoring process and potentially the optimization of the entire process. The designed algorithm can be implemented in a biomass boiler controller, enabling automatic regulation of the combustion process based on continuous analysis of measurement data and identification of combustion stages.

The research confirmed the significant utility of numerical models and their simulation results, which serve as the basis for analyzing the combustion process using neural networks. The application of convolutional neural networks allowed for the determination of combustion phases, even in such a complex process as straw ignition. Identifying the phase of the process based on data and trends of selected parameters enables faster responses in the installation without operator intervention. As a result of the research, the effectiveness of detecting process stage changes by the convolutional network was achieved, expressed using the confusion matrix through the F1-score (the harmonic mean between precision and recall) at a level close to 96%.

6. References

- [1] Liu, Hao, et al. "Control of NOx emissions of a domestic/small-scale biomass pellet boiler by air staging." *Fuel* 103 (2013): 792-798.
- [2] Böhler, Lukas, et al. "Carbon monoxide emission models for small-scale biomass combustion of wooden pellets." *Applied energy* 254 (2019): 113668.
- [3] Szubel, M., A. Dernbecher, and T. Dziok. "Determination of kinetic parameters of pyrolysis of wheat straw using thermogravimetry and mathematical models." *IOP Conference Series: Earth and Environmental Science*. Vol. 214. No. 1. IOP Publishing, 2019.
- [4] A. Le Guennec, S. Malinowski, and R. Tavenard, "Data augmentation for time series classification using convolutional neural networks," in *ECML/PKDD workshop on advanced analytics and learning on temporal data*, 2016.

- [5] Chen, Yaoran, et al. "2-D regional short-term wind speed forecast based on CNN-LSTM deep learning model." *Energy Conversion and Management* 244 (2021): 114451.
- [6] Lim, Su-Chang, et al. "Solar power forecasting using CNN-LSTM hybrid model." *Energies* 15.21 (2022): 8233.
- [7] Li, Nan, and Yong Hu. "The deep convolutional neural network for NOx emission prediction of a coal-fired boiler." *Ieee Access* 8 (2020): 85912-85922.
- [8] Li, Zhenhua, Guanghong Li, and Bin Shi. "Prediction of Oxygen Content in Boiler Flue Gas Based on a Convolutional Neural Network." *Processes* 11.4 (2023): 990.
- [9] P. Bałazy, P. Gut, and P. Knap, "Neural classifying system for predictive maintenance of rotating devices.," in *IOP Conference Series: Materials Science and Engineering*, 2022, p. 12013.
- [10] P. Knap, K. Lalik, and P. Bałazy, "Boosted Convolutional Neural Network Algorithm for the Classification of the Bearing Fault form 1-D Raw Sensor Data," *Sensors*, vol. 23, no. 9, p. 4295, 2023.
- [11] Knap, Paweł, Krzysztof Lalik, and Patryk Bałazy. "Boosted Convolutional Neural Network Algorithm for the Classification of the Bearing Fault Form 1-D Raw Sensor Data." *Sensors* 23.9 (2023): 4295.
- [12] K. Lalik, J. Kozak, S. Podlasek, and M. Kozek, "Self-Powered Wireless Sensor Matrix for Air Pollution Detection with a Neural Predictor," *Energies (Basel)*, vol. 15, no. 6, p. 1962, 2022.
- [13] Macukow, Bohdan. "Sieci neuronowe, historia badań i podstawowe modele." (2020).
- [14] M. Szubel, K. Papis-Fraczek, and S. Podlasek, "Impact of the air supply system configuration on the straw combustion in small scale batch-boiler-experimental and numerical studies," *Renew Energy*, vol. 220, p. 119540, 2024.
- [15] K. Sornek, "Prototypical biomass-fired micro-cogeneration systems—energy and ecological analysis," *Energies (Basel)*, vol. 13, no. 15, p. 3909, 2020.

Review of feature selection methods for Predictive Maintenance Systems

Urszula Jachymczyk, Paweł Knap
AGH University of Krakow, Poland
ulajach1@gmail.com, pknapp@agh.edu.pl

Abstract: The development of Industry 4.0 and Predictive Maintenance Systems allows for effective utilization of equipment by incorporating ML methods for identifying tool condition. However, including large number of Condition Indicators for machinery monitoring increases computational complexity, hence the response of the system elongates. Therefore, it is important to check the utility of indicators and reduce them. In this paper, we investigate different feature selection methods: AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion), Random Forest, Lasso Regression (L1 Regularization) for NASA Gearbox Fault Detection Dataset, PHM 2009. We processed the raw data and calculated CI from time domain, frequency domain and envelope. An SVM Classifier model was trained on full collection of indicator and reduced, then performances were compared. The obtained results highlight the advantage of feature selection, proving that effective PdM systems can be based on diminished number of Health Indicators.

Keywords: FAULT DETECTION, NASA GEARBOX DATASET, FEATURE SELECTION, AIC, L1 REGULARIZATION, RANDOM FOREST

1. Introduction

In industry, monitoring of machinery condition is usually based on 100+ Condition Indicators (CI, also called Health Indicators, HI) calculated from sensors signals. They are usually based on vibration signals which carry extensive information about the state of machine, identifying degradation patterns that indicate potential failures. CI are usually calculated directly from time domain, as well as from frequency domain or envelope/envelope spectrum. For PdM systems working in real time, any additional computations slow down the detection of possible failure. Additionally, numerous CI often tend to carry similar information about the state of a machinery and a large number of irrelevant features can adversely affect the performance of system, by incorporating unnecessary details and noise in classification process. Therefore, feature selection or feature reduction methods should be considered.

Feature selection is a search problem in a space of feature subsets. 4 families of feature selection methods can be distinguished:

- *Filter methods* rely solely on properties of the data to select features, without involving any learning algorithm [1]. They are fast and computationally intensive (e.g. Correlation Coefficient, Chi-Square Test).
- *Wrapper methods* evaluate feature subsets by training and testing a specific learning algorithm [1]. They use the model's performance as the evaluation criterion, making them more accurate but computationally expensive (e.g. AIC, BIC which uses Forward Selection or Backward Elimination).
- *Embedded methods* perform feature selection during the model training process. They integrate the selection with the learning algorithm, benefiting from both the filter and wrapper approaches [2] (e.g. Lasso Regularization, Ridge Regularization, Elastic Net, Three-based method).
- *Hybrid methods* combine aspects of filter, wrapper, and embedded methods to leverage their respective strengths and mitigate their weaknesses. These methods involve a multi-stage process where filter methods are used for an initial broad selection of features, followed by wrapper or embedded methods [2] (e.g. Random Forest, XGBoost, LightGBM).

In this paper, AIC, BIC wrapper methods, Lasso Regression embedded method, Random Forest hybrid method will be tested on classification of bearings faults in mechanical gearbox. Rolling element, such as bearings, are the core and vulnerable components in the machinery whose health condition directly affects in the performance, efficiency and lifespan of the machines [3].

2. Dataset

A methodology for bearing health assessment and fault classification is developed and evaluated for gearbox vibration data

provided by the Prognostics and Health Management Society for the 2009 data challenge competition [4]. However, data originally was unlabeled, therefore a labeled dataset (very similar to this used in the competition, available in [5]) provided in November 2009 was used to perform supervised learning. The data is representative of generic industrial gearbox (fig. 1). Two geometries were used for measurements: one using a spur gears, the other using spiral cut (helical) gears.

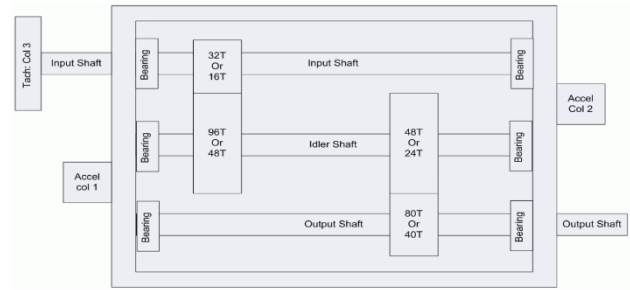


Fig. 1 Schematic of Gearbox Used in PHM 2009 Challenge Data [4].

The measurements consisted of two accelerometers signals along with a tachometer signal with sample rate 200/3 [kHz]. The gearbox contains several mechanical elements, including three shafts, 4 gears, and 6 bearings. The data provided consisted of 160 data sets, in which the gearbox was tested under 5 different speeds (30 [Hz] and 1800 [RPM], 35 [Hz] and 2100 [RPM], 40 [Hz] and 2400 [RPM], 45 [Hz] and 2700 [RPM], 50 [Hz] and 3000 [RPM]), 2 different loads (high and low). In this paper, detection of bearing faults was prioritized. On tab.1 presented are considered bearings and their faults (IS - Input Shaft, ID - Idler Shaft, OS - Output Shaft, :IS - Input Side, :OS - Output Side).

Table 1: Gearbox bearings and their defects faults for 8 operating conditions.

Bearing	IS:IS	ID:IS	OS:IS	IS:OS	ID:OS	OS:OS
Case						
Spur 1	Good	Good	Good	Good	Good	Good
Spur 2	Good	Good	Good	Good	Good	Good
Spur 3	Good	Good	Good	Good	Good	Good
Spur 4	Ball	Good	Good	Good	Good	Good
Spur 5	Inner	Ball	Outer	Good	Good	Good
Spur 6	Inner	Ball	Outer	Good	Good	Good
Spur 7	Inner	Good	Good	Good	Good	Good
Spur 8	Good	Ball	Outer	Good	Good	Good

3. Signal processing and feature extraction

The health assessment and fault classification algorithms requires valid input features to accurately evaluate the condition of gearbox components and level of damage. Great significance has extraction and selection of the most appropriate features that correlate with bearings fault signatures and fault severity [6]. The methods for signal processing and feature extraction used to derive various condition indicators from vibration signals are discussed below. Most commonly employed scalar state indicators, calculated using acceleration signals are described in [7].

Time domain indicators

Features derived from the raw time signal can offer a comprehensive understanding of the vibration levels exhibited by the monitored gearbox, as well as the distribution of the vibration data. The PdM systems typically utilize wideband diagnostic indicators which are primarily computationally efficient, resulting in quick responses, and also less sensitive to disturbances and signal variations. Utilized time domain indicators for bearings faults detection:

- Peak-to-peak (PP),
- Root mean square (RMS),
- Crest factor (CF),
- Shape factor (SF),
- Standard deviation (STD),
- Kurtosis,
- Velocity root mean square (VRMS).

Frequency domain indicators

The advantage of using simple scalar time-domain indicators is that no knowledge about the characteristic fault frequencies is required and thus the number and complexity of the components is not taken into account. This simplifies the analysis procedure significantly, with the trade-off being that it does not provide any information about the exact location of component exhibiting the anomalous behavior, therefore more component-specific frequency-domain indicators are used [8]. Frequency analysis can detect faults at an early stage, even before they become apparent in the time domain. The following general frequency condition indicators were calculated:

- Mean frequency (MF),
- Frequency Cener (FC),
- 3 max amplitude peaks and corresponding frequencies.

Envelope indicators

We also calculated more component-specific frequency-domain from envelope spectrum. The envelope spectrum is probably the most used tool to inspect the cyclostationary behavior of signals. Most mechanical faults of bearings or gears induce some form of cyclostationary behavior in the observed vibration signals [8]. This phenomenon causes the envelope spectrum to exhibit discrete peaks at the corresponding fault frequencies which arise from the background noise and thus their amplitude is distinguishable from the spectral noise. A faulty rolling bearing usually emit the following frequencies [9]:

- FTF (Fundamental Train Frequency) – frequency used for overall bearing condition assessment,

$$FTF = \frac{f}{2} \cdot \left(1 - \frac{d \cdot \cos \theta}{D_p}\right) \quad (1)$$

- BPFO (Ball Pass Frequency of the Outer Race) - frequency emitted by a rolling element bearing when it impacts a defect in the outer race,

$$BPFO = \frac{n \cdot f}{2} \cdot \left(1 - \frac{d \cdot \cos \theta}{D_p}\right) \quad (2)$$

- BPFI (Ball Pass Frequency of the Inner Race) – frequency emitted by a rolling element bearing when it impacts a defect in the inner race,

$$BPFI = \frac{n \cdot f}{2} \cdot \left(1 + \frac{d \cdot \cos \theta}{D_p}\right) \quad (3)$$

- BSF (Ball Spin Frequency) - frequency at which bearing balls rotate about their own axes, indicate defects on inner or outer race.

$$BSF = \frac{D_p \cdot f}{2d} \cdot \left(1 - \left(\frac{d \cdot \cos \theta}{D_p}\right)^2\right) \quad (4)$$

D_p – pitch diameter of bearing,
 n – number of balls,
 f – rotating frequency,
 d – rotating elements diameter,
 θ – contact angle.

4. Feature selection methods

The purpose of feature selection is to identify and choose the most relevant and informative subset of features from a larger set of available features or variables. The majority of efforts in fault prediction an classification concentrate on identifying the most effective models neglecting cleaning the training data and selection of best parameter values from it. Occasionally, the data used for model training have considerable variability, which can lead to a reduction in model accuracy. The effectiveness of Predictive Maintenance systems depends directly from features chosen to identify the condition of a machinery resulting in it successfulness or, in contrary, its failure.

AIC and BIC

AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) are both statistical measures used in the field of model selection and evaluation [10]. They provide a quantitative way to balance the trade-off between model fit and complexity, penalizing overly complex models to prevent overfitting. AIC and BIC can also be used for feature selection by evaluating different subsets of features [11]. The model can be trained with different combinations of features and the subset of features that results in the lowest AIC or BIC value is selected.

$$AIC = -2 \cdot \ln(L) + 2 \cdot k \quad (5)$$

$$BIC = -2 \cdot \ln(L) + \ln(n) \cdot k \quad (6)$$

L - maximum likelihood of the model,
 k - number of parameters in the model,
 n - sample size.

In both criteria, the first term ($-2 \cdot \ln(L)$) penalizes models based on the goodness of fit (likelihood), with lower values indicating a better fit. The second term penalizes models complexity.

For the criteria to be the most accurate, it is advised to iterate over all possible combinations of features (exhaustive search), but this approach can be computationally expensive for datasets with a large feature space [12]. To make the process less computationally expensive, a forward or backward selection method can be used. Forward selection starts with no features and iteratively adds the features that reduces AIC and BIC until no further improvement is observed. Backward selection starts with all features and iteratively removes the features. Those approaches quickly converges to solution but considers only local optima.

L1 Regularization

The L1 regularization (also called Lasso regularization), can be used in Logistic Regression model. Additional penalty is added to a cost function which forces the learning algorithm to not only fit the data, but also keep the model weights as small as possible, preventing overfitting [13]. After training the Lasso logistic regression model, the coefficients of the features are examined. Features with non-zero coefficients are considered important and selected for the final model, while features with coefficients close to zero are considered less important and may be excluded. The strength of the regularization penalty (λ) can be adjusted.

Random Forest

Using Random forest is a hybrid feature selection method which works by leveraging the importance scores assigned to each feature during the training process; it considers interactions between features and automatically handles feature importance calculation without the need for explicit assumptions about the data distribution. Scikit-Learn measures a feature's importance by looking at how much the tree nodes that use that feature reduce impurity on average (across all trees in the forest) [13].

5. Methodology

The measurement data with accelerometers signals was loaded and split into data frames with 50% overlap to obtain the larger amount of training points.

For classification of faults of 6 bearings in gearbox a multilabel classification was necessary – each training instance can be assigned to multiple classes (e.g. inner race failure for IS:IS and ball failure for ID:IS tab.1). However, the bearings IS:OS, ID:OS, OS:OS did not encounter any breakage, so they were excluded from fault detection - there were 3 bearing remaining. The labels were created according to tab.2.

Labels 4, 7, 8, 10, 12 never occurred so the multilabel classification was conducted using 7 occurring labels.

Table 2: Bearing faults and corresponding labels (G -Good, B – Ball fault, I – Inner race fault, O – Outer race fault).

	Bearing 1				Bearing 2				Bearing 3			
Fault	G	B	I	O	G	B	I	O	G	B	I	O
Label	1	2	3	4	5	6	7	8	9	10	11	12

For each data frame BPFO, BPFI, FTF, BSW were calculated, which depended on bearing operating condition for measurements with different speed (tab. 3). The details of bearings geometry required for calculation of those component-specific indicators was provided by the authors of a dataset.

Table 3: Calculated characteristic bearings frequencies for various operating frequencies.

	30 Hz	35 Hz	40 Hz	45 Hz	50 Hz
BPFO [Hz]	91,57	106,83	122,09	137,35	152,62
BPFI [Hz]	148,43	173,17	197,91	222,65	247,38
BSF [Hz]	5,84	6,81	7,78	8,75	9,73
FTF [Hz]	11,45	13,35	15,26	17,17	19,08

Time domain data was transformed to frequency domain and envelope spectrum. As presented on fig. 2, characteristic frequencies were much more distinguishable for envelope spectrum than from raw data spectrum, proving its usefulness for fault detection. Plot was created for spur 5 using input accelerator signal, which should detect the inner race fault for bearing 1.

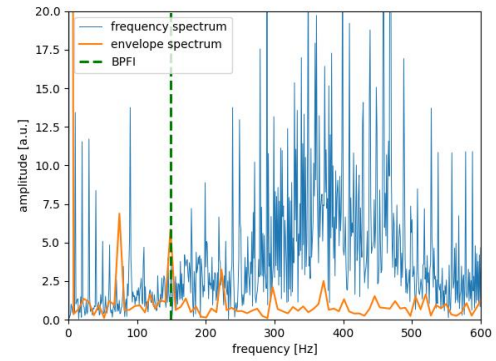


Fig. 2 Comparison of frequency and envelope spectrum for identification of characteristic bearings frequencies.

On prepared data, all aforementioned in section 3 Condition Indicators were calculated. As for BPFO, BPFI, FTF, BSP – when peak frequencies were present in envelope spectrum a “1” value was assigned, if not – “0”.

SVM Classifier was initially tested on full collection of indicators, then the feature selection was conducted using 4 methods. For AIC and BIC a forward selection approach was used because exhaustive search was computationally expensive for calculating combinations of 38 features. Feature selection using Random forest model started with finding satisfactory model parameters and then the feature importances were directly acquired. For L1 regularization, Logistic Regression model was trained with penalty parameter equal to L1. Coefficient results returned by model allowed to assess the usefulness of individual indicators. For RF and LR we decided to halve the number of indicators to 20 (the initial number was 38) by establishing suitable thresholds. SVM model was tested on 4 reduced indicators set using 10-fold cross validation (to average results across 10 different subsets of a dataset and enable utilization of all points for training), after adjusting its parameters (using grid search). As an evaluation a weighted F1-score was used, which calculates F1 score for each label as the harmonic mean of precision and recall. Each label has assigned a weight based on the number of true instances for that label in the dataset - it is particularly useful in multi-label classification tasks where the labels are imbalanced, as it provides a more comprehensive evaluation of the model's performance across all labels.

6. Results

After applying aforementioned feature selection methods, dataset dimension significantly decreased. However, each one of the algorithms chose different subset of features (in – signal from accelerometer mounted on input side, out - accelerometer mounted on output side). Best features according to:

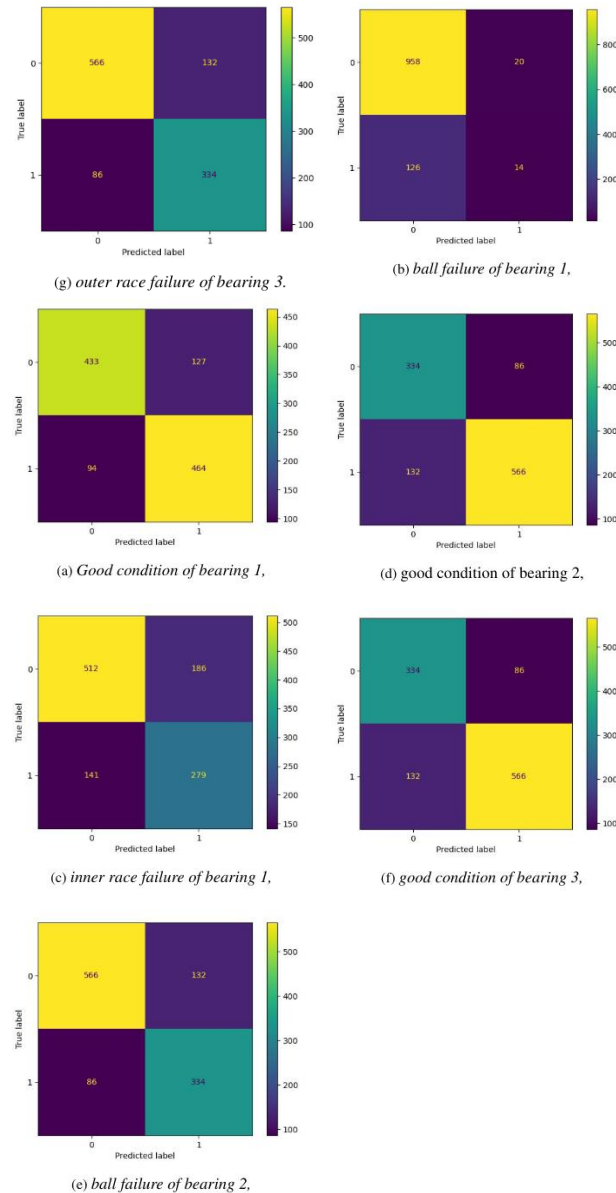
- AIC: pp_in, sf_in, std_in, kurt_in, fc_in, rms_in, BPFO_out, freq3_out;
- BIC: pp_in, sf_in, std_in, fc_in, std_in, freq3_out;
- RF: pp_in, rms_in, cf_in, sf_in, mf_in, fc_in, sf_out, std_out, kurt_out, vrms_out, mf_out, fc_out, peak1_in, freq1_in, peak2_in, freq3_in, peak1_out, peak2_out, peak3_out, freq3_out;
- L1: pp_in, rms_in, sf_in, std_in, vrms_in, mf_in, fc_in, pp_out, rms_out, cf_out, sf_out, std_out, kurt_out, vrms_out, mf_out, fc_out, peak1_in, peak3_in, peak2_out, peak3_out.

For every chosen subset of features tweaking parameters of SVM model was required to adapt model to new feature space (tab.4).

Table 4: Comparison of SVM model parameters and averaged classification results for different feature selection methods..

Feature selection method	SVM model parameters		Averaged F1-score results
	C	γ	
No feature selection	275	0,015	0,792
AIC	1200	0,1	0,683
BIC	50	0,2	0,653
Random Forest	1500	0,01	0,807
L1 regularization	600	0,01	0,821

Also confusion matrices of all 7 possible labels for best performing SVM model and dataset were generated (fig. 3).

**Fig. 3** Final confusion matrices for each label

7. Conclusions

Results of this paper prove that selection and final choice of features plays crucial role in performance of PdM systems. Through thorough data analysis and testing various indicators reduction methods we were able to improve the performance of SVM model by 3% while using only half of available CI. The best results were achieved by L1 Regularization algorithm and the worst by BIC. Both AIC and BIC are very sensitive and rejected too many indicators. Additionally, when dealing with a bigger number of features, method is computationally intractable and not possible to use in practice. Surprisingly, even though Random Forest is most

state-of-art among presented algorithms, it did not obtain best results.

Using each of the proposed methods resulted in different subsets of features – it leads to a conclusion, that one perfect set of indicators that would be selected by all algorithms do not exist. It depends from selection criteria the methods use, which should be adjusted to requirements of individual PdM system.

Only one from presented component-specific CI from envelope spectrum was selected by one of feature selection algorithms. It can be a result of noisiness of data caused by failures of other gearbox components.

An SVM Classifier model achieved satisfactory results for every subset of dataset therefore is sufficient for bearing faults classification task. It has been observed that the model parameters vary significantly depending on set of indicators serving as an input. Model has the biggest problem with differentiability of ball failure in bearing 1, which would require further adjustments and/or additional indicators.

1. Talavera, L. (2005). An Evaluation of Filter and Wrapper Methods for Feature Selection in Categorical Clustering. In: Famili, A.F., Kok, J.N., Peña, J.M., Siebes, A., Feelders, A. (eds) *Advances in Intelligent Data Analysis VI*. IDA 2005. Lecture Notes in Computer Science, vol 3646. Springer, Berlin, Heidelberg.
2. A. Jović, K. Brkić and N. Bogunović, "A review of feature selection methods with applications," *2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, Opatija, Croatia, 2015, pp. 1200-1205
3. D. Neupane and J. Seok, "Bearing Fault Detection and Diagnosis Using Case Western Reserve University Dataset With Deep Learning Approaches: A Review," in *IEEE Access*, vol. 8, pp. 93155-93178, 2020
4. PHM 2009 Data challenge Competition, (2009). [<http://www.phmsociety.org/competition/09>]
5. PHM 2009 Data challenge Competition - labeled version form November. [https://github.com/mingzhangPHD/Supplementary-Dataset/tree/main/PHM2009%20Gearbox/PHM_Society_2009_Competition_Expanded_txt]
6. Al-Atat, Hassan & Siegel, David & Lee, Jay. (2011). A Systematic Methodology for Gearbox Health Assessment and Fault Classification. *International Journal of Prognostics and Health Management*. 2. 10.36001/ijphm.2011.v2i1.1339.
7. Vikas Sharma and Anand Parey. „A review of gear fault diagnosis using various condition indicators”. W: Elsevier 144 (2016), s. 253–263.
8. Cédric Peeters, Jérôme Antoni, Jan Helsen, “Blind filters based on envelope spectrum sparsity indicators for bearing and gear vibration-based condition monitoring”, *Mechanical Systems and Signal Processing*, Volume 138, 2020, 106556
9. Roque, Antonio & Silva, Tiago & Calado, João & Dias, J.. (2009). An approach to fault diagnosis of rolling bearings. *WSEAS Transactions on Systems and Control*. 4.
10. Rossi R, Murari A, Gaudio P, Gelfusa M. Upgrading Model Selection Criteria with Goodness of Fit Tests for Practical Applications. *Entropy*. 2020; 22(4):447
11. Albahli S, Yar GNAH. Defect prediction using akaike and bayesian information criterion. *Comput Syst Sci Eng*. 2022;41(3):1117-1127
12. Khalili, Abbas. (2011). An Overview of the New Feature Selection Methods in Finite Mixture of Regression Models. *Journal of the Iranian Statistical Society JIRSS*. 2.
13. Aurélien Géron. „Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow, Concepts, Tools, and Techniques to Build Intelligent Systems (second edition)”. Sebastopol, CA: O'Reilly Media, Inc., 2019.

Deep Eye - Intelligent Device for Creating Digital Models of Underwater Topography

Zuzanna Schab, Marta Żelasko, Krzysztof Lalik, Krzysztof Mamcarz
AGH university of Krakow, Poland
klalik@agh.edu.pl

Abstract: This project endeavors to advance underwater mapping technologies through the development of algorithms tailored for mapping, analyzing, and creating digital models of bathymetry. A key aspect of this research involves the integration of advanced technologies, specifically High-Definition cameras and a purpose-designed sonar system. By leveraging these tools, high-resolution point cloud data and precise representations of submerged geographical structures can be obtained. The project also aims to conduct a waterproof device, operated by a diver, incorporating both HD cameras and sonar sensors. This device will enhance data collection and mapping capabilities in challenging underwater environments, addressing complexities associated with underwater terrains. Several critical questions are central to this research. Primarily, the investigation seeks to understand how artificial intelligence algorithms can be optimized to harness the collective potential of HD cameras and the specially crafted sonar system. Additionally, the research endeavors to evaluate the comparative efficacy and accuracy of the proposed sonar system and HD cameras in underwater mapping, particularly under adverse environmental conditions where visibility is compromised by high particle concentrations. Through these efforts, this project aims to contribute to the advancement of underwater mapping technologies, offering insights into optimizing sensor integration and algorithm development for enhanced mapping accuracy in challenging underwater environments.

Keywords: UNDERWATER IMAGING, UNDERWATER 3D POINT CLOUD, UNDERWATER PHOTOGRAMMETRY, FEATURE MATCHING

1. Introduction

In the realm of marine exploration and resource management, the accurate depiction of underwater terrain holds paramount importance. Submerged landscapes harbor a wealth of ecological diversity, geological features, and invaluable resources, yet their detailed mapping has long presented a formidable challenge to scientists and researchers. Traditional sonar systems, which utilize sound waves to map underwater topography, have been indispensable tools in oceanographic research for decades. However, these traditional methodologies, while providing invaluable insights into submerged terrain, often suffer from limitations such as resolution constraints and an inability to capture fine-scale features. The integration of sonar systems with High-Definition (HD) cameras has emerged as a promising approach to overcome these limitations and revolutionize our ability to unveil the mysteries of the deep.

Recent advances in the field have seen the development and application of sophisticated models and technologies aimed at enhancing underwater mapping capabilities. Authors in [1] presented a theoretical model explaining the imaging mechanism of underwater topography by radar, highlighting the significance of current-induced surface effects. This early work laid the foundation for understanding how underwater features could be detected using surface observations. Building on such foundational theories, a novel framework incorporating deep learning techniques, specifically the Gate Recurrent Unit (GRU) model, to improve water depth inversion in complex marine environments were introduced in [2]. Their study demonstrated significant improvements in accuracy and applicability of inversion models, setting a new standard for underwater topographic mapping.

The integration of mathematical modeling and computational analysis further enriches the field. Authors in [3] utilized GPS and echo sounder data to construct underwater topography models, providing insights into sedimentation changes over time. Such approaches underscore the importance of combining multiple data sources to achieve more accurate and dynamic representations of underwater landscapes.

Enhancements in image processing technologies have also played a critical role in advancing underwater mapping. In [4] an efficient underwater image enhancement method that significantly improved visibility and detail in underwater images were proposed. Similarly, [5] developed a Bayesian retinex algorithm for enhancing underwater images, addressing common issues such as color casts and low visibility. These advancements in image enhancement

techniques are crucial for the effective integration of HD cameras in underwater mapping systems.

Moreover, a multi-view underwater image enhancement method using an embedded fusion mechanism, which further refined image quality and detail retrieval, were introduced in [6]. Their approach demonstrated the potential of combining multiple enhancement algorithms to achieve superior results, particularly in challenging underwater conditions.

The capture and processing of 3D point cloud data have also seen significant advancements. Authors in [7] proposed a methodology for capturing point cloud data in turbid underwater environments, addressing issues such as turbidity and distortion. This method proved effective in obtaining accurate 3D representations of underwater structures, essential for detailed topographic mapping.

Deep learning techniques have been increasingly employed to enhance underwater mapping capabilities. Authors in [8] utilized deep learning methods for multi-dimensional underwater point cloud detection, achieving high accuracy in detecting and classifying underwater objects. Similarly, convolutional neural networks (CNN) with point cloud registration to improve the precision of underwater object pose estimation were combined in [9]. These studies highlight the potential of AI-driven approaches in refining underwater mapping and object detection processes.

Lastly, Authors in [10] explored the use of visual SLAM (Simultaneous Localization and Mapping) algorithms for autonomous underwater robot navigation. The integration of ORBSLAM2 and point cloud processing demonstrated the feasibility of using minimal instrumentation setups for effective underwater navigation and mapping.

This project endeavors to advance underwater mapping technologies through the development of algorithms tailored for mapping, analyzing, and creating digital models of bathymetry. By integrating advanced technologies, specifically High-Definition cameras and a purpose-designed sonar system, this research aims to obtain high-resolution point cloud data and precise representations of submerged geographical structures. The development of a waterproof device operated by a diver, incorporating both HD cameras and sonar sensors, will enhance data collection and mapping capabilities in challenging underwater environments. The key questions central to this research involve optimizing artificial intelligence algorithms to harness the collective potential of HD cameras and sonar systems, and evaluating the comparative efficacy and accuracy of these tools under adverse environmental conditions. Through these efforts, this project aims to contribute to the

advancement of underwater mapping technologies, offering insights into optimizing sensor integration and algorithm development for enhanced mapping accuracy in challenging underwater environments.

2. Design Considerations

In addition to our primary design considerations, we have meticulously engineered the device to withstand the challenging conditions of underwater operation up to depths of 30 meters. Waterproofing is not merely a feature but a fundamental aspect of the device's design, ensuring its functionality and integrity even in the most demanding marine environments. By implementing robust seals and enclosures, we have fortified the device against water ingress, corrosion, and pressure, guaranteeing reliable performance throughout underwater missions.

Buoyancy has been a focal point in our design, as we recognize its pivotal role in facilitating ease of operation and mitigating the risk of device loss. By carefully calibrating the buoyant properties of the device, we have achieved a delicate balance that enables effortless manipulation by divers while preventing the device from sinking to unreachable depths. This buoyancy management system not only enhances user experience but also safeguards against potential accidents or equipment loss, ensuring uninterrupted data collection and mission success.

Furthermore, we have prioritized user comfort, ergonomics, and compactness in the design process, recognizing the importance of seamless integration into the diver's equipment ensemble. Its streamlined form factor and lightweight construction contribute to enhanced maneuverability and agility underwater, empowering divers to navigate complex environments with ease and confidence.

Given the inherent variability of the underwater landscape, we have meticulously optimized the shape of the device to a cylinder, to minimize hydrodynamic resistance and facilitate seamless navigation through confined spaces such as underwater caves and crevices.

3. Sonar System

For our sonar system we use a set of hydrophones and acoustic transducers. Similar to how a microphone captures sound in the air, a hydrophone functions to detect acoustic signals beneath the water's surface.

The majority of the hydrophones operate on a distinctive property of certain ceramics, generating a minute electrical current in response to fluctuations in underwater pressure. When submerged in the water, a ceramic-based hydrophone generates low-voltage signals across a board spectrum of frequencies, reacting to underwater sounds originating from various directions.

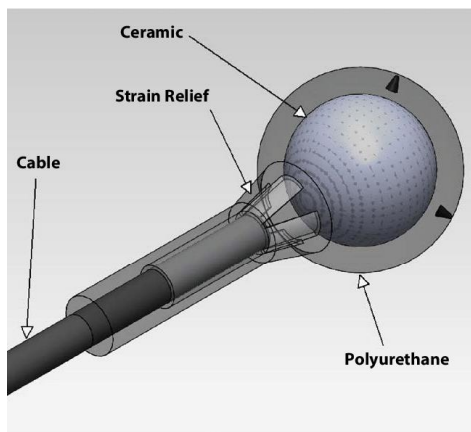


Fig. 1. Schematic of an example ceramic hydrophone

Through the amplification and recording of these electrical signals, hydrophones can accurately measure underwater sounds.

While a single hydrophone has the capacity to capture sounds from all directions, employing multiple hydrophones arranged in an array, yields signals that can be manipulated to enhance sensitivity beyond that of a singular device.

A point cloud comprises myriad data points suspended in space, collectively representing an object or a defined area such as a landscape or specific objects.



Fig. 2. Example of an underwater object

Each of these 3D points possesses a distinct set of X, Y, and Z geometric coordinates, essentially serving as a digital counterpart to corresponding points on the actual object.

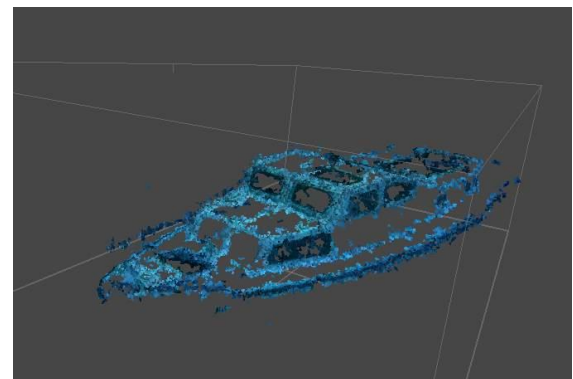


Fig. 3. Point cloud of underwater object

Point clouds are relatively straightforward to create and can be generated using freely available software, capable of storing highly detailed information with ease. Furthermore, their accuracy, particularly in the case of geo-referenced point clouds, is exceptional, enabling precise measurements. In most cases, constructing a point cloud for a landscape involves stitching together numerous high-resolution images captured by drones to produce a detailed output.

The level of detail in the resulting point cloud is directly influenced by the quantity of high-resolution images collected. On presented pictures we can see one of the pictures taken by one of our members. We then created a point cloud of the object which was later used to create a 3D model.

Feature matching is a critical step in many computer vision applications, including object recognition, image stitching, and 3D reconstruction. One widely used algorithm for this purpose is the FLANN (Fast Library for Approximate Nearest Neighbors) algorithm. Before matching features, it is essential to detect and describe these features accurately in each image. The Scale-Invariant Feature Transform (SIFT) algorithm is one of the most effective techniques for this task.

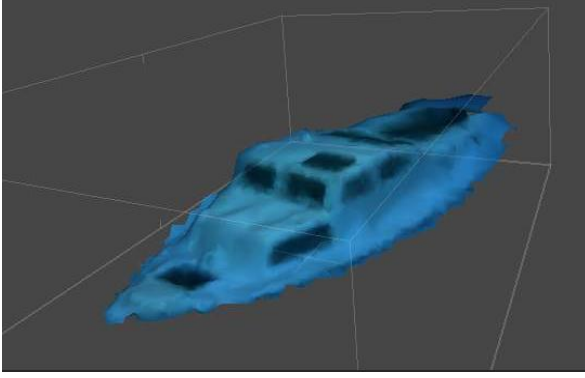


Fig. 4. 3D model created from the point cloud

SIFT is designed to detect and describe local features in images, which remain invariant to scale changes, rotation, and partially invariant to changes in illumination and 3D viewpoint. The first step in SIFT is constructing a scale-space representation of the image. This is done by convolving the image $I(x,y)$ with Gaussian kernels of varying standard deviations (σ), creating a series of progressively blurred images. This process can be described by Equation (1).

$$G(x,y,\sigma) = I(x,y)G(x,y,\sigma) \quad (1)$$

where $G(x,y,\sigma)$ is the Gaussian blur operator with standard deviation σ . This operation results in multiple blurred images at different scales, which help in detecting features that are invariant to scale changes.

The next step involves creating the Difference of Gaussian (DoG) images. DoG is computed by subtracting one blurred image from another, each having a different σ :

$$D(x,y,\sigma) = G(x,y,k\sigma) - G(x,y,\sigma) \quad (2)$$

where k is a constant multiplicative factor. This process results in a set of images that highlight the edges and important features in different scales.

SIFT detects keypoints by searching for local extrema in the DoG images. Each pixel is compared with its 8 neighbors in the current image and its 9 neighbors in the scale above and the scale below. If a pixel is a local extremum (either a maximum or minimum), it is considered a potential keypoint. Not all detected keypoints are equally useful. To refine the keypoints, SIFT uses the Taylor series expansion of the scale-space function to interpolate the location of the extrema accurately. Keypoints with low contrast (below a certain threshold) or those that are poorly localized along edges are discarded. Each keypoint is assigned one or more orientations based on the local image gradient directions. This step ensures that the keypoints are invariant to rotation. The orientation of a keypoint is determined by computing the gradient magnitude and direction around the keypoint and creating an orientation histogram. The highest peak in the histogram indicates the dominant orientation. Finally, a descriptor is created for each keypoint. This descriptor is a vector that describes the local image region around the keypoint. The image gradients are measured in the region around the keypoint, and these gradients are weighted by a Gaussian window. The resulting values are then transformed into a vector that is normalized to enhance invariance to changes in illumination.

Once features have been detected and described using SIFT, the next step is to match these features between images. FLANN (Fast Library for Approximate Nearest Neighbors) performs fast and approximate nearest neighbor searches. It is particularly useful when dealing with large datasets where exact nearest neighbor search would be computationally expensive.

The basic idea behind FLANN is to use a combination of randomized kd-trees and hierarchical k-means trees to organize the data, making the search process much faster.

To match features between two images, FLANN first builds an index of the feature descriptors from one image. For each feature descriptor in the second image, FLANN searches the index to find the closest matching descriptor. This process results in a set of matched feature pairs. To improve the robustness of the matching process, FLANN incorporates techniques to reject outliers. One common method is the ratio test proposed by [11], which compares the distance of the closest neighbor to that of the second-closest neighbor. Matches are accepted if the ratio is below a certain threshold, indicating a good match. By integrating these methods, it becomes possible to accurately match features between images, enabling various advanced computer vision tasks such as 3D modeling, image stitching, and object recognition in underwater environments.

6. Results

Our project aimed to advance underwater mapping technologies by integrating high-definition cameras and sonar systems to create accurate digital models of underwater terrains. The field tests conducted in both controlled and challenging environments provided valuable insights into the effectiveness and limitations of our approach.

Field Testing at Diving Pool

We conducted initial tests at the Diving pool, a controlled environment known for its excellent visibility. The primary objective was to generate a detailed point cloud of an actual boat wreck. The results from this test were highly successful, demonstrating the capability of our integrated system to produce high-resolution 3D models of underwater objects.

Table 1: Point Cloud Accuracy in Pool

Metric	Value
Average Point Density (points/m ²)	500000
Mean Distance Between Points (mm)	1,5
Standard Deviation of Distance (mm)	0,2
Total Number of Points	500000
Coverage Area (m ²)	10
Processing Time (minutes)	45

The results from Pool highlighted the system's ability to capture fine details and produce a dense and accurate point cloud, with an average point density of 500,000 points per square meter and a mean distance between points of 1.5 mm.

Field Testing at Zakrzówek

Subsequent tests were performed at Zakrzówek, an environment characterized by significantly poorer visibility due to high particle concentration. The objective was to assess the performance of our imaging and sonar systems under less ideal conditions.

Table 2: Point Cloud Accuracy in Zakrzówek

Metric	Value
Average Point Density (points/m ²)	120000
Mean Distance Between Points (mm)	5,0
Standard Deviation of Distance (mm)	0,8
Total Number of Points	120000
Coverage Area (m ²)	10
Processing Time (minutes)	60

The results from Zakrzówek indicated a significant reduction in point cloud density and accuracy, with an average point density of 120,000 points per square meter and a mean distance between points of 5.0 mm. This decrease in performance underscores the challenges posed by poor visibility and the limitations of photographic methods in such conditions.

The comparative analysis between the two test sites clearly illustrates the strengths and weaknesses of our system under different environmental conditions. In the controlled environment of Diving Pool, the high visibility allowed for precise and detailed point cloud generation. However, in Zakrzówek, the poor visibility conditions significantly impacted the quality of the point cloud, highlighting the need for complementary data from sonar systems to maintain accuracy.

The data clearly shows that while high-definition cameras can provide excellent results in clear water conditions, their effectiveness diminishes significantly in turbid environments. This finding underscores the importance of integrating sonar data to complement imaging techniques and ensure reliable and accurate underwater mapping across varying conditions.

The results of our tests underscore the critical role of integrating multiple technologies for underwater mapping. The high-definition cameras provided excellent results in clear water conditions, but their limitations in turbid environments highlighted the necessity for supplementary sonar data. Our system's ability to adapt to different environments by combining these technologies ensures robust and accurate mapping, paving the way for advancements in underwater exploration and resource management.

7. Summary

In this project, we successfully advanced underwater mapping technologies by integrating sophisticated sonar systems and high-definition cameras, coupled with advanced image processing and deep learning algorithms. Our efforts focused on developing robust algorithms for mapping, analyzing, and creating digital models of underwater terrains.

We conducted a series of tests at the Diving pool, where we successfully generated a detailed point cloud of an actual boat wreck. This demonstration confirmed the effectiveness of our integrated approach in a controlled environment with good visibility.

Further tests were carried out at Zakrzówek, an area with significantly poorer visibility. These tests revealed a critical limitation: imaging based solely on photographs proved insufficient under these challenging conditions. The reduced visibility highlighted the dependency of photographic methods on clear water conditions, underlining the necessity for supplementary techniques to ensure accurate underwater mapping.

Our findings indicate that the integration of sonar systems with HD cameras is essential for overcoming the limitations posed by adverse underwater environments. The sonar system, comprising hydrophones and acoustic transducers, provided reliable data even in turbid conditions, where traditional photographic methods failed. The use of SIFT for feature detection and FLANN for feature matching between images further enhanced our ability to process and integrate data from multiple sources.

The development of a waterproof device, designed to operate effectively up to depths of 30 meters, was a crucial aspect of our project. This device, incorporating robust seals, buoyancy management, and ergonomic considerations, facilitated efficient data collection and maneuverability in underwater environments. The cylindrical shape minimized hydrodynamic resistance, enabling smooth navigation through confined spaces.

Our research demonstrates that while high-resolution imaging plays a vital role in underwater mapping, it must be complemented by sonar data to address the challenges of visibility and detail capture. The integration of multiple data sources and the application of advanced algorithms ensure a more comprehensive and accurate representation of underwater terrains.

Future work will focus on further enhancing the integration of sonar and imaging technologies, improving the robustness of our algorithms, and expanding the application of our methods to more

diverse and challenging underwater environments. This project has laid a strong foundation for ongoing advancements in underwater mapping technologies, offering significant potential for future research and practical applications in marine exploration and resource management

6. References

- [1]. Alpers, W., & Hennings, I. (1984). A theory of the imaging mechanism of underwater bottom topography by real and synthetic aperture radar. *Journal of Geophysical Research: Oceans*, 89(C6), 10529-10546.
- [2]. Leng, Z., Zhang, J., Ma, Y., & Zhang, J. (2020). Underwater topography inversion in liaodong shoal based on GRU deep learning model. *Remote Sensing*, 12(24), 4068.
- [3]. Iwakami, S., Tamega, M., Sanada, M., Mohri, M., Iwakami, Y., Okamoto, N., & Watanabe, M. (2021). Mathematical modeling and computational analysis of underwater topography with global positioning and echo sounder data. *Journal of Applied Mathematics and Physics*, 9(5), 1171-1179.
- [4]. Zhang, W., Zhuang, P., Sun, H. H., Li, G., Kwong, S., & Li, C. (2022). Underwater image enhancement via minimal color loss and locally adaptive contrast enhancement. *IEEE Transactions on Image Processing*, 31, 3997-4010.
- [5]. Zhuang, P., Li, C., & Wu, J. (2021). Bayesian retinex underwater image enhancement. *Engineering Applications of Artificial Intelligence*, 101, 104171.
- [6]. Zhou, J., Sun, J., Zhang, W., & Lin, Z. (2023). Multi-view underwater image enhancement method via embedded fusion mechanism. *Engineering Applications of Artificial Intelligence*, 121, 105946.
- [7]. Lin, H., Zhang, H., Li, Y., Wang, H., Li, J., & Wang, S. (2020). 3D point cloud capture method for underwater structures in turbid environment. *Measurement Science and Technology*, 32(2), 025106.
- [8]. Tsai, C. M., Lai, Y. H., Sun, Y. D., Chung, Y. J., & Shortis, M. (2021). Multi-dimensional underwater point cloud detection based on deep learning. *Sensors*, 21(3), 884.
- [9]. Wang, X. X., Gao, J., & Feng, L. (2020, August). Recognition and 3D pose estimation for underwater objects using deep convolutional neural network and point cloud registration. In *2020 International Conference on System Science and Engineering (ICSSE)* (pp. 1-6). IEEE.
- [10]. Hidalgo, F. (2020, October). ORBSLAM2 and point cloud processing towards autonomous underwater robot navigation. In *Global Oceans 2020: Singapore-US Gulf Coast* (pp. 1-4). IEEE.
- [11]. Balestrieri, E., Daponte, P., de Vito, L., & Lamonaca, F. (2021). Sensors and Measurements for Unmanned Systems: An Overview. *Sensors*, 21(4). <https://doi.org/10.3390/s21041518>
- [12]. Muja, M., & Lowe, D. (2009). Flann-fast library for approximate nearest neighbors user manual. Computer Science Department, University of British Columbia, Vancouver, BC, Canada, 5, 6.

Unified Framework for PdM Algorithm Development: The pdm-tools Architecture

Paweł Knap^{1,*}, Urszula Jachymczyk¹
AGH University of Krakow, Poland¹
pknap@agh.edu.pl

Abstract: This paper proposes a novel architecture for *pdm-tools*, a dedicated library designed to streamline the development process of predictive maintenance (PdM) algorithms within the Industry 4.0 paradigm. *pdm-tools* tackles the time-consuming nature of PdM algorithm development by encompassing five key components within a unified framework. *pdm-tools* tackles the time challenge by offering a unified workflow: generate synthetic vibration data (gearboxes, bearings, shafts) for rapid prototyping, eliminating initial real-world data collection. Data preparation and feature extraction ensure readiness through scaling, filtering, and extracting diverse condition indicators, empowering robust algorithms. Finally, user-defined model development and evaluation allow training, optimization, and selection of optimal PdM algorithms (machine learning/deep learning). By integrating these functionalities within a single architecture, *pdm-tools* empowers rapid development and evaluation of PdM algorithms, ultimately leading to a more efficient and effective implementation of predictive maintenance strategies within the Industry 4.0 landscape.

Keywords: PREDICTIVE MAINTENANCE, FAULT IDENTIFICATION, ANOMALY DETECTION, REMAINING USEFUL LIFE

1. Introduction

Predictive Maintenance (PdM) is an advanced maintenance strategy that leverages data analysis techniques and condition-monitoring technologies to predict when equipment failure might occur [1]. Unlike traditional maintenance approaches, such as reactive maintenance, which involves repairing equipment after a failure, or preventive maintenance, which schedules regular maintenance based on time or usage, PdM aims to minimize downtime and maintenance costs by performing maintenance activities just in time before a failure happens [2].

PdM relies on real-time data collected from various sensors and historical data to identify patterns and anomalies that indicate potential issues. By applying machine learning algorithms and statistical models, PdM systems can forecast the remaining useful life of machinery and recommend optimal maintenance actions [3, 4, 5]. This proactive approach not only enhances the reliability and efficiency of equipment but also extends its operational lifespan, thereby contributing to overall operational excellence [6, 7, 8].

libraries such as PyTorch, Scikit-learn, and TensorFlow. Although there are numerous repositories on GitHub related to PdM, none of them provide a complete framework that is reusable, adjustable, and scalable, which presents a significant gap in the current landscape of PdM development tools.

To address these challenges, we propose a unified *pdm-tools* framework designed to streamline the development of Predictive Maintenance (PdM) solutions. This framework facilitates the generation of synthetic vibration data for critical components such as gearboxes, bearings, and shafts, thereby mitigating the need for initial real-world data collection and expediting the prototyping process. Comprehensive data preparation and feature extraction methods ensure the data is preprocessed through scaling, filtering, and the extraction of diverse condition indicators, which are crucial for developing robust predictive algorithms. Additionally, the framework supports user-defined model development and evaluation, enabling the training, optimization, and selection of optimal PdM algorithms, encompassing both machine learning and deep learning techniques.

2. Framework overview.

The proposed framework encompasses a comprehensive suite of components designed to facilitate the development of Predictive Maintenance (PdM) solutions. This section provides an overview of the key components comprising the framework architecture (Fig 2.).

The framework includes a robust data generation module capable of producing synthetic vibration data and simulating faults in critical machinery components such as bearings, gearboxes, and shafts. This functionality enables users to generate data representative of real-world scenarios, facilitating the rapid prototyping of PdM solutions.

A critical aspect of PdM solution development involves preprocessing raw data to ensure its suitability for algorithmic analysis. The framework offers a flexible preprocessing module where users can perform tasks such as scaling, filtering, and normalization, particularly crucial when working with real data. This preprocessing step enhances the quality and reliability of the data, laying a solid foundation for subsequent analysis.

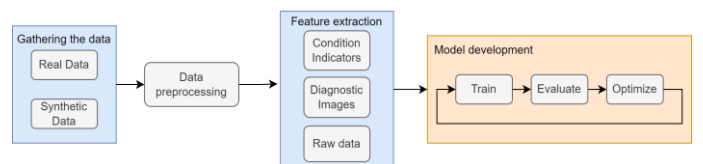


Fig. 2 High-level framework overview.

Fig. 1 Predictive Maintenance tasks diagram.

Predictive Maintenance (PdM) encompasses three core tasks (Fig 1.): anomaly detection, fault identification, and estimation of the remaining useful life (RUL). Anomaly detection involves monitoring equipment to identify deviations from normal operational behavior, signaling potential issues before they escalate. Fault identification goes a step further by diagnosing the specific nature and cause of detected anomalies, enabling targeted corrective actions [9, 10]. Lastly, the estimation of the remaining useful life predicts the time remaining before an asset is likely to fail, allowing for optimal scheduling of maintenance activities to prevent unexpected breakdowns and minimize downtime. Together, these tasks ensure a comprehensive and proactive approach to maintaining equipment health and performance.

Currently, most Predictive Maintenance (PdM) solutions are developed using scripts written in MATLAB or Python. These scripts often leverage a variety of MATLAB toolboxes and Python packages to create new algorithms. While MATLAB offers a comprehensive and well-suited solution for PdM, its high cost can be prohibitive. Consequently, in the fields of Deep Learning (DL) and Machine Learning (ML), scientists often prefer open-source

The feature extraction module provides users with a range of options tailored to their specific modeling requirements. Whether utilizing diagnostics images, condition indicators, or raw data, this module offers diverse feature extraction techniques to capture relevant information from the data. By extracting meaningful features, users can enhance the predictive capabilities of their models and uncover valuable insights into equipment health.

Central to the framework is the model development module, where users can design, evaluate, and optimize custom PdM models. Leveraging a combination of machine learning and deep learning methodologies, users have the flexibility to create models tailored to their unique application domains. This module also facilitates the optimization of model hyperparameters, ensuring optimal performance and generalization.

In the final stage of PdM solution development, users can evaluate the performance of their custom models and compare them against pre-existing implementations. This allows for comprehensive benchmarking and validation, enabling users to assess the efficacy of their solutions and make informed decisions regarding model selection and refinement. Additionally, users can leverage pre-existing models as starting points for their own implementations, further expediting the development process.

3. Data Generation

The data generation component of the framework employs sophisticated techniques to simulate various machine malfunctions and generate synthetic vibration data [11]. Through the application of specific amplitude and frequency modulation methods, the framework can accurately replicate fault conditions within machinery components. Certain faults, such as those occurring in shafts and gearboxes, are represented as sinusoidal components, while others, such as bearing faults, are modeled as decaying pulses. Moreover, the framework allows for the application of randomly generated frequency response functions corresponding to the machine under consideration, a technique commonly employed in modal analysis to identify resonant frequencies. This function is characterized by non-periodic and non-monotonic amplitude modulation, reflecting the dynamic nature of vibration signals produced by machines operating at varying speeds. By combining these features and parameters, the framework can simulate a range of faults including:

- shaft imbalance,
- bearing defects,
- and gearbox faults,

thereby enabling the generation of synthetic data representative of diverse operational conditions. This parametrization capability further enhances the versatility of the framework, facilitating the simulation of different fault scenarios for comprehensive testing and analysis.

4. Data preprocessing.

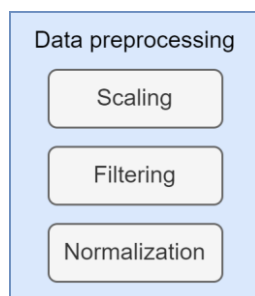


Fig. 3 Data preprocessing component.

The data preprocessing component of the framework is integral to ensuring that the data is suitably prepared for subsequent analysis

and model development. This component is comprised of three key parts (Fig. 3.):

- **Scaler** - The scaler module facilitates the scaling of measured values according to specified measurement parameters, like amplification and sensitivity.
- **Filtering** - The filtering module provides tools to design and apply various types of filters to the data. Users can implement low, high and band-pass filters (Fig. 4.).
- **Normalization** - The normalization module applies z-score normalization to the data, transforming it into a standard normal distribution. This process ensures that the data has a mean of zero and a standard deviation of one, facilitating improved performance and stability in machine learning algorithms.

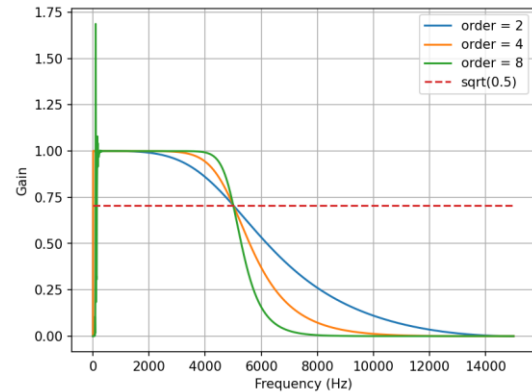


Fig. 4 Example of a band-pass filter (from 2 to 5000 [Hz]).

5. Feature extraction

The feature extraction component of the framework provides a comprehensive suite of tools to calculate various indicators and diagnostic images that are essential for developing robust predictive maintenance models [12]. This component supports the extraction of features in the time domain, frequency domain, and specific envelope indicators related to bearing faults, as well as the generation of diagnostic images.

Time Domain Indicators

The framework allows for the calculation of several time domain indicators, including:

- **Peak-to-peak (PP)**: The difference between the maximum and minimum values in the signal.
- **Root mean square (RMS)**: The square root of the mean of the squares of the values, reflecting the power content of the signal.
- **Crest factor (CF)**: The ratio of the peak value to the RMS value, indicating the presence of peaks.
- **Shape factor (SF)**: The ratio of the RMS value to the mean of the absolute values, providing insight into the waveform shape.
- **Standard deviation (STD)**: A measure of the amount of variation or dispersion in the signal.
- **Kurtosis**: A measure of the "tailedness" of the probability distribution of the signal.
- **Velocity root mean square (VRMS)**: The RMS value of the velocity signal.

Frequency Domain Indicators

In the frequency domain, the framework can calculate:

- **Mean frequency (MF)**: The average frequency content of the signal.
- **Frequency Center (FC)**: The center of mass of the power spectrum.
- **Three maximal amplitude peaks and their corresponding frequencies**: The highest peaks in the frequency spectrum and their associated frequencies.

Envelope Indicators

For bearing fault detection, the framework includes envelope indicators that are related to specific bearing fault frequencies:

- Fundamental Train Frequency (FTF): Related to the cage or train of the bearing.
- Ball Pass Frequency of the Outer Race (BPFO): The frequency at which a ball passes over a defect on the outer race.
- Ball Pass Frequency of the Inner Race (BPFI): The frequency at which a ball passes over a defect on the inner race.
- Ball Spin Frequency (BSF): The frequency at which a ball spins within the bearing.

Diagnostic Images

The framework also provides capabilities for generating various diagnostic images, which are useful for visual analysis and model input:

- Fast Fourier Transform (FFT): Converts the time-domain signal into its frequency components.
- Power Spectral Density (PSD): Estimates the power distribution of a signal over frequency.
- Short-Time Fourier Transform (STFT): Represents the signal in both time and frequency domains.
- Continuous Wavelet Transform (CWT): Provides a time-frequency representation of the signal.
- Kurtogram: A visual tool to identify the frequency bands where kurtosis is maximized, indicating potential faults.

6. Model development

The model development, evaluation, and comparison component of the framework offers users a robust environment to design, optimize, and assess their predictive maintenance (PdM) models. Users can develop their models utilizing popular machine learning libraries such as TensorFlow and Scikit-learn, depending on the specific algorithm requirements [13].

Model Development

The framework supports the creation and customization of models using a wide array of machine learning (ML) and deep learning (DL) algorithms. By leveraging TensorFlow for DL models and Scikit-learn for ML algorithms, users can construct models tailored to their specific PdM applications.

Optimization Algorithms

A key feature of the framework is its integration of advanced optimization algorithms to enhance model performance. Users can utilize:

- Grid Search - A systematic approach to hyperparameter tuning by exhaustively searching through a specified parameter grid.
- Random Search - An alternative to grid search that randomly samples hyperparameters, often yielding comparable performance with reduced computational cost.
- Bayesian Optimization - A probabilistic model-based approach that efficiently navigates the hyperparameter space to find optimal settings.
- Hyperband - A resource-efficient optimization algorithm that dynamically allocates resources to promising hyperparameter configurations, particularly suited for DL models.

Base Solutions and Comparison

The framework provides a set of pre-implemented base solutions that users can employ for benchmarking purposes. These baseline models facilitate a straightforward comparison, enabling users to

evaluate the performance of newly developed algorithms against established standards.

Automated Evaluation

To streamline the evaluation process, the framework automatically generates confusion matrices (Fig 5.) and computes essential performance metrics, including precision, accuracy, recall, and the F1 score, where applicable. These metrics provide a comprehensive assessment of model performance, aiding users in the selection and refinement of their PdM algorithms.

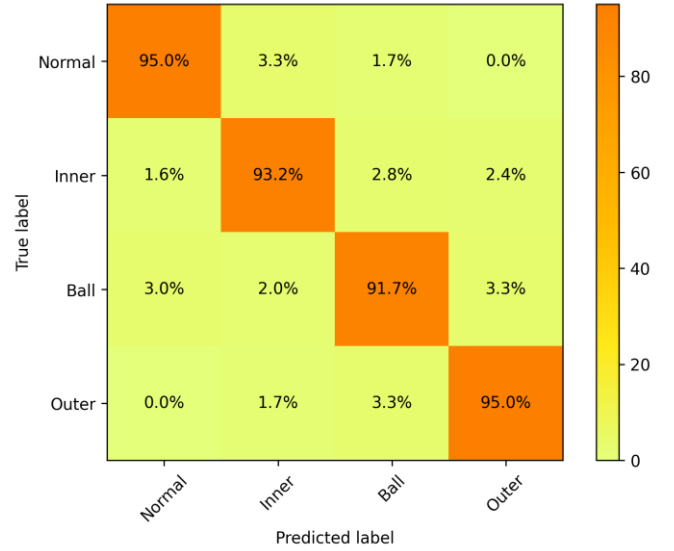


Fig. 5 Example confusion matrix.

7. Conclusions

The proposed architecture allows for the rapid, scalable, and adjustable development of Predictive Maintenance (PdM) solutions. By leveraging well-established packages such as NumPy, SciPy, TensorFlow, Matplotlib, and Scikit-learn, the framework ensures robustness and flexibility in handling specific tasks associated with PdM.

One of the key advantages of this framework is its capability to generate synthetic data, which can be particularly useful in providing additional information to enhance algorithm performance. This feature significantly reduces the initial dependency on real-world data collection, thereby accelerating the prototyping phase.

Furthermore, the framework enables users to benchmark their solutions against predefined baseline models. This comparative capability is a significant advantage, as it allows for an objective evaluation of new algorithms, facilitating the identification of the most effective approaches.

However, it is important to note that the current framework does not address the deployment phase of PdM solutions. Deployment is a critical aspect that involves integrating developed models into real-world systems, ensuring they operate efficiently under actual operating conditions. Future work should focus on incorporating deployment capabilities to provide a comprehensive end-to-end solution for PdM.

Overall, the architecture presented offers a robust platform for the development of PdM solutions, with strengths in synthetic data generation, comprehensive data preprocessing, and advanced model development and evaluation. By addressing the deployment phase in future iterations, the framework can become an even more powerful tool for predictive maintenance applications.

References

1. Hu, Yang, et al. "Prognostics and health management: A review from the perspectives of design, development and decision." *Reliability Engineering & System Safety* 217 (2022): 108063.M.
2. Ran, Yongyi, et al. "A survey of predictive maintenance: Systems, purposes and approaches." *arXiv preprint arXiv:1912.07383* (2019).
3. Krupitzer, Christian, et al. "A survey on predictive maintenance for industry 4.0." *arXiv preprint arXiv:2002.08224* (2020).
4. Tiddens, Wieger, Jan Braaksma, and Tiedo Tinga. "Decision framework for predictive maintenance method selection." *Applied Sciences* 13.3 (2023): 2021.
5. Namuduri, Srikanth, et al. "Deep learning methods for sensor based predictive maintenance and future perspectives for electrochemical sensors." *Journal of The Electrochemical Society* 167.3 (2020): 037552.
6. Lalik, Krzysztof, et al. "Simulation Modeling for Mass Customization Furniture Production: Investigating Production Volume and Machine Load in scope of Industry 4.0 standard." *Industry 4.0* 8.4 (2023): 116-122.
7. Lalik, Krzysztof, et al. "The steam pressure impacts reducing system for a biomass cogenerator based on monitoring of the frequency characteristics of the steam actuator." *Vibrations in Physical Systems* 30.2 (2019).
8. Lalik, Krzysztof, et al. "Self-powered wireless sensor matrix for air pollution detection with a neural predictor." *Energies* 15.6 (2022): 1962.
9. Bałazy, P., Paweł Gut, and Paweł Knap. "Neural classifying system for predictive maintenance of rotating devices." *IOP Conference Series: Materials Science and Engineering*. Vol. 1239. No. 1. IOP Publishing, 2022.
10. Knap, Paweł, Krzysztof Lalik, and Patryk Bałazy. "Boosted Convolutional Neural Network Algorithm for the Classification of the Bearing Fault Form 1-D Raw Sensor Data." *Sensors* 23.9 (2023): 4295.
11. Jablonski, Adam. *Condition monitoring algorithms in Matlab*. Springer Nature, 2021.
12. Oppenheim, Alan V. *Discrete-time signal processing*. Pearson Education India, 1999.
13. Aurélien Géron. „Hands-on Machine Learning with Scikit Learn, Keras, and TensorFlow, Concepts, Tools, and Techniques to Build Intelligent Systems (second edition)”. Sebastopol, CA: O’Reilly Media, Inc., 2019.

Simulation experiment for the follow-up controller of the MIMO system

Kamil Pieprzycki¹, Urszula Jachymczyk²
 AGH University of Krakow
 Krakow, Poland
 pieprzycki.kp@gmail.com, ulajach1@gmail.com

Abstract: Controlling Multi-Input Multi-Output (MIMO) systems, such as portal conveyors, poses significant challenges due to their inherent complexity and variability. Traditional control methods often fall short in handling the dynamic and nonlinear nature of these systems. This paper presents a novel reinforcement learning (RL) approach, leveraging the twin-delayed deep deterministic policy gradient (TD3) algorithm, to develop a follow-up controller that is robust to changes in system parameters. Our simulation experiments demonstrate the effectiveness of this method.

Keywords: REINFORCEMENT LEARNING, TD3, LOAD MASS VARIATION, SIMULATION STUDY, MIMO SYSTEMS

1. Introduction

The increasing complexity and nonlinearity of modern control systems present significant challenges for achieving effective control. MIMO systems [1], which involve multiple inputs and outputs interacting simultaneously, exemplify these challenges. Traditional methods based on differential equations often fail to capture the chaotic and unstable behavior of such systems accurately. Several factors contribute to the complexity of control systems:

- **State Variables:** Modern control systems include numerous state variables that describe the system's current state, making management difficult.
- **Interconnected Subsystems:** Many control systems consist of interconnected subsystems that introduce complex interactions and feedback loops.
- **Parameter Uncertainty:** Production variability and environmental conditions introduce uncertainty in system parameters.
- **Nonlinearity:** Nonlinear systems exhibit disproportionate responses to changes in state variables and control actions, complicating control design.

Traditional methods struggle with these issues, necessitating robust and adaptive control strategies. Reinforcement Learning (RL) offers a promising alternative by enabling systems to adapt to dynamic environments and maintain optimal performance. RL approaches, particularly those leveraging deep learning techniques, can handle the complexities of nonlinear and dynamic systems more effectively than traditional methods. The RL agent learns optimal control strategies through a process of trial and error, continuously improving its performance based on feedback from the environment. This adaptability makes RL a powerful tool for controlling complex systems.

2. Object of study

In this chapter, we will focus on presenting a specific object on which RL control research will be carried out. Portal conveyor has been chosen as the object, which is an example of a multidimensional system. The study will be carried out using a simplified mathematical model that considers the variable dynamics of the system due to the variable weight of the load and the variable length of the sling.

Object presentation

The object consists of cart that moves in two axes and a sling that moves in one axis and swings in two planes. The following is an overview diagram of the crane (Fig. 1).

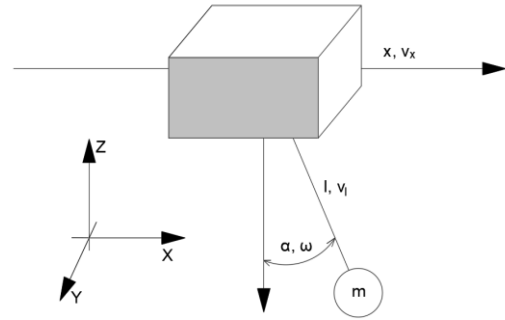


Fig. 1 Final confusion matrices for each label

The model can be described by the state variables shown in Table I.

Table 1: Portal Conveyor Variables

Symbol	Description
x	X-axis cart position
v_x	X-axis cart velocity
α	Angle between the lift-line and Z axis
ω	Angular velocity of the load on XZ plane
L	Length of the lifted load
v_l	Sling extraction speed

The following equations (1-6) are representations of the object's state-space, where F_x/F_l are forces acting on cart/payload and m_c/m_l are weights of the cart/payload.

$$\dot{z}_1 = \dot{x} = z_2 \quad (1)$$

$$\dot{z}_2 = \ddot{x} = \frac{F_x}{m_c + m_l} + \frac{m_l g}{m_c + m_l} \alpha \quad (2)$$

$$\dot{z}_3 = \dot{\alpha} = z_4 \quad (3)$$

$$\dot{z}_4 = \ddot{\alpha} = \frac{-F_x}{l(m_c + m_l)} - \frac{g(m_c + 2m_l)}{l(m_c + m_l)} \alpha \quad (4)$$

$$\dot{z}_5 = \dot{l} = z_6 \quad (5)$$

$$\dot{z}_6 = \ddot{l} = \frac{F_l}{2m_l} \quad (6)$$

Non-linearity of the object

Portal Conveyor are nonlinear systems because they cannot be described by linear equations. This is due to several factors,

including the nonlinear dependence of force on speed, torque on angular velocity, and frictional force on speed and angle of rotation. The variable length of the sling and the variable weight of the load also create additional nonlinearity in the crane's behavior. To simplify the mathematical model, assumptions were made about a rigid sling and small rocking angles.

3. Methodology

Overview

This study endeavors to develop and assess the efficacy of a reinforcement learning (RL) algorithm specifically designed for the control of object described in paper. The system consists of an agent and the environment with which the agent interacts. The conveyor is modeled using state space equations, which are recalculated in each iteration to account for the changing sling length and load mass. This approach ensures that the simulation accurately reflects the non-linear behavior of the system, particularly under varying load conditions. In our proposed system, rewards are determined by penalizing deviations from desired performance metrics, including precise load positioning and minimization of sling sway during operation. The RL agent, functioning as the conveyor controller, learns optimal actions through a process of trial and error, iteratively refining its strategy based on the rewards received from the environment [2]. This iterative learning process is visually depicted in Fig. 2, illustrating the dynamic interaction between the RL agent and its environment.

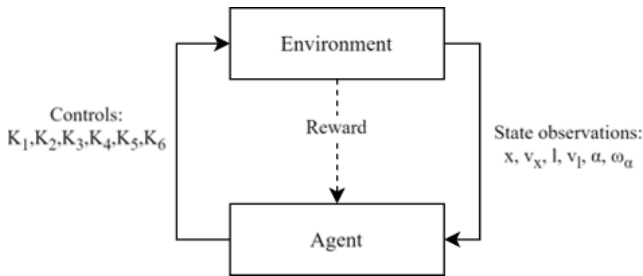


Fig. 2 System scheme.

Reward function

Another thing worth mentioning is the reward feature. The reward function (7) assesses how close the agent's behavior was to ideal in a given iteration [3]. The higher the value of the reward function, the more satisfactory the agent's performance has been achieved. The type of reward function that the authors chose can be called a penalty function, because the function coefficients are negative [4]. The closer to the target regulation point, the smaller the penalty, hence the maximum, but unattainable, reward value from a single simulation is zero.

$$\text{Reward} = -a * e - b * \alpha \quad (7)$$

Hyperparameters a and b should be selected depending on the weights of individual parts of the function. Parameter e is the carriage linear position deviation, a & b are the weights of function components. Taking into account the value of the position error, the swinging angle of the sling

4. Network training

Neural Network Architecture

The decision to employ the Reinforcement Learning Deep Deterministic Policy Gradient (rDDPG) agent stemmed from its aptness for environments characterized by continuous action spaces and observations. Within the rDDPG framework, both the actor and critic functions are realized as deep neural networks [5]. This

algorithm, abbreviated as Deep Deterministic Policy Gradient, represents an actor-critic, model-free methodology rooted in the deterministic policy gradient method. The actor network is constructed with fully connected layers, followed by leaky rectified linear unit (ReLU) activations, culminating in a hyperbolic tangent (tanh) layer responsible for outputting the action values [6]. Conversely, the critic network features distinct pathways for state and action inputs, which are then fused and processed through additional layers to yield the Q-value [7]. For a comprehensive understanding of the neural network architecture, please refer to Tables 2-5 for detailed information.

Table 2: Actor

Layer	Model	Activations
Input	8	-
Dense	64	LeakyReLU
Dense	16	LeakyReLU
Dense	8	Tanh

Table 3: Critic state path

Layer	Model	Activations
Input	8	-
Dense	64	LeakyReLU
Dense	32	-

Table 4: Critic action path

Layer	Model	Activations
Input	4	-
Dense	32	-

Table 5: Critic common path

Layer	Model	Activations
Dense	2	LeakyReLU
Output	1	-

Additionally, the interaction between the actor and critic is depicted in Fig. 3, showcasing their collaborative dynamics in the training process.

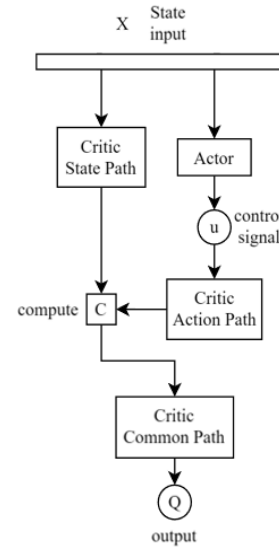


Fig. 3. Network architecture.

Training

In the training phase of our reinforcement learning (RL) algorithm, it is imperative to acknowledge the computational intensity of this process, as each episode within the Simulink environment involves the RL agent interacting with the environment, learning optimal actions through trial and error, and updating its policy based on received rewards. Throughout this iterative process, the agent strives to maximize its cumulative

reward, ultimately honing its decision-making capabilities to achieve the desired control objectives. An important indicator of the learning progress is the convergence of the average reward towards zero, signifying that the agent is increasingly proficient in navigating the environment and accomplishing its goals. Fig. 4 offers a visual representation of this learning progress, depicting the evolution of the agent's performance over time. Analyzing this progression provides valuable insights into the dynamics of the training process, including the rate of learning and potential areas for improvement. By harnessing the computational power of our hardware setup and monitoring the training progress through visualizations such as Fig. 3, we can effectively guide the RL agent towards achieving optimal performance. This rigorous training process lays the foundation for robust and reliable control of non-linear objects, ensuring safety, efficiency, and precision in real-world operations.

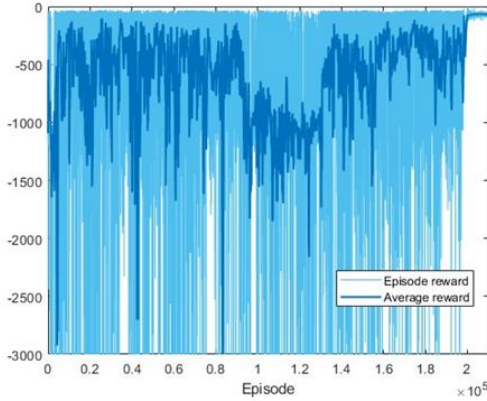


Fig. 4. Training process.

For each episode, the agent takes actions in 1000 iterations, as the time sample is set to 10 ms, resulting in an overall simulation time of 10 seconds. The training process stopped after almost 210,000 simulations (Fig. 3). During the learning process, the agent tuned the policy correctly more often than incorrectly. Finally, after 200,000 simulations, the agent fine-tuned the policy such that each subsequent simulation was performed properly. The final tuning of the neural network weights continued for the next 10,000 simulations. The decision to stop the learning process was made due to the reward level remaining at an appropriate level.

4. Results

In this section, we present the outcomes from simulations that attained the highest overall reward. We evaluated our system's performance under two different masses: Mass A, which is 1kg and Mass B, which is 5kg, utilizing specified control indicators to ensure effective regulation. The simulations for crane movement lasted 10 seconds in both cases. The plotted data, including dynamic variables x , v_x , α , ω , l , v_l correspond to the most successful cases. The Performance Evaluation for Mass A is illustrated in Fig. 5. The graph above illustrates the response of the system when subjected to varying conditions associated with Mass A. Through comprehensive analysis, it was observed that the system adequately met the predefined control indicators, including settling time, overshoot, and steady-state error. These results indicate that the system effectively regulated the dynamics of Mass A, demonstrating its capability to maintain stability and achieve desired performance levels.

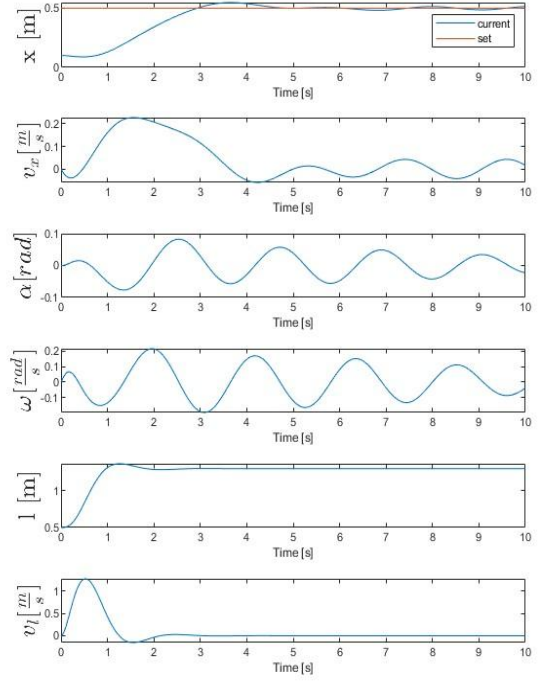


Fig.5 Performance of crane with load mass A [1kg].

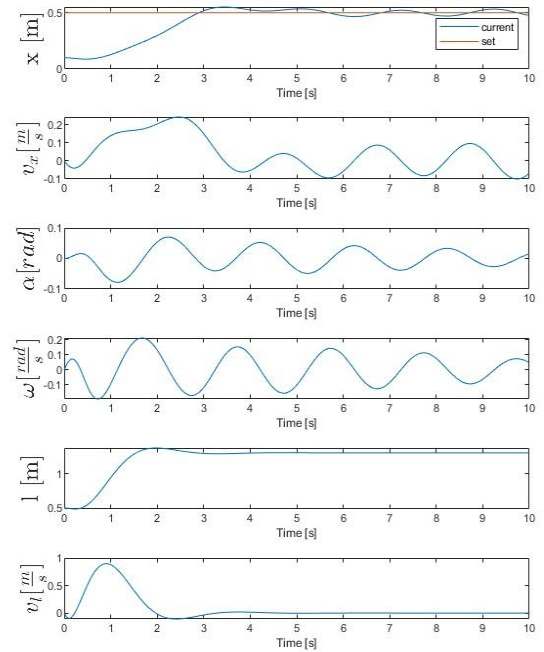


Fig.6 Performance of crane with load mass A [1kg].

Similarly, the graph representing the performance evaluation for Mass B (Fig. 6) showcases the system's response under different operational parameters. Upon detailed examination, it was evident that the system consistently satisfied the prescribed control criteria. Notably, the settling time, overshoot, and steady-state error remained within acceptable bounds, underscoring the system's ability to regulate the dynamics of Mass B effectively. To further analyze the system's behavior, we plotted Gain values against the weight. These plots provide additional insights into how the system's performance varies with different masses.

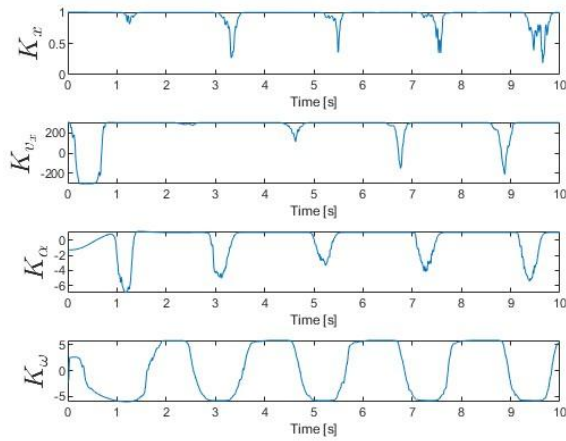


Fig. 7 Gains for mass A [1kg].

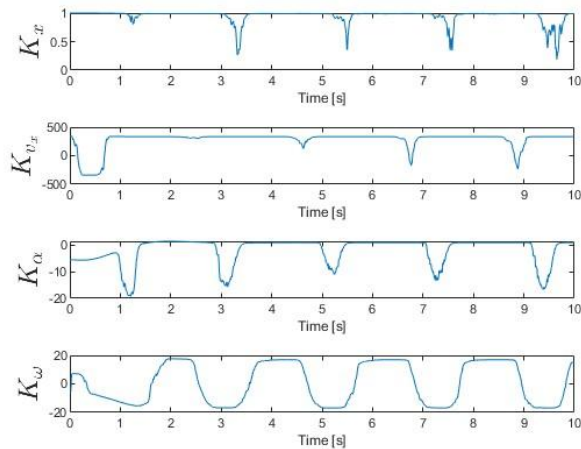


Fig. 8 Gains for mass B [5kg].

Overall, the results obtained from the evaluation of our system for both Mass A and Mass B demonstrate its robustness and reliability in achieving satisfactory regulation. By successfully meeting the specified control indicators under varying mass conditions, our system showcases its versatility and applicability across a diverse range of scenarios and environments. These findings substantiate effectiveness of our system in providing stable and controlled operation, thereby fulfilling the intended objectives. For comparison with the generally used LQR controller, the feedback gains for the linear-quadratic controller would need to be recalculated for each time instance [7]. The reinforcements produced by the agent for the two different masses are very similar in shape to each other.

5. Conclusions

In summary, our approach has yielded promising results, demonstrating the efficacy of our system in regulating the dynamics of masses A and B. The performance evaluation revealed that our system successfully met the specified control indicators, including settling time, overshoot, and steady-state error, for both Mass A and Mass B. These findings underscore the validity and effectiveness of our approach in addressing the challenges associated with dynamic mass regulation. The consistent achievement of satisfactory regulation across varying mass conditions indicates the robustness and reliability of our system. Encouraged by these results, we are motivated to continue our research efforts in this domain. Moving forward, we plan to explore further optimizations and enhancements to our system, with the aim of extending its applicability to broader contexts and addressing more complex challenges. One significant step in this direction involves conducting experiments on a real

system, rather than relying solely on simulations. By transitioning from simulations to real-world testing, we aim to validate the performance of our system in practical scenarios and gain deeper insights into its behavior under real-world conditions. Leveraging the insights gained from this study, we are confident in our ability to refine and advance our system. We envision incorporating advanced control strategies and algorithms, leveraging the latest developments in Reinforcement Learning and control theory. Through ongoing research and experimentation, we remain committed to realizing the full potential of our approach and making meaningful contributions to the field of dynamic mass regulation. Furthermore, our future experiments on real crane systems will allow us to validate the scalability and adaptability of our approach across different industrial settings. By collaborating with industry partners and crane operators, we aim to tailor our system to specific operational requirements and address any practical challenges that may arise. Ultimately, our goal is to develop innovative solutions that not only improve the efficiency and safety of crane operations but also pave the way for the adoption of advanced control technologies in various industrial applications.

6. References

- Derigent, William, Olivier Cardin, and Damien Trentesaux. "Industry 4.0: contributions of holonic manufacturing control architectures and future challenges." *Journal of Intelligent Manufacturing* 32.7 (2021): 1797-1818.
- Fulton, Nathan, and André Platzer. "Verifiably safe off-model reinforcement learning." *International Conference on Tools and Algorithms for the Construction and Analysis of Systems*. Cham: Springer International Publishing, 2019.
- Bałaży, Patryk, Paweł Gut, and Paweł Knap. "Positioning algorithm for AGV autonomous driving platform based on artificial neural networks." *Robotic Systems and Applications* 1.2 (2021): 41-45.
- Zhu, Meixin, Xuesong Wang, and Yinhai Wang. "Human-like autonomous car-following model with deep RL." *Transportation research part C: emerging technologies* 97 (2018): 348-368.
- Darwish, Ahmed, Momen Khalil, and Karim Badawi. "Optimising public bus transit networks using deep RL." *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2020.
- Lalik, K., Kozek, M., Dominik, I., Łukasiewicz, P. (2020). Adaptive MRAC Controller in the Effector Trajectory Generator for Industry 4.0 Machines. In: Bartoszewicz, A., Kabziński, J., Kacprzyk, J. (eds) *Advanced, Contemporary Control. Advances in Intelligent Systems and Computing*, vol 1196. Springer, Cham.
- Lalik, K.; Kozek, M.; Dominik, I. Autonomous Machine Learning Algorithm for Stress Monitoring in Concrete Using Elastoacoustical Effect. *Materials* 2021, 14, 4116. <https://doi.org/10.3390/ma14154116>
- Zhang, Zidong, Dongxia Zhang, and Robert C. Qiu. "Deep RL for power system applications: An overview." *CSEE Journal of Power and Energy Systems* 6.1 (2019): 213-225.

Expanding the Capabilities of Data Transmission Technology to Improve the Efficiency of Multi-Channel Communication Systems

Tatiana Sestacova, Gherman Sorochin

Technical University of Moldova, Chisinau, Republic of Moldova

tatiana.sestacova@sde.utm.md, gherman.sorochin@sde.utm.md

Abstract: The article, been a continuation of previous studies, proposes and analyzes one of the ways to increase the efficiency of multi-channel communication systems through the simultaneous use of several derivative systems of Walsh functions with different generating functions, which are necessary for constructing pseudo-random sequences (PSR). These sequences, in turn, are used as spreading codes when generating noise-like signals in multi-channel data transmission systems with channel separation (SDMA) according to the signal shape or code. The correlation properties of the PSR are studied and it is shown that, in the contrast of the known ones, the proposed approach makes it possible to increase the number of channels in the system without increasing the channel bandwidth that improves the efficiency of a multi-channel communication system.

Keywords: CDMA SYSTEM, NOISE-LIKE SIGNALS, PSEUDO-RANDOM SEQUENCES, WALSH FUNCTION DERIVATIVE, GENERATING FUNCTION.

1. Introduction

It is known that from the point of view of using time and frequency resources, the most effective are multi-channel systems with channel separation according to signal shape [1-5]. On the other hand, to increase the noise immunity of such systems, it is desirable to use broadband noise-like signals (NLS), since such signals are little sensitive to powerful narrow-band interference, to multipath propagation, and have data transmission secrecy [2, 7, 8].

Increasing the number of multiplexed channels without increasing the number of physical lines is possible by endowing signals with special individual characteristics that would lead to the distinction of multiplexed channels for the purpose of their further separation. In general, such signs can be the parameters of signal carriers: amplitude, frequency, and phase in the case of continuous modulation, temporal position, duration or shape of pulses in discrete modulation.

In systems with channel division by signal form (code), each channel is assigned a specific signal form, which is a distinctive feature, an identification “address” of a given channel (subscriber). At the same time, the information itself can be introduced into a broadband signal in several ways. For example, a narrowband information signal is multiplied by a pseudo-random sequence (spreading code) with a period T , consisting of N bits of duration τ_0 each. In this case, the NLS base is numerically equal to the number of extension code elements.

Since the correlation characteristics of NPS are completely determined by the correlation characteristics of spreading codes [2, 4, 6], the decoding of such signals on the receiving side is carried out by a correlation receiver, the basis of which is a correlator, which is a multiplier and integrator connected in series, calculating the correlation function (CF) input signal from a copy of the extension code stored in memory.

2. Theoretical background

Discrete signals with the best structure of the cross-correlation function (CCF) include Walsh functions, and the number of functions in the Walsh system is equal to $N = 2^m$, where m is an integer [8, 9, 11].

Orthogonal extension codes, such as Walsh sequences, can be used only if all users of a single CDMA channel are synchronized to within a small fraction of an elementary signal. Due to the fact that the mutual correlation of different shifts of Walsh sequences is not zero, pseudo-random sequences are required to be used in the absence of precise synchronization.

On the other hand, from the point of view of reducing multiple access interference, spreading codes (or NLs) should have a cross-correlation function as small as possible (ideally zero).

In order to improve the properties of *aperiodic* autocorrelation functions (AACF) and cross-correlation functions (ACCF), Walsh signal systems often build so-called derivative signal systems [2, 10]. A derivative is a signal obtained as a result of element-by-element (symbol-by-symbol) multiplication of two signals – the original and the generating producing. The signal from

the Walsh matrix acts as the original signal, and the spreading code acts as the generating signal. A system consisting of derivative signals is called a derivative.

A signal is chosen as a generating signal such that the derivative system has good correlation properties. As a rule, this is a signal with good AACF, and it is found using a computer or other method.

The discrete signals with the best ACF structure include Barker signals (codes), M-sequences, de Bruijn sequences [1, 2, 12]. Only seven Barker codes are known, and there are only two even sequences (two and four characters long). The length of M-sequences is $L = 2^n - 1 = 2^3 - 1 = 7$ symbols, and the length of de Bruijn sequences is $L = 2^n = 2^3 = 8$ symbols.

But as noted above, the number of functions in the Walsh system, as well as the length of each of its functions, is equal to a power of two, i. e. for $n = 3$, the length of the Walsh function is eight. Therefore, for the generating function, it is better to use de Bruijn sequences to construct derivative systems of Walsh functions. For $n = 3$, the following de Bruijn sequences can be obtained using the appropriate nonlinear feedback shift registers: 00010111, 10111000, 11101000.

To increase the efficiency of a multichannel communication system, it is desirable to use as many information sources (subscribers) as possible. The number of subscribers is determined by the number of different identifying "addresses" of channels (subscribers), i. e. the number of functions in the Walsh derivative system (spreading codes). Consequently, for $n = 3$ only 8 channels can be organized.

It is proposed to take another way: **for a given bandwidth, an increase in the number of channels (subscribers) can be achieved using several systems of derivative Walsh functions obtained by different generating functions.** For example, for the de Bruijn sequences mentioned above, the following derivatives of the Walsh system can be obtained (Fig. 1).

Thus, with the presented Walsh function derivative systems, it is possible to create a multi-channel system with 32 channels ***without increasing the communication channel bandwidth.***

1	1	1	1	1	1	1	1
1	1	1	1	-1	-1	-1	-1
1	1	-1	-1	-1	-1	1	1
1	1	-1	-1	1	1	-1	-1
1	-1	-1	1	1	-1	-1	1
1	-1	-1	1	-1	1	1	-1
1	-1	1	-1	-1	1	-1	1
1	-1	1	-1	1	-1	1	-1

a)

-1	-1	-1	1	-1	1	1	1
-1	-1	-1	1	1	-1	-1	-1
-1	-1	1	-1	1	-1	1	1
-1	-1	1	-1	-1	1	-1	-1
-1	1	1	1	-1	-1	-1	1
-1	1	1	1	1	1	1	-1
-1	1	-1	-1	1	1	-1	1
-1	1	-1	-1	-1	-1	1	-1

b)

1	-1	1	1	1	-1	-1	-1
1	-1	1	1	1	-1	1	1
1	-1	-1	-1	-1	1	-1	-1
1	-1	-1	-1	1	-1	1	1
1	1	-1	1	1	1	1	-1
1	1	-1	1	1	-1	-1	1
1	1	1	-1	-1	-1	1	-1
1	1	1	-1	1	1	-1	1

c)

1	1	1	-1	1	-1	-1	-1
1	1	1	-1	-1	1	1	1
1	1	-1	1	-1	1	-1	-1
1	1	-1	1	1	-1	1	1
1	-1	-1	-1	1	1	1	-1
1	-1	-1	-1	-1	-1	-1	1
1	-1	1	1	1	-1	1	-1
1	-1	1	1	1	1	-1	1

d)

Fig. 1. Systems of derivatives of Walsh functions:

- a) Walsh's original system – W_8 ; b) Derivative system of Walsh functions with generating function $-1-1-11-1111 - D_1$
c) Derivative system of Walsh functions with generating function $1-1111-1-1-1 - D_2$; d) Derivative system of Walsh functions with generating function $1-1111-1-1-1 - D_3$

B. Experimental data

Let us consider the correlation properties of the derivatives of the Walsh systems presented in Fig. 1, for example, the 3rd and 5th from the D_1 system, the 4th and 6th from the D_2 system and the 1st and 7th from the D_3 system. Moreover, we will consider auto-correlation functions (ACF) and cross-correlation functions (CCF) of the selected functions both within the systems themselves and between systems. The results of calculating the correlation functions of the selected derivatives of the Walsh functions are presented below.

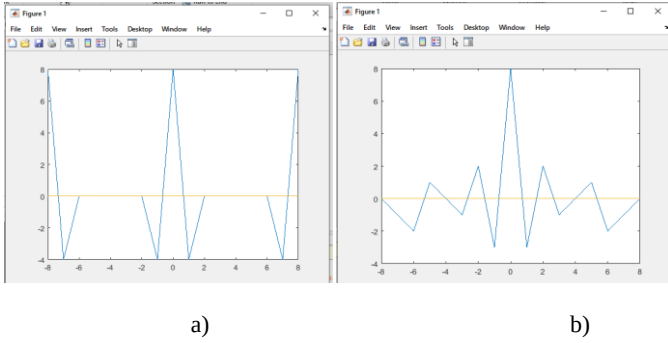


Fig. 2. Correlation characteristics of the 3rd derivative of the Walsh function from the D_1 system
a) PAKF; b) AAKF

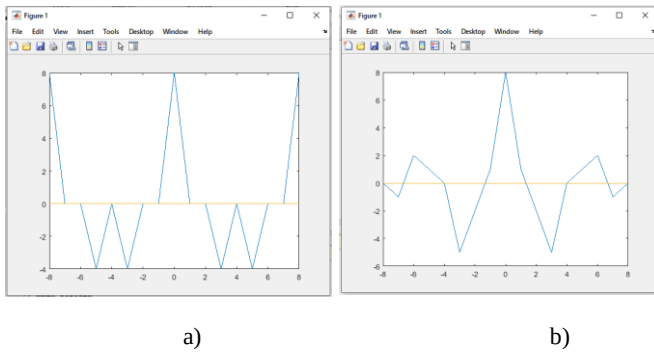


Fig. 3. Correlation characteristics of the 5th derivative of the Walsh function from the D_1 system
a) PAKF; b) AAKF

As a comparison, we present the autocorrelation characteristics of the 5th original Walsh function (Fig. 4) and the cross-correlation characteristic of the 3rd and 5th derivatives of the Walsh functions from the D_1 system (Fig. 5)

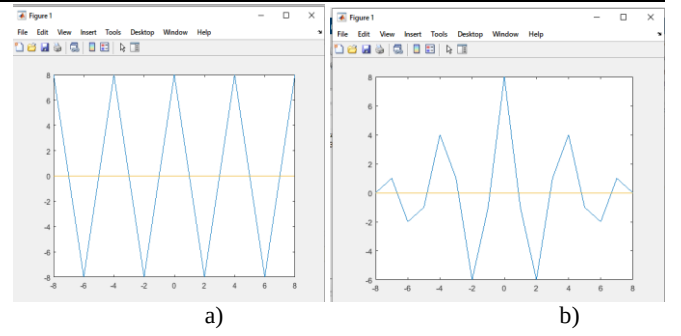


Fig. 4. Correlation characteristics of the 5th original of the Walsh function
a) periodic PAKF; b) aperiodic AAKF

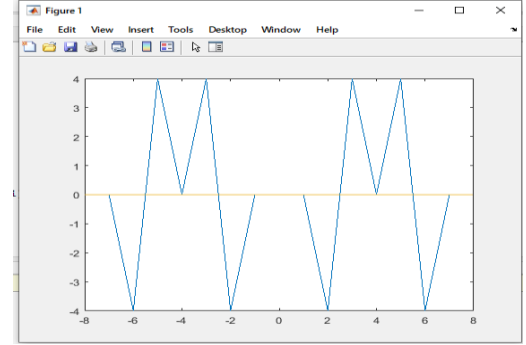


Fig. 5. Cross-correlation characteristic of the 3rd and 5th derivatives of the Walsh functions from the D_1 system

Similarly, the correlation characteristics of the original and derivative Walsh functions, as well as their mutual characteristics from the D_2 and D_3 systems, were studied. The results are presented in Fig. 6, 7.

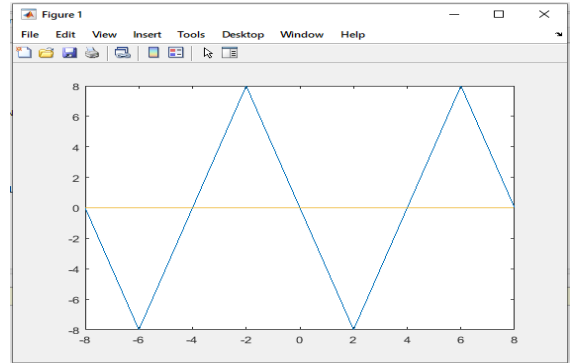


Fig. 6. Cross-correlation characteristic of the 2nd and 3rd original Walsh functions

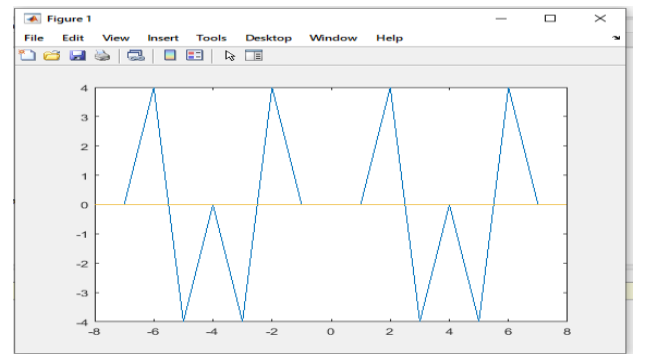


Fig. 7. Cross-correlation characteristic of the 1st and 7th derivatives of the Walsh functions from the D_3 system

Let us now consider what form the mutual correlation characteristics of the derivatives of the Walsh functions, which are taken from different systems, have. For example, the cross-correlation characteristic between the 3rd derivative of the Walsh function from the D_1 system and the 7th derivative of the D_3 system (Fig. 8), or the cross-correlation characteristic between the 3rd derivative of the Walsh function from the D_1 system and the 6th derivative from the D_2 system (Fig. 9).

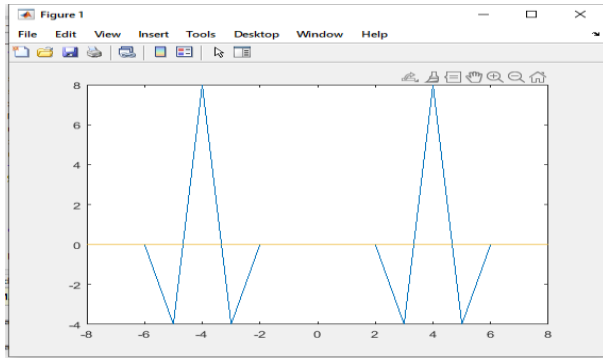


Fig. 8. Cross-correlation characteristic between the 3rd derivative of the Walsh function from the D_1 system and the 7th derivative of the Walsh functions from the D_3 system

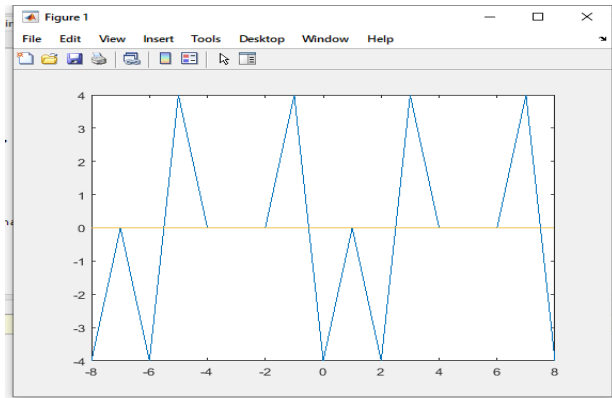


Fig. 9. Cross-correlation characteristic between the 3rd derivative of the Walsh function from the D_1 system and the 6th derivative of the Walsh functions from the D_2 system

C. Results and discussion

As shown by the analysis of the correlation characteristics of the original and derivative Walsh functions (Fig. 2, Fig. 3), the autocorrelation characteristics (PACF and AACF) of the derivatives of the Walsh functions have *better indicators* than the autocorrelation characteristics of the original Walsh functions (Fig. 4). The central lobe has an amplitude equal to the number of symbols (Fig. 8), the suppression coefficient is equal to the ratio of the amplitude of the central peak of the ACF to the maximum amplitude of the side lobes and lies in the range from 1.6 to 2.67. This makes it possible to distinguish such signals well when exposed to noise and ensure reliable synchronization of decoding devices.

The analysis of cross-correlation characteristics of the investigated derivative Walsh functions taken from separate systems (Fig. 5, Fig. 6, Fig. 7) and from different systems (Fig. 8, Fig. 9) showed that the considered derivative Walsh functions have satisfactory characteristics. The CCF between the 3rd and 5th derivatives of the Walsh function from system D_1 have emission amplitudes in the range of ± 4 and zero CCF value at $\pm \tau$ shifts, similar characteristics have the 1st and 7th derivatives of the Walsh function from system D_3 with a slightly different emission location.

Similar results were obtained when comparing the mutual correlation characteristics of other variants of the derivatives of the Walsh functions taken from the systems D_1 , D_2 , D_3 .

Conclusion

The analysis of correlation properties of the studied spreading codes based on derivative Walsh functions allows to draw the following conclusions:

1. The *correlation properties* of the derived Walsh functions have *much better* correlation characteristics than the original Walsh functions.
2. The *cross-correlation functions* of the derived Walsh functions taken from the derived Walsh system itself or from different systems have satisfactory characteristics. They can be used in multichannel communication systems with the application of solvers, which do not react to CCF outliers within the specified limits, in order to reduce the level of multiple access interference.
3. The derivatives of the Walsh function ($L=8$ digits) have a large amplitude of the central peak of the ACF equal to the sequence length, but the amplitude of the side lobes increases slightly. This should be taken into account when designing multichannel systems to ensure *reliable synchronization* of receiving devices under the influence of noise.
4. The large length of the spreading code based on the Walsh functions derivatives allows distributing the signal energy over the spectrum, increasing the noise immunity of the system, providing good protection against unauthorized access, and improving electromagnetic compatibility with neighboring radio engineering systems. But it is necessary to provide special protection measures against side lobe effects on the processing of input data.
5. **The use of several systems of Walsh derivative functions with different generating functions allows increasing the number of channels in the communication system without increasing the bandwidth of the communication system, which increases the efficiency of such systems.**

Search and design of PRS of arbitrary length is a real practical task. Obtaining all the sequences of the ensemble based on Walsh function derivatives with an analysis of the influence of generating function type on their correlation properties is necessary for the optimal use of considered sequences in the communication system. This, in turn, determines the quality of the communication system, the number of system subscribers, protection against unauthorized access, and, consequently, the economic aspects of systems.

At the same time, these sequences must have certain characteristics such as low threshold of cross-correlation and a pronounced peak of autocorrelation with minimal levels of side lobes, etc. The results can be used in the development of broadband communication systems and information transmission systems with protection from unauthorized access.

References

- [1] Варакин Л.Е. Системы связи с шумоподобными сигналами. - Москва: Радио и связь, 1985 – 384 с.
- [2] Мазурков М.И. Системы широкополосной радиосвязи. – Одесса: Наука и техника, 2009.—344 с.
- [3] Solomon, W. Golomb and Guang, Gong. Signal Design for Good Correlation, Cambridge, Cambridge University Press, 2005. - 458 p.
- [4] Feher K. Wireless Digital Communications: Modulation and Spread Spectrum Applications, 1995. – 524 p.
- [5] Gantmakher V.E., Bystrov N.E., Chebotarev D.V. Noise-like signals. Analysis, synthesis and processing — St. Petersburg: Science and Technology, 2005.
- [6] Ипатов В. Широкополосные системы и кодовое разделение сигналов. Принципы и приложения. М.: Техносфера. 2007. - 488 с.
- [7] Бессарабова А.А., Венедиктов М.Д., Ледовских В.И. Разделение каналов по форме в широкополосных системах

передачи информации.. Воронеж: Воронеж. гос. техн. ун-т, 2007.- 144 с.

[8] Никитин Г.И. Применение функций Уолша в сотовых системах связи с кодовым разделением каналов – СПб: СПбГУАП, 2003.- 86 р.

[9] Беспалов М.С., Складенко В.А. Функции Уолша и их приложения. – Владимир: Изд-во Владим. гос. ун-та, 2012. - 35 р.

[10] Sorochin, Gherman, Şestacova, Tatiana Comparative analysis of composite Barker codes and composite Walsh functions. - Innsbruck, Austria//Österreichisches Multiscience Journal №38/2021, p. 72-78.

[11] Senin A.I. Correlation properties of sequences based on M-sequences and Walsh sequences. - М.: Vestnik of Bauman Moscow State Technical University. Instrumentation Series № 5, 2014, pp.88-97.

[12] Khachatryan L. G. Methods of construction of de Bruijn sequences // Discrete Mathematics, 1992. Т. 3. № 4, с. 62-78.

Enhancing OCR Accuracy for ID-1 Documents with Security Features through Machine Learning-driven Image Optimization

Alba Haveriku¹, Besjana Muraku¹, Brunela Karamani¹, Hakik Paci^{1,*}
Polytechnic University of Tirana, Albania¹

alba.haveriku@fti.edu.al, besjana.muraku@fti.edu.al, bkaramani@fti.edu.al, hpaci@fti.edu.al

Abstract: OCR technology is widely used in various applications, including document digitization, data extraction, and document management systems. The OCR technology has seen significant advancements in recent years, especially with the integration of machine learning and artificial intelligence techniques. These advancements have led to substantial improvements in accuracy, particularly for standard fonts and clear document images. However, challenges still exist, especially when dealing with low-quality images, noisy images, handwritten text, or documents with unusual fonts. Some documents like ID cards, driving licenses, etc. use some security features like deliberate errors, OVI (Optical Variable Ink), Rainbow print, Guilloche pattern, fine line, and microprint to protect the documents from being counterfeited. These security elements also generate noise on the image to perform the OCR. In this paper, we present a way to enhance the OCR accuracy for ID-1 documents with security features through machine learning-driven image optimization. Albanian driving license images on the personalization process are used as a dataset to train the model. During the training process, the model of the ID-1 card is presented with a dataset containing input features (such as images, texts, or numerical data) along with corresponding labels or outcomes. After training, the model and the implemented algorithm to optimize the image for the OCR process are implemented in real-life application.

Keywords: OCR, ID-1 CARD, LASER ENGRAVING, MACHINE LEARNING, IMAGE OPTIMIZATION

1. Introduction

The ID-1 cards are common in our daily lives, like our ID cards, bank cards, driver's licenses, and membership cards. Most people have a bunch of these cards for different purposes. Let's break down the main ways we use them.

Identification: One of the main uses of an ID-1 card is to confirm who the holder is and make sure all the info is legit. Usually, these IDs come from official organizations like government agencies, etc. They are products with security features to make sure they're hard to counterfeit. Example of those cards, national IDs, passports, and driver's licenses. They usually have different fields like, name, surname birth date, a photo, and the signature of the holder, making it easy to match the holder's data [1,2, 3]. Those types of cards are engraved with laser tech.

Access: Organizations and companies give ID-1 cards to their employees to control who can go where, track work hours, and offer discounts on stuff like vending machines. In smaller companies where everyone knows each other, a photo might not be necessary [1,3]. These cards often have machine-readable features, so employees can swipe them to access restricted areas, access equipment for different purposes like printers, or get special deals at a mall for discounts without needing a photo.

Payment: Using cash is becoming less common, with credit and debit cards taking over. These cards have security features like magnetic strips (on older cards), contact chips, or contactless chips. They can be used at different stores and for online transactions. They might not have many visual security features because the security is built into the chip and the banking systems.

Tracking Services: ID cards can help businesses organize and access customer info to improve services. For example, in a medical business, the customer database might have insurance details, medications, health history, and current treatments. When a client comes in, their ID can be used to quickly access this info and answer questions about their history, payments, and treatment plans.

ID cards can have various security features to make it harder to counterfeit. Some security features are added when personalizing or printing the cards, while others are included during the production phase. When considering security, it's important to find ways to streamline processes with barcodes, magnetic strips, contact, and contactless chips [3].

The security features during production are divided into four levels based on how difficult they are to create and verify.

Level 1: These are security features that the public can easily check without special tools, like guilloche, tactile effects, and CLI/MLI features.

Level 2: These security features can be checked with simple tools such as magnifiers and UV lamps, like microprinting, UV color, and thermochromic ink.

Level 3: These security features require sophisticated laboratory equipment and are meant for qualified forensic laboratories, such as digital watermarking.

Level 4: These are highly confidential security measures known only to a small group of experts, often only the manufacturer.

To combat counterfeiting, various security features can be used in the production and personalization of ID cards [1]. Some key security features include:

Photo: Photos are a basic security feature on ID cards as they provide a quick way to verify the holder's identity. While they reduce the likelihood of fraud, they are not foolproof as photos can be replicated using optical scanners or copiers. Combining photos with a database that verifies the cardholder's photo can enhance security, especially when using technologies like barcodes, QR codes, or RFID with HID prox iclass or Mifare. A simple swipe or tap of these technologies can instantly bring up the user's record with a matching photo for comparison. This security feature can be applied to the card through color printing or more securely through laser engraving, which engraves on the second or third layer of the card.

Barcodes: Barcodes are widely used to read data on simple and low-cost devices. Information can be encoded onto an ID card during personalization or printing. One-dimensional barcodes are used for short codes, while two-dimensional barcodes like QR codes store more data and can provide security features through data encryption.

Magnetic Stripe: Magstripe cards contain a band of magnetic material on the back, storing readable and writeable information. Most magstripe badges today use high coercivity (hico), but some applications still use low coercivity (lo-co) cards. However, mag stripes are less commonly used now due to their susceptibility to copying, making them a target for fraud.

Smart Card/Proximity: In some cases, simply using a card for identification or reading barcode information isn't enough. Smart cards or proximity cards are used to store more digital data during card usage, such as a tachograph card storing a driver's identification and driving data. This data is transmitted securely using a card reader/writer via contact or contactless methods. These

security features are often used for secure door access control and time and attendance applications. Proximity cards store unique data on each card and transmit information to the card reader via an antenna coil. This process can use low-frequency 125Khz technology like HID prox or Mifare, as well as more secure high-frequency 13.56Mhz technology relying on smart chips for encrypted data on HID iclass Seos and Mifare DesFire EV2 cards.

UV Ink: This security feature is added to cards during production or sometimes during personalization using printing techniques. It involves using special ink that's only visible under UV light, which helps to lower the risk of counterfeiting [2]. It's a cost-effective technology, and the feature isn't visible to the naked eye unless the card is held under a UV blacklight. You can typically find this feature on national ID cards, driving licenses, and other similar cards.

Holograms: This visual security feature is commonly used today, especially on ID cards that hold value for the cardholder. Holograms are extremely difficult to copy due to their unique visual effects [3]. There are only a few companies that produce personalized holograms, and they can't be replicated using computer scanning/alteration or standard printing color equipment, making them highly secure. Personalized holograms, whether as a surface overlay or embedded into the card, offer the highest level of security for any application requiring enhanced security.

Microtext: This security feature is one of the toughest to reproduce. Micro text printing is widely used in passports, ID cards, and currencies. It can be combined with intentional errors to confuse potential fraudsters [2, 3]. To the naked eye, micro text looks like a regular thin line on an ID card, but when examined with a microscope or magnifying glass, it reveals repeating text such as "authentic," "genuine," or "valid." This security feature can't be printed using a regular printer; it requires industrial equipment. Therefore, it must be implemented during card production, limiting the option for personalization.

Laser Engraving: Laser engraving is a highly secure method of card personalization, usually in monochrome, that etches features into the second or third layer of the card [2, 3]. This provides tamper-proof and highly durable personalization, making forgery and manipulation nearly impossible. Any attempt to alter engraved information will visibly damage the card. Different laser powers can be used to create a tactile feature during personalization.

In many cases, it is necessary that the data written on the card should be read with the purpose of using them as digital data. A simple way to enable data reading is the use of 2-dimensional barcodes, which offer the possibility to store a lot of data. However, the barcode may not represent the real data on the card, which may not have been written in full or may have been modified later. Another form that is very useful and advanced is the OCR technology, but this technology does not offer 100% accuracy in reading the data.

The use of security elements such as deliberate errors, OVI (Optical Variable Ink), Rainbow print, Guilloche pattern, fine line, and microprint, thermochromic inks will affect the accuracy of the result obtained by OCR for reading the data.

In this article, we aim to optimize the accuracy of OCR for reading cards with a security element using machine-learning.

1.1 ID cards on transportation

Transportation is a huge field where ID-1 cards are super common. People use these cards for things like driving licenses, tachograph cards, vehicle registration cards, Driver Certificate of Professional Competence (CPC), ADR cards, and more. In Europe, almost every country uses an ID-1 card for driving licenses. If you're authorized to drive under TachoNet, you got to have a smart ID-1 card with a chip and special software for those fancy smartcards [1].

A digital tachograph is a fancy electronic device that's installed in road vehicles. It automatically (or semi-automatically) shows, keeps track of, and records the movements of the vehicle and the working hours of professional drivers. It's all about keeping an electronic record of what the driver's up to. A tachograph card is like a little chip card that's used in the digital tachograph. It verifies the owner of the card and records all the important data. The chip in these cards is super important and must be secure to make sure nobody messes with the data. That's why these cards are expensive.

So, a driving license card or a CPC card is basically an official document that proves you're allowed to drive different kinds of vehicles like motorcycles, cars, trucks, or buses on public roads. In Europe, pretty much every country uses these ID-1 cards for driving licenses. Now, these cards must meet certain security requirements outlined in EU directives. But you won't find smart cards being used for driving licenses or CPC cards in most European countries because they don't really share data digitally. Instead, these cards have some visual security features like micro text, guilloche, UV printing, and holograms. You know, stuff that makes it hard to counterfeit them. As for the cost, it varies depending on how many security features they have, but they're usually pretty affordable.

Almost all the transportation cards are laser engraved to make them stand out from the printed ones. The laser engraving is done in the middle layers of the card, while the printing happens on the first and last layers, you know, the outside of the card. It's a cool combination that ensures these cards are unique and easy to identify.

2. The OCR process and the problems

The OCR technique enables us to read data from different images by converting them into characters and thus digital information will be generated. It is a very studied and completed technology with the aim of increasing the performance and accuracy of reading data from images. Its foundations begin in the 1970s and it is constantly in decline due to the extensive use of this technology [4,5]. This technology is used in many different fields such as banking, finance, airports, transportation, handwriting reading, etc.

OCR technology has been further expanded to include new technologies such as:

- Optical Word Recognition, this technology is used to identify words for different languages when separated by spaces.
- Intelligent character recognition (ICR), aims to recognize handwritten cards, including the use of machine learning.
- Intelligent word recognition (IWR), aims to recognize handwritten words, including the use of machine learning.

The implementation of OCR has several stages to make it possible to read the data.

Pre-processing, in this phase the images are processed to increase the accuracy of reading the data [6]. This process includes:

- De-skewing, since the images generated by the scan are not always in the right position, it may be necessary to rotate them by several degrees to make the rows/lines as readable as possible.
- Binarization, this is a very important process in OCR technology because most engines of this technology work on black and white images. It is very important that the boundary between a black and white point is well defined.
- Cleanup from non-glyph elements.
- Detection of lines and words.
- Normalization of the ratio of the dimensions of the figure and its size.
- Segmentation and isolation of crankcases.

Text recognition is the process to identify and convert the characters.

To determine the list of possible characters for identification with OCR, there are two main algorithms.

- Pattern recognition, mainly used for typewritten and handwritten text that can be easily read.
- Feature extraction, mainly used for handwritten text.

Software such as Tesseract use a two-pass approach to character recognition. The second pass is known as adaptive recognition and uses letter shapes recognized with high confidence in the first pass to better recognize the remaining letters in the second pass. This is useful for unusual fonts or low-quality scans where the font is distorted.

The use of neural networks trained for character recognition enables reading not only individual characters but the entire line [7].

Post-Processing, the accuracy of OCR can be improved if the result is limited by a lexicon or a limited number of characters. This is used in cases where for a certain area we know which characters or words are allowed to be used.

Knowing the grammar of the language being scanned can also help determine whether a word is likely to be a verb or a noun, for example, allowing for greater accuracy.

The combination of security elements such as Guilloche, fine lines, microtext, OVI, etc. They will affect the accuracy of the OCR process since these security elements are visual elements.

This was also noticed in the tests done on the personalization of ID-1 cards with a security element compared to the tests done on white ID-1 cards.

To increase the accuracy of OCR in cards where we have a visual security element and personalized laser data, we propose to build an algorithm using Convolutional Neural Network (CNN).

3. Optimized Algorithm using CNN to increase the accuracy of OCR

A convolutional neural network, or CNN for short, is a cool type of artificial neural network that's mainly used for recognizing and processing images. It's great at spotting patterns in images. However, to make CNN work its magic, it needs loads of labeled data points for training. And when it comes to training, CNNs need some serious processing power, like a GPU or an NPU, to deliver results fast enough to be useful [7].

Convolutional Neural Networks (CNN) are like the superheroes of the tech world. They're used in all sorts of cool applications and can do some seriously impressive stuff [8,9]. One of their early triumphs was recognizing handwritten digits - you know, like those old-school math problems we used to solve with pen and paper. Since then, CNNs have come a long way, with new layers and fancy computer vision techniques constantly being added to make them even better. One of the places you'll often find CNNs flexing their muscles is in the ImageNet Challenge. It's like a big showdown where they go head-to-head with different datasets of sketches. They're always trying out different combinations to see which one comes out on top. So yeah, CNNs are awesome. They're constantly evolving and finding new ways to blow our minds in the world of computer vision.

To train the CNN algorithm, which will identify when a character is engraved with a laser or is part of the security element of the card, we will use a dataset which was obtained with real data of over 200,000 images of driving licenses.

The Albanian driving license contains visual security elements, and their distribution reduces the accuracy of the OCR process for reading the data on the card. The data in the datasets are grayscale images on both sides of the map, and we also have data that were measured by laser. The accuracy of the cards is considered 100%

because they have been verified manually during the quality control process.

The laser engraving process on poly-carbonate cards produces the tactile effects because the laser does not write on surface layers, but it writes on other layers. This effect can be detected with hand easily, but it is hard with 2 dimensional scanners. Anyway, this is an element that can be detected by training CNN algorithm.

This algorithm is implemented using python language and the main library we used are as follow:

- TFLearn – Deep learning library used to create layers of CNN algorithm.
- Numpy – To process the image matrices.
- Open-cv – To process the image like converting them to grayscale etc.
- Tesseract is an open-source OCR engine maintained by Google.

Tesseract is famous for being really good at reading text from images. It can handle all sorts of image types, like scanned documents and photos, and it's super accurate. Plus, it supports over 100 languages! If you want even better results, you can customize things like how the pages are broken up and the language models used. And the cool thing is, Tesseract is all about community contributions. It's also really easy to use with Python, so developers can easily add OCR capabilities to their apps.

In the following figure we have an image of an Albanian driving license personalized with laser engraving technology.

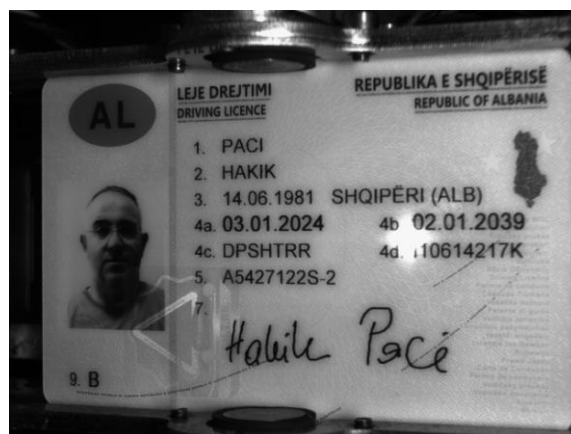


Fig. 1 Personalized driving license

Extracting personalized data with OCR, from this image using tesseract library do not provide a high accuracy because security visual elements interfaces with personalized letters.

In Figure 2, we generated using our algorithm an image that is cleaner for better OCR accuracy. This algorithm is trained to detect the contours of laser engraved elements.

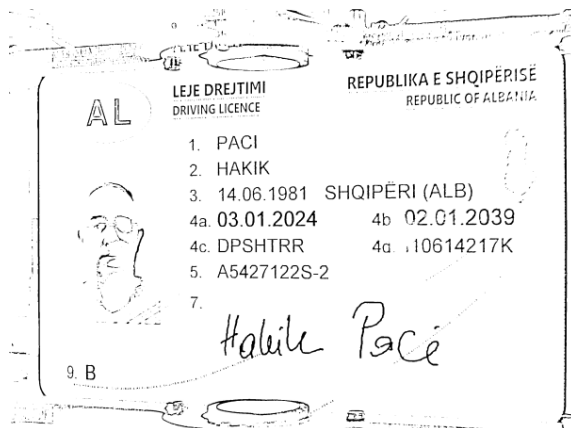


Fig. 2 OCR Optimized image of the driving license

After the implementation and training of the algorithm to real personalization system there have been personalized more than 250,000 cards and the automatic quality control process that uses the optimized images has a better accuracy with more than 99.7%.

Usually, the OCR algorithm fails for characters that are not in Albanian alphabet because the number of driving licenses for those type is very limited, and it needs more data for training process.

4. Conclusions

The OCR technology is very useful and well developed with very high accuracy and helps digitalization process in many different fields.

Sometimes the images quality is very good and can be archived very good quality on the OCR process but sometimes is necessary to change and optimize the image quality to ensure a better result. Usually, some basic images effects like rotating the angle of image with some few degrees, changing the contrasts, brightness, gamma, etc. can provide a higher accuracy rate in OCR process.

The images modification is not always the best way to archive better results because the image source is very good and does not need to be modified with basic effects.

CNN is a very powerful way to solve and to optimize results in different IT solutions. It needs to be well trained and can get better and better with the increase in the number of images used to train CNN.

In this paper, we present a way to enhance the OCR accuracy for ID-1 documents with security features through machine learning-driven image optimization. Albanian driving license images on the personalization process are used as a dataset to train the model. During the training process, the model of the ID-1 card is presented with a dataset containing input features (such as images, texts, or numerical data) along with corresponding labels or outcomes. After training, the model and the implemented algorithm to optimize the image for the OCR process are implemented in real-life application (Albanian driving license personalization quality control).

The results archived from the automatic quality control process are very promising and this solution can also be used on other cases to get higher results from the OCR even when the different graphics affect the images for the OCR process.

5. References

1. H. PACI, Digital signature implementation on ID-1 cards as a personalization security feature, SECURITY & FUTURE, ISSN: 2535-0668, YEAR VI, ISSUE 1 / 2022, pp. 32-34.
2. <https://www.thalesgroup.com/en/markets/digital-identity-and-security/government/security-features/levels-identity-security>
3. <https://www.advantidge.com/id-card-security-features/>
4. Ray Smith Daria Antonova Dar-Shyang Lee, "Adapting the Tesseract Open-Source OCR Engine for Multilingual OCR", The International Workshop on Multilingual OCR (2009), Barcelona, Spain, 2009, Article No.: 1, Pages 1–8. DOI: 10.1145/1577802.1577804.
5. Minghao Li, Tengchao Lv, Jingye Chen, Lei Cui, Yijuan Lu, Dinei Florencio, Cha Zhang, Zhoujun Li, Furu Wei, "TrOCR: Transformer-based Optical Character Recognition with Pre-trained Models", The Thirty-Seventh AAAI Conference on Artificial Intelligence, Washington DC, USA, 2023, pp. 13094-13112, DOI: 10.48550/arXiv.2109.10282
6. Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. "Deep Residual Learning for Image Recognition.", 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp. 770-778, 2016, DOI: 10.1109/CVPR.2016.90
7. Kartik Dutta, Praveen Krishnan, Minesh Mathew and. Jawahar C. V, "Improving CNN-RNN Hybrid Networks for Handwriting Recognition," 16th International Conference on Frontiers in Handwriting Recognition (ICFHR), Niagara Falls, NY, USA, 2018, pp. 80-85, DOI:10.1109/ICFHR-2018.2018.00023.
8. Karen Simonyan, Andrew Zisserman "Very Deep Convolutional Networks for Large-Scale Image Recognition", 3rd International Conference on Learning Representations, {ICLR} 2015, San Diego, CA, USA, 2015, abs/1409.1556
9. Tao Wang, David J. Wu, Adam Coates, Andrew Y. Ng. "End-to-End Text Recognition with Convolutional Neural Networks", 21st International Conference on Pattern Recognition (ICPR2012), Tsukuba, Japan, 2012, pp. 3304-3308
10. Hakik Paci, Dorian Minarolli, Evis Trandafili, Stela Paturri, "Albanian Handwritten Text Recognition using Synthetic Datasets and Pre-Trained Models," WSEAS Transactions on Information Science and Applications, vol. 21, pp. 264-271, 2024, DOI:10.37394/23209.2024.21.25
11. Hoo-Chang Shin, Holger R. Roth, Mingchen Gao, Le Lu, Ziyue Xu, Isabella Nogues, Jianhua Yao, Daniel Mollura, and Ronald M. Summers, "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning", IEEE Transactions on Medical Imaging, vol. 35, pp. 1285-1298, 2016, DOI:10.1109/TMI.2016.2528162
12. In-Jung Kim, and Xiaohui Xie, "Handwritten Hangul recognition using deep convolutional neural networks", International Journal on Document Analysis and Recognition (IJDAR), vol.18, pp. 1-3, 2015, DOI:10.1007/s10032-014-0229-4.
13. Ali Asghar, Leghari Mehwish, Hakro Dil, Awan Shafique, Jalbani Dr, Pakistan Nawabshah, "A Novel Approach for Online Sindhi Handwritten Word Recognition using Neural Network". Sindh University Research Journal SURJ (Science Series), Vol. 48(1), pp. 213-216, 2016.
14. Yudong Liang, Jinjun Wang, Sanping Zhou, Yihong Gong, and Namming Zheng, "Incorporating image priors with deep convolutional neural networks for image super resolution", Neurocomputing, vol. 194, pp. 340-347, 2016, DOI: 10.1016/j.neucom.2016.02.046

Using artificial neural networks to model climate data to adapt transport infrastructure to climate change

Stoyanka Ivanova

University of Architecture, Civil Engineering and Geodesy (UACEG) – Sofia, Bulgaria

siva_fce@uacg.bg

Abstract: This paper examines an innovative approach for modeling the influence of climate parameters on transport infrastructure using artificial neural networks. Through them, detailed climatic data are generated by geographical positions and monthly and annual maps are created for Bulgaria's territory using the following parameters: surface temperature, diffuse fraction, horizontal solar irradiation, and average albedo of the terrain. Average ground temperature and monthly solar irradiation are essential for maintenance planning and developing strategies to adapt to extreme weather conditions, such as heat waves or frost, which can affect the condition and performance of the road surface. Average monthly temperatures can be used to design effective systems to prevent icing of road surfaces and improve drainage systems. By demonstrating the capabilities of accurate modeling and analysis, this paper highlights the importance of applying artificial neural networks in planning and improving the resilience of transport infrastructure against climate change.

Keywords: CLIMATE CHANGE, ARTIFICIAL NEURAL NETWORKS, TRANSPORT INFRASTRUCTURE, SOLAR RADIATION

1. Introduction

Climate change represents one of the most serious challenges of modern times, affecting various aspects of life and the economy. Transport infrastructure is particularly vulnerable to extreme weather conditions such as high temperatures, heat waves, heavy rainfall, and snowfall. These climatic factors can damage road surfaces, reducing their durability and increasing maintenance and repair costs. In this context, adapting transport infrastructure to climate change is essential to ensure its sustainability and functionality [1].

The current development aims to demonstrate the possibility of using artificial neural networks (ANN) to model climate parameters that are important for transport infrastructure. An essential task is to investigate how ANNs can analyze, store, and feed on demand by finely interpolating various climatic factors essential for road infrastructure design and maintenance. Using data from multiple sources, such as PVGIS and NASA POWER, and integrating this data into ANNs, aims to create reliable and accurate models to support informed decision-making for adapting transport infrastructure to climate change.

2. Methodology

2.1. Description of data used and characteristics

Artificial Neural Networks (ANN) represent a powerful tool for modeling climate parameters [2]. Here are the climate factors selected for this stage of development and how each may affect the climate assessment of transport infrastructure:

Ground temperature (at a height of 2 m above the ground):

- Crucial for determining the thermal load on road surfaces, which can cause them to expand or contract.
- It affects the choice of building materials that must withstand temperature fluctuations.
- An important factor in planning winter road maintenance, such as de-icing and snow removal.

Diffuse fraction of solar irradiation:

- An indirect measure of cloudiness appears.
- Its value is related to visibility and road safety, especially in areas with high cloud cover.
- Determines the need for additional lighting on road sections.

Global horizontal solar irradiation:

- Crucial for determining the thermal load from solar radiation on road surfaces, which can lead to their expansion or contraction.
- It affects the choice of road surface materials resistant to UV radiation.

Global inclined solar irradiation (on inclined surfaces):

- It is used to assess the potential for using solar panels in road infrastructure, such as snow protection equipment.

- This is important for the design of road facilities, such as bridges and tunnels, where shadows can affect safety and visibility.

Albedo (reflectivity) of the terrain:

- It affects the reflection of sunlight from road surfaces, which can affect the temperature and resistance of materials.
- The parameter is essential in calculating reflected solar irradiation, and it is a component in calculating global inclined solar irradiation.
- Albedo reflects success in designing adequate road surfaces to minimize the heat island effect in urban areas.

Each of these factors can be analyzed and modeled using ANNs to provide accurate and reliable data to support the design and maintenance of road infrastructure in different climates.

More broadly, this includes calculating the probability of road icing, snow accumulation, road conditions under different weather scenarios, and the potential impact of climate change on road safety and durability.

2.2. Sources of climatic data for the territory of Bulgaria

Finding climatic data of Bulgarian origin is a significant and complex challenge. Therefore, data on foreign origin was used. Three primary data sources were used to train and test the ANN (PVGIS [3], NASA POWER [4], and NEO – NASA Earth Observations [5]). Choosing GIS data with climate data for training Artificial Neural Networks (ANNs) instead of traditional data from weather stations has several key advantages:

Broad geographic scope and coverage: PVGIS and NASA POWER data offer global coverage with high spatial resolution. This is particularly important for ANN training as it allows the analysis of climate patterns and parameters in different geographic regions, including hard-to-reach and little-explored areas.

Data content and diversity: PVGIS, NASA POWER, and NEO – NASA Earth Observations provide an extensive set of weather and climate data, including solar radiation, temperature changes, humidity, wind conditions, and more. This variety of information is essential for developing complex and accurate models through ANNs.

Constant update and availability: Data from the satellite systems are updated regularly, which ensures continuous access to up-to-date information. This is important for maintaining the relevance and accuracy of ANN models that can respond to ongoing climate change.

Standardization and compatibility: Data from these sources are standardized and easily compatible with different analytical tools and software platforms, facilitating their integration and processing in the context of ANN.

These advantages make PVGIS and NASA POWER highly suitable for ANNs, providing the rich and diverse data set needed to create reliable and accurate predictive models.

2.2.1. PVGIS (Photovoltaic Geographical Information System) [3]

This system provides detailed data on solar irradiance and meteorological parameters that are critical for modeling solar energy and related climate aspects. PVGIS information can be used to analyze the influence of solar radiation on road conditions.

As stated on the PVGIS website, the JRC-PVGIS database is built based on satellite measurements between 1998 and 2016. The data is provided by the Climate Monitoring Satellite Application Facility (CM-SAF) and has a time step of 1 hour and a spatial resolution of approximately 2.5 km (1'30"). The PVGIS-CMSAF database covers the same territory as the Meteosat satellite, which includes Europe, Africa, and parts of South America.

Later, the CM-SAF team created a new solar database called PVGIS-SARAH using a different algorithm successfully implemented in Asia in the previous version of PVGIS. This database also covers the PVGIS-CMSAF territory. In 2020, an updated version called PVGIS-SARAH2 was published.

PVGIS's spatial resolution is about 2.5 km. These solar data are of a new generation and significantly better in terms of spatial and temporal resolution and accuracy.

Data is available online on request for a specific geographic location defined by latitude and longitude. To cover the territory of Bulgaria with high resolution, tens of thousands of requests for providing data on the monthly values of diffuse and global solar irradiation must be made manually and processed. The data is available via a Non-Interactive Service API, which allows data to be downloaded at the desired resolution with a program (e.g., VBA in the MS Excel environment) that makes API requests for locations specified with their geographic coordinates. The denser the network and the more requests, the more detailed information can be retrieved.

2.2.2. NASA POWER (Prediction of Worldwide Energy Resources) [4]

Solar data for the territory of Bulgaria can be found in different formats and on different NASA websites, being offered with different angular resolutions and organization. These data cover solar horizontal irradiation for the period 1981 to 2021, as well as its diffuse component since 2001. The data can be viewed using the visualization software POWER Data Access Viewer, available online.

This database provides extensive information on climate parameters collected by various NASA satellite missions. Data from NASA Power includes temperature, humidity, wind conditions, and other important weather indicators critical to understanding the climate factors affecting road infrastructure.

Using this data will allow ANN to train and test models that can predict climate conditions with high accuracy and reliability. This, in turn, will improve the ability to plan and react to different weather conditions, thereby increasing the safety and efficiency of the road infrastructure.

Upon request, NASA Power generates text files (in ASCII, CSV, and other formats) with the necessary data, which can be further processed with other software. Data can be organized by day (in the menu Temporal Average: Daily) or month (in the menu Temporal Average: Interannual).

According to the information on the site, the solar data is from MERRA-2 (Modern-Era Retrospective Analysis for Research and Applications, Version 2), a relatively low-resolution NASA climate database: $0.5^\circ \times 0.625^\circ$. On the same site, other valid climate data with a resolution of 0.5° are available, such as information on winds (direction and strength at different heights), humidity, temperature (at ground level and 2 m above it), atmospheric pressure, etc. Through interpolation, the system can provide this information for a specific location.

In addition to the data for past years, monthly climate data summarized (averaged) over a longer period is also available through the POWER Data Access Viewer (in the Temporal Menu Average: Climatology). Among them are data on heating and cooling degrees, terrain albedo, solar irradiation, temperatures, and others.

2.2.3. NEO - NASA Earth Observations [5]

Data on the NEO—NASA Earth site Observations are organized into five main areas: Atmosphere, Energy, Land, Life, and Ocean. Each of these areas includes between 10 and 20 datasets. Monthly solar values, arranged in maps, are available in the Energy area under the name "Solar Insolation." Unfortunately, this data includes only the global cumulative solar energy and lacks information on its diffuse component.

Various useful and interesting data with different resolutions can be downloaded from this site:

- 1.0° (matrix of size 360×180 elements);
- 0.5° (matrix size 720×360);
- 0.25° (matrix size 1440×720);
- 0.1° (matrix size 3600×1800).

This site offers climate data and information on the altitude of the earth's surface in meters, with a maximum resolution of 0.1° . For Bulgaria's latitude, this corresponds approximately to a resolution of 8.16 km in the east-west direction and 11.12 km in the north-south direction. These data were used for the relief in Bulgaria.

2.3. Downloaded data

The neural networks are trained with data downloaded mostly from PVGIS [3] and the NASA POWER website [4].

Data on diffuse fraction, global horizontal irradiation, and temperature by month from 2005 to 2020, with a resolution of 0.1° , were downloaded from the PVGIS site via API requests. For this task, data was downloaded in CSV format from the SARAH2 database for an area slightly larger than that of Bulgaria, covering a latitude of 41° to 44.5° and a longitude of 22° to 29° (see Fig. 1) for the period from 2005 to 2020.

In the east-west direction, the area is divided into 70 parts, and in the north-south direction—into 35, for a total of 2450 regions with dimensions of $0.1^\circ \times 0.1^\circ$.

The downloaded data are averaged and saved in a format suitable for ANN training. This format includes information on the month, latitude and longitude, a diffuse fraction (D/G), mean daily solar irradiation (G), and mean diurnal temperature for the respective month. Locations inside the Black Sea or the Mediterranean Sea, identified by the zero temperature values provided by PVGIS, were removed from the file. As a result, 26,412 records remain in the database.

POWER Data Access Viewer site allows climate data to be downloaded from a specified point or region (POWER Regional Data Access) from 1981 to 2020. Albedo data starts in 1984 and has a resolution of 0.5° , based on the MERRA2 database, which has a resolution of $0.5^\circ \times 0.625^\circ$. In general, this is a relatively low resolution that will not provide a high-quality end result.

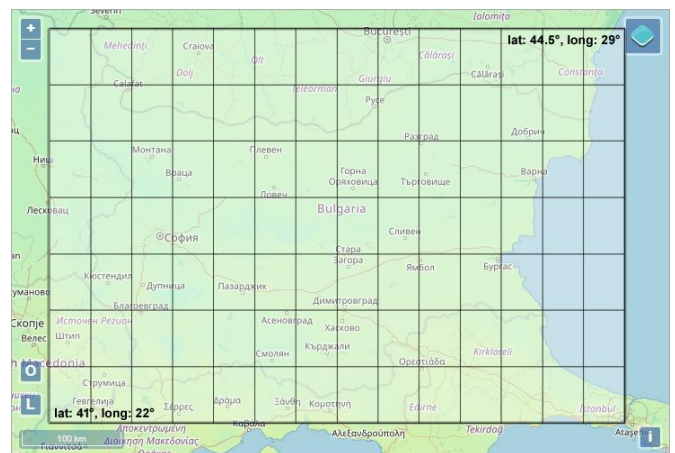


Fig. 1. Dimensions of the analyzed territory of Bulgaria and adjacent areas. Dividing the territory into parts.

Data has been downloaded in ASCII or CSV format for an area with latitude 41° to 44.5° and longitude 22° to 29° from 2005 to 2020. The downloaded file has a header of 10 lines and 1568 lines of albedo data by month (16 years × 98 regions on the territory of Bulgaria). Some initial analyses were performed on these data.

Elevation data downloaded from NEO – NASA Earth Observations with the same resolution.

When working with ANNs, it is desirable to work with territories/values with a slightly more extensive range than necessary to increase accuracy for the desired range.

3. Modeling with artificial neural networks

3.1. Trained neural networks

The artificial ones used neural networks were trained as part of the developments in the monograph “Artificial Neural Networks for Energy Efficiency” [2]. Four of them (see Fig. 2) are suitable for the present development:

- ANN with 5 input neurons, 30 neurons in the hidden layer, and 1 neuron in the output layer to determine surface temperatures based on a given month, geographic coordinates, elevation, and terrain slope. Training terminates after 47,000 epochs, with a correlation coefficient between training and resulting values $R^2 = 0.997$ and the test set – with $R^2 = 0.997$.
- ANN with 4 input neurons, 40 neurons in the hidden layer, and 1 neuron in the output layer to determine the diffuse fraction (ratio between monthly values of diffuse and global solar irradiation, i.e., a measure of average cloud cover) based on a given month, geographically location coordinates and altitude. Training terminates after 108,300 epochs, with a correlation coefficient between training and output values of $R^2 = 0.964$, and with the test set, with $R^2 = 0.962$.
- ANN with 5 input neurons, 20 neurons in the hidden layer, and 1 neuron in the output layer to determine the horizontal solar irradiation based on a given month, geographical location coordinates, altitude, and diffuse fraction. Training terminates after 23,000 epochs, with a correlation coefficient between training and output values of $R^2 = 0.999$, and with the test set, with $R^2 = 0.999$.
- An ANN with 4 input neurons, 40 neurons in the hidden layer, and 1 neuron in the output layer to determine monthly albedo values of the surrounding terrain based on a given month, geographic coordinates, altitude, and temperature. Training terminates after 3,788,910 epochs, with a correlation coefficient between training and output values of $R^2 = 0.979$ and the test set, with $R^2 = 0.976$.

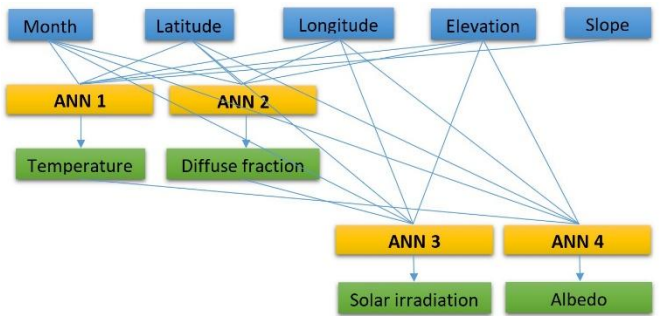


Fig. 2. Schema of the connections between inputs (in blue) and outputs (in green) in the project's integrated four artificial neural networks (in yellow).

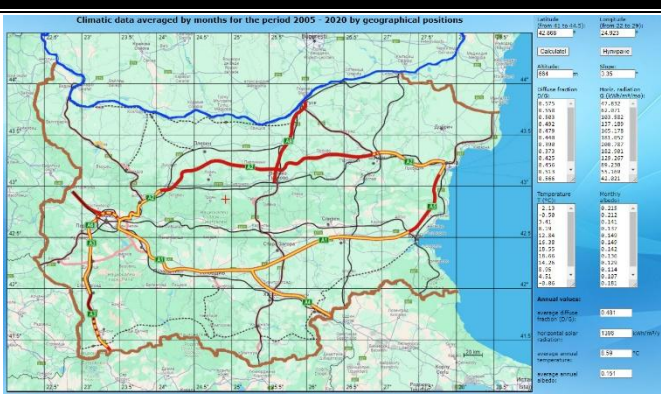


Fig. 3. Page content for ANN retrieval of monthly and annual climate data (temperature, diffuse fraction, horizontal solar irradiation, and albedo) by geographic location.

3.2. Using trained ANNs in HTML pages with Javascript code

The approach of using trained ANNs in HTML pages with JavaScript programming code included is developed in the same monograph [2]. This approach has been further developed for this project, and the web pages are available in a small website at [6]:

1) **Page for climate data by geographical positions.** Data on average monthly temperatures, diffuse fraction, horizontal solar irradiation (kWh/m²/mo), and albedo for all 12 months are displayed for a given geographical position on a map of Bulgaria. The displayed points on the map can be settlements, highways, and roads of different classes. In addition, summary values of these parameters for the year are also output, as shown in Fig. 3.

2) **Page for monthly and annual climate data maps.** Monthly maps (for a specified month) and annual maps with variable resolution (0.1°, 0.05°, 0.02°, 0.01°) are drawn for selected parameters: surface temperature, diffuse fraction, horizontal solar irradiation, and albedo.

The page displays a map of meters' elevation (see Fig. 4) and terrain slope. The monthly maps are divided into four groups for relevant climate parameters, 12 for each month. There are four annual maps, one for each parameter.

Fig. 5 demonstrates an example of page work for compiling monthly and annual climate maps using ANN (for temperature, diffuse fraction, horizontal solar irradiation, and albedo). The map shows the average diffuse fraction for the month of January (average for the period 2005 – 2020).

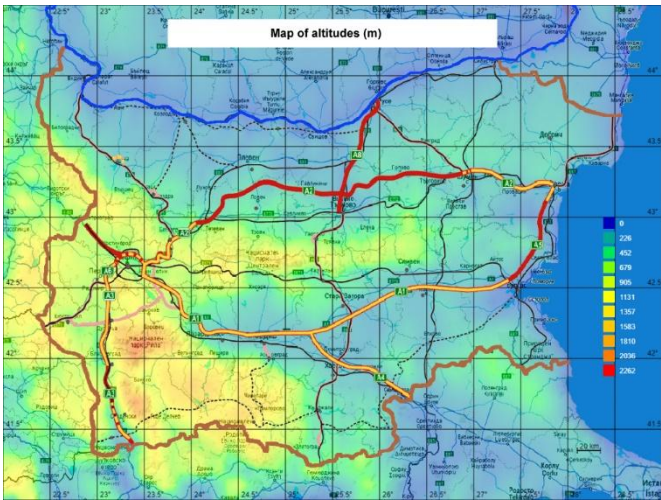


Fig. 4. Map of altitudes in meters of the considered territory of Bulgaria.

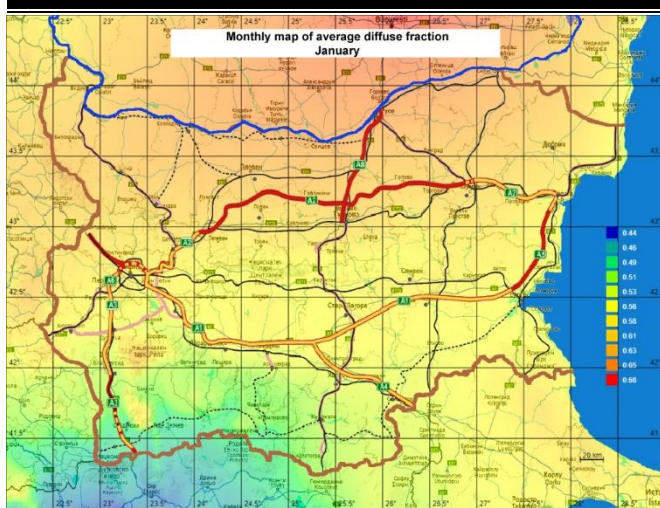


Fig. 5. An example of the work of the page for compiling monthly and annual climate maps using ANN (for temperature, diffuse fraction, horizontal solar irradiation, and albedo). The map shows the average diffuse fraction for January (average for the period 2005 – 2020).

4. Applications of the developed pages

4.1. Page for climate data by geographic position

Climate data for 2005 – 2020 by geographical positions provides valuable information that can be used to adapt transport infrastructure to climate change. Here are some possible applications of this data:

Road infrastructure design and improvement: Elevation and average monthly temperature data can help design roads that are resilient to extreme weather conditions, such as heavy rainfall, snowfall, or high temperatures.

Maintenance and rehabilitation planning: Analysis of horizontal solar irradiation and albedo can be used to plan the maintenance of road surfaces, taking into account potential wear and degradation from UV radiation and temperature changes.

Optimizing lighting: Global horizontal solar irradiation and diffuse fraction data can be used to optimize the planning and placement of street lighting, choosing the most efficient locations and angles to maximize the use of natural light and reduce energy costs.

Adaptation to climate change: Average annual temperatures and solar radiation data can be used to analyze climate change trends and adapt transport infrastructure to cope with expected changes, such as increasing the resistance of road surfaces to high temperatures or improving drainage systems to cope with increased rainfall.

Developing sustainable transport solutions: Climate information can support the development of sustainable transport solutions, such as building green corridors to reduce urban heat and improve air quality and planning cycling and pedestrian routes adapted to local climatic conditions.

Prevention of erosion and landslides: Elevation data can help analyze the risk of erosion and landslides, especially in areas with significant climate change. This is key to preventing damage to road infrastructure.

These applications illustrate how detailed information on climate conditions can be used to improve and adapt transport infrastructure to current and future climate challenges, thereby increasing its resilience and efficiency.

4.2. Page for monthly and annual climate data maps

The monthly and yearly maps of climate data for 2005 – 2020 offer a valuable tool for analysis and planning, especially in adapting transport infrastructure to climate change. Here are some possible applications of the data provided in this area:

Infrastructure planning and design: Elevation and terrain slope data can be used to plan new roads, railways, and bridges, considering the potential risks of flooding, erosion, and landslides.

This is particularly important in areas where climate change may increase the frequency and intensity of these phenomena.

Adaptation to extreme temperature conditions: Monthly and annual temperature maps provide information on changes in climate conditions, which is critical for adapting transport infrastructure to extreme temperature conditions. This may include choosing materials that are more resistant to thermal expansion and contraction and planning maintenance and repair activities.

Optimizing the use of solar energy: Global horizontal solar irradiation and albedo data can be used to optimize the placement and design of solar panels on the roofs of transport facilities, such as stations, bus stops, and parking lots, to produce clean energy.

Water resources management: Analysis of climate data can help manage water resources, especially in planning drainage systems for roads and railways. This is important to prevent flooding and ensure the safety of passengers and vehicles.

Improving infrastructure resilience: Annual climate data maps can be used to assess long-term climate trends, which is essential for designing sustainable transport infrastructure that can withstand changes in climate conditions.

Education and research: The data provided can be used in educational programs and research aimed at understanding and adapting to climate change and developing innovative solutions to reduce the environmental impact of transport infrastructure.

These applications highlight the importance of access to accurate and detailed climate data for developing strategies to adapt transport infrastructure to future climate challenges.

5. Conclusions

The research presents an innovative approach for modeling the influence of climatic parameters on the transport infrastructure through artificial neural networks. The main conclusions of the present study are the following:

- **Applicability of ANNs for modeling climate data:** ANNs are proving to be a powerful tool for analyzing and modeling climate parameters critical to transportation infrastructure. Using historical climate data and interpolation, ANNs can generate accurate and reliable forecasts for various climate conditions.
- **Improving infrastructure resilience:** Using ANNs to model climate data provides an opportunity to enhance the resilience of transport infrastructure against climate change. This includes designing anti-icing systems, optimizing drainage systems, and selecting appropriate building materials.
- **Expanding planning and maintenance capabilities:** Models built using ANNs provide detailed monthly and annual maps of climate parameters. These maps can be used to plan road maintenance and rehabilitation and adapt infrastructure to extreme weather conditions.
- **Versatility of application:** The information generated through ANN can be used in transport infrastructure and other fields such as agriculture, water resources management, and urban planning. This highlights the importance of accurate climate data for various sectors of the economy.

Future research may focus on expanding the database and improving the accuracy of the models, as well as investigating other climate parameters that may affect transportation infrastructure. In addition, integrating ANNs with different analytical tools can lead to more effective and sustainable climate change adaptation solutions.

Acknowledgments

The development was financed under the “GIS-based analyses of climate change impacts on road infrastructure to build adaptive capacity” project at the UACEG Research, Consultancy, and Design Center.

6. References

1. Republic of Bulgaria, National Strategy for Adaptation to Climate Change and Action Plan, 2019.

- <https://www.moew.government.bg/bg/klimat/mejdunarodni-pregovori-i-adaptaciya/adaptaciya/>
2. St. Ivanova. Artificial neural networks for energy efficiency, UACEG – Sofia, 2022.
 3. Photovoltaic Geographical Information System (PVGIS): https://joint-research-centre.ec.europa.eu/photovoltaic-geographical-information-system-pvgis_en
 4. NASA POWER (Prediction of Worldwide Energy Resources): <https://power.larc.nasa.gov/>
 5. NASA Earth Observations: <https://neo.gsfc.nasa.gov/>
 6. Home page: <https://ait.uacg.bg/Transport/>
 7. Page for climate data by geographical positions: <https://ait.uacg.bg/Transport/climate-location.html>
 8. Page for monthly and annual climate data maps: <https://ait.uacg.bg/Transport/climate-map.html>

Modeling solar data using artificial neural networks for solar applications in transport infrastructure

Stoyanka Ivanova ^{1,*}

University of Architecture, Civil Engineering and Geodesy (UACEG) – Sofia, Bulgaria ¹
siva_fce@uacg.bg

Abstract: This paper examines an approach using artificial neural networks for innovative modeling of solar data that is needed to realize solar applications for transport infrastructure purposes. Through this modeling, detailed solar data are generated by geographical positions, and monthly and annual maps are created for the territory of Bulgaria for horizontal solar irradiation with its diffuse and direct components and for inclined and reflected solar irradiation, according to Norio Igawa's model. Diffuse fraction and horizontal and inclined solar irradiation can be helpful in designing solar applications in road infrastructure, such as power signaling systems and street lighting. By demonstrating the capabilities of accurate modeling and analysis of solar data, this paper highlights the importance of applying artificial neural networks in planning and improving the resilience of transport infrastructure against climate change. Using solar energy in transport infrastructure reduces carbon emissions and strengthens environmental sustainability.

Keywords: SOLAR DATA MODELING, ARTIFICIAL NEURAL NETWORKS, TRANSPORT INFRASTRUCTURE, CLIMATE CHANGE ADAPTATION, SOLAR ENERGY APPLICATIONS

1. Introduction

Climate change poses a severe challenge to transport infrastructure, with increased temperatures, intense rainfall, and extreme weather leading to damage to road surfaces, erosion, and flooding [1]. These factors reduce the durability of infrastructure objects and increase maintenance and repair costs. In this context, solar energy plays a crucial role in preventing and adapting to the harmful effects of climate change. The use of solar technologies in transport infrastructure, such as solar panels for street lighting and powering signal systems, not only reduces carbon emissions but also increases transport networks' energy efficiency and reliability. Solar solutions contribute to the sustainable development and adaptability of the transport network, providing better protection and functionality in changing climate conditions.

Solar energy offers a sustainable and environmentally friendly source of energy that is essential for modern transport infrastructure. Implementing solar technologies in transport infrastructure achieves a significant reduction in carbon emissions and dependence on fossil fuels. Solar applications improve transportation networks' energy efficiency and reliability while contributing to resilience and adaptability to climate change. These technologies provide better protection against extreme weather conditions and contribute to the long-term sustainability of transport infrastructure.

This research aims to demonstrate the application of developed artificial neural networks (ANN) for modeling solar data for use in solar applications in transportation infrastructure. The study focuses on two main web pages that visualize modeling results and provide valuable data for users.

The scope of the study includes:

- Using already developed neural networks described in the monograph "Artificial Neural Networks for Energy Efficiency" [2] for modeling solar data;
- generation of detailed solar data and maps by geographical positions and generation of monthly and annual maps for the territory of Bulgaria for various types of solar irradiation;
- development of web pages that provide interactive tools for visualization and analysis of solar data;
- and there are opportunities to improve the sustainability of the transport infrastructure through the use of solar energy.

Various methods for modeling solar data vary in complexity and accuracy. Traditional methods include empirical models and statistical analyses based on historical data and meteorological observations. However, these methods may be limited by the accuracy and range of available data.

With the development of artificial intelligence and machine learning, artificial neural networks (ANNs) have established themselves as powerful tools for modeling solar data [2, 3]. ANNs can process large volumes of data and discover complex relationships between various parameters, leading to more accurate predictions and modeling of solar irradiation.

Although ANNs offer significant advantages, their successful application requires careful data preparation, model optimization, and validation of results. The present study seeks to integrate the best practices and methodologies in artificial neural networks to provide reliable and accurate models for solar applications in Bulgaria's transport infrastructure.

2. Methodology

2.1. Description of data

Artificial Neural Networks (ANN) are a powerful tool for modeling climate parameters. In this development, the following climatic factors are modeled, which are essential for the assessment of the impact of climate on transport infrastructure:

- Surface temperature (at a height of 2 m above the ground);
- Diffuse fraction of solar irradiation;
- Horizontal global solar irradiation;
- Inclined global solar irradiation (on inclined surfaces);
- Albedo (reflectivity) of the terrain.

2.2. Sources of climatic data for the territory of Bulgaria

Finding climate data of Bulgarian origin is a significant and complex challenge. Therefore, data of foreign origin from three primary sources were used to train and test the ANN: PVGIS, NASA POWER, and NEO – NASA Earth Observations.

PVGIS (Photovoltaic Geographical Information System) [4]:

It provides detailed data on solar irradiation and meteorological parameters critical for modeling solar energy and related climate aspects. The data covers the period 2005-2020 with a resolution of 0.1°.

NASA POWER (Prediction of Worldwide Energy Resources) [5]: This site offers extensive information on climate parameters collected by various NASA missions, including temperature, humidity, wind conditions, and other important meteorological indicators. The data covers the period 1981-2020.

NEO - NASA Earth Observations [6]: It organizes data into five main areas: atmosphere, energy, land, life, and ocean. It provides valuable information at various resolutions, including data on the height of the Earth's surface and global solar insolation.

2.3. Data processing and pre-processing

API requests were used to download high-resolution data from PVGIS and NASA POWER for Bulgaria's territory. The data were downloaded in ASCII or CSV format and converted to a suitable format for ANN training.

The data are normalized to improve the accuracy and efficiency of the ANN. Missing values and anomalies are handled and removed to ensure quality model training.

2.4. Selection of modeling parameters

The following modeling parameters are selected:

- **Geographical coordinates:** They include latitude and longitude, which are crucial to accurately predicting climate parameters.
- **Meteorological parameters:** They include temperature, diffuse fraction, horizontal and inclined solar irradiation, and terrain albedo.
- **Physical characteristics of the field:** They include altitude and terrain slope, which affect the distribution of solar radiation and temperature fluctuations.

The current methodology provides a comprehensive and structured approach for processing and modeling solar data through artificial neural networks. These networks are essential for planning and optimizing solar applications, and this is the first of their kind for Bulgaria.

2.5. Scope of data

The neural networks are trained with data downloaded mostly from PVGIS [4] and the NASA POWER website [5].

Data on diffuse fraction, horizontal global irradiation, and temperature by month from 2005 to 2020, with a resolution of 0.1° , were downloaded from the PVGIS site via API requests. For this task, data was downloaded in CSV format from the SARA2 database for an area slightly larger than that of Bulgaria, covering a latitude of 41° to 44.5° and a longitude of 22° to 29° (see Fig. 1) for the period from 2005 to 2020.

In the east-west direction, the area is divided into 70 parts, and in the north-south direction—into 35, for a total of 2450 regions with dimensions of $0.1^\circ \times 0.1^\circ$.

The downloaded data are averaged and saved in a format suitable for ANN training. This format includes information on the month, latitude and longitude, a diffuse fraction (D/G), mean daily solar irradiation (G), and mean diurnal temperature for the respective month. Locations inside the Black Sea or the Mediterranean Sea, identified by the zero temperature values provided by PVGIS, were removed from the file. As a result, 26,412 records remain in the database.

POWER Data Access Viewer site allows climate data to be downloaded from a specified point or region (POWER Regional Data Access) from 1981 to 2020. Albedo data starts in 1984 and has a resolution of 0.5° , based on the MERRA2 database, which has a resolution of $0.5^\circ \times 0.625^\circ$. Generally, this relatively low resolution will not provide a high-quality final result.

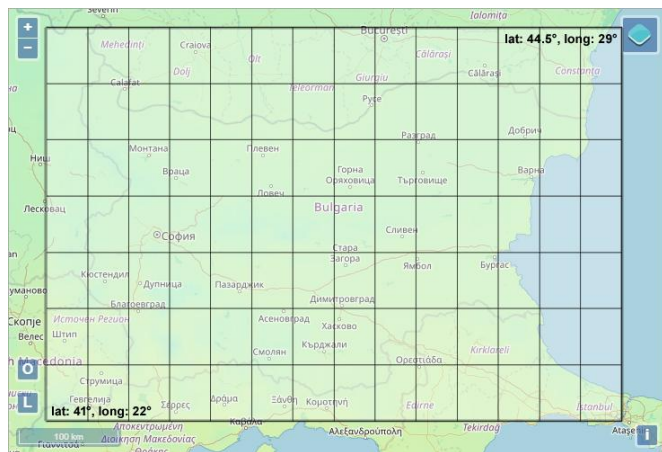


Fig. 1. Dimensions of the analyzed territory of Bulgaria and adjacent areas. Dividing the territory into parts.

Data has been downloaded in ASCII or CSV format for an area with latitude 41° to 44.5° and longitude 22° to 29° from 2005 to 2020. The downloaded file has a header of 10 lines and 1568 lines of albedo data by month (16 years $\times$ 98 regions on the territory of Bulgaria). Some initial analyses were performed on these data.

Elevation data downloaded from NEO – NASA Earth Observations with the same resolution.

When working with ANNs, it is desirable to work with territories/values with a slightly larger range than necessary to increase accuracy for the desired range.

3. Modeling with neural networks

3.1. Trained neural networks

Used and artificial neural networks were trained as part of the developments in the monograph “Artificial Neural Networks for Energy Efficiency” [2]. Five of them are adequate to the present development:

- ANN with 5 input neurons, 30 neurons in the hidden layer, and 1 neuron in the output layer to determine surface temperatures based on a given month, location, elevation, and terrain slope. Training terminates after 47,000 epochs, with a correlation coefficient between training and resulting values $R^2 = 0.997$ and the test set – with $R^2 = 0.997$.
- ANN with 4 input neurons, 40 neurons in the hidden layer, and 1 neuron in the output layer to determine the diffuse fraction (ratio between monthly values of diffuse and global solar irradiation, i.e., a measure of average cloud cover) based on a given month, geographically location coordinates and altitude. Training terminates after 108,300 epochs, with a correlation coefficient between training and output values of $R^2 = 0.964$, and with the test set, with $R^2 = 0.962$.
- ANN with 5 input neurons, 20 neurons in the hidden layer, and 1 neuron in the output layer to determine the horizontal solar irradiation based on a given month, geographical location coordinates, altitude, and diffuse fraction. Training terminates after 23,000 epochs, with a correlation coefficient between training and output values of $R^2 = 0.999$, and with the test set, with $R^2 = 0.999$.
- ANN with 4 input neurons, 40 neurons in the hidden layer, and 1 neuron in the output layer to determine monthly albedo values of the surrounding terrain based on a given month, location, altitude, and temperature. Training terminates after 3,788,910 epochs, with a correlation coefficient between training and output values of $R^2 = 0.979$ and the test set, with $R^2 = 0.976$.
- ANN with 4 input neurons, 35 neurons in the hidden layer, and 1 neuron in the output layer to determine cumulative monthly global irradiation on inclined surfaces by monthly mean day length, monthly mean sun elevation angle at noon, diffuse fraction, plane azimuth (exposure) relative to the south direction, the slope of the plane relative to the horizon. It calculates G_t / G , a ratio between monthly global inclined irradiation and monthly global horizontal irradiation. Its training terminates after 334,800 epochs, with a correlation coefficient between training and output values of $R^2 = 0.9955$ and the test set, with $R^2 = 0.9966$. The training values are created by applying the model of Norio Igawa [7] of diffuse radiance in the SOLARES software [8]. Details of the approach are given in the author's monograph [2].

3.2. Using trained ANNs in HTML pages with Javascript code

The approach of using trained ANNs in HTML pages with JavaScript programming code included is developed in the same monograph [2]. This approach has been further developed for this project, and the web pages are available in a small website [9]:

1) **Page for solar potential by geographic positions** [10]. Numerical data for horizontal, inclined, and reflected solar irradiation ($\text{kWh/m}^2/\text{mo}$) for all 12 months are displayed at a given geographical position on a map of Bulgaria. In addition, summary values of these parameters for the year are also displayed, as shown in Fig. 2.

2) **Page for monthly and annual solar data maps** [11]. Monthly maps (for a given month) and annual maps with variable resolution (0.1° , 0.05° , 0.02° , 0.01°) are drawn for selected parameters: horizontal solar irradiation, direct and diffuse horizontal solar irradiation, inclined and reflected solar irradiation (at set slope and orientation of the receiving surface). The page displays monthly

maps, divided into 12 groups for corresponding solar parameters. There are also 6 annual maps, one for each parameter.



Fig. 2. Page content for ANN extraction of monthly and annual solar data (horizontal, reflected, and inclined solar irradiation) by geographic positions.

Fig. 3 shows an example of page operation for compiling monthly and annual solar maps using ANN (horizontal, reflected, and inclined solar irradiation). The map shows global annual inclined solar irradiation at azimuth 0° and slope 33° on the receiving plane. This is the optimal tilt for PV panels inclined in a southerly direction.

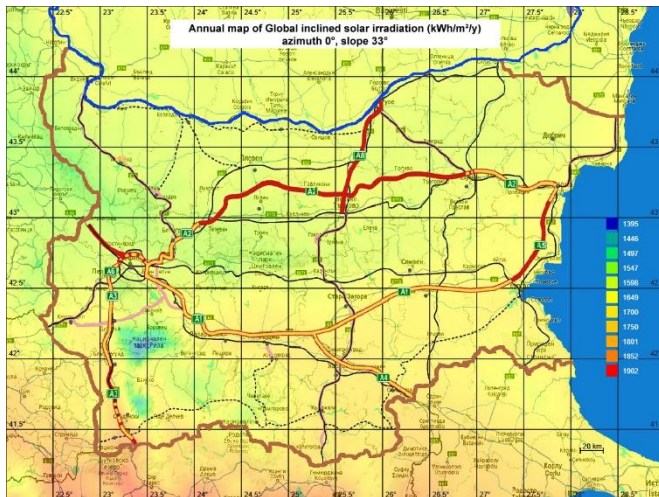


Fig. 3. An annual map of global inclined solar irradiation in kWh/m^2 with south exposure and a slope 33°, generated within page [11].

Fig. 4 shows an example of a monthly solar map of the global oblique solar irradiance in June with east exposure and slope 10°. Often, PV panels are placed in east/west orientation, with a 10-15° tilt, for a larger amount of energy generated before or after noon.

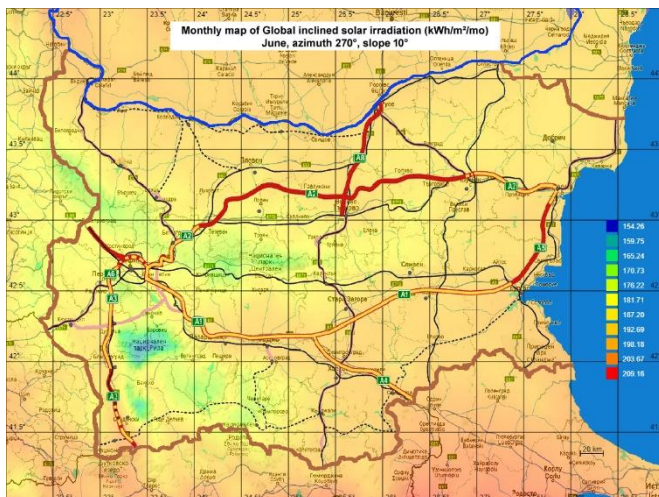


Fig. 4. A monthly map of global inclined solar irradiation in kWh/m^2 with eastern exposure and a slope 10° for June, generated within page [11].

4. Applications

4.1. Page for solar potential by geographic location

The analysis of the solar potential in Bulgaria, based on detailed data on solar irradiation by month and geographical position, can significantly impact the adaptation of the transport infrastructure to climate change. Here are some possible applications of this data in this area:

Integrating solar energy into transportation infrastructure: Data on solar potential can help integrate solar panels into the roofs of bus stops, parking lots, and other elements of transportation infrastructure. This can reduce dependence on traditional energy sources and improve energy efficiency.

Route planning for electric vehicles: Solar potential data can help plan routes and locations for solar-powered electric vehicle charging stations, providing a more sustainable transportation alternative.

Adaptation to changing climate conditions: Understanding solar potential can help adapt transport infrastructure to changing climate conditions, considering expected temperature increases and precipitation changes. This could include building more sustainable road surfaces and drainage systems.

Optimizing street and tunnel lighting: Solar data can be used to optimize street and tunnel lighting by integrating solar panels to reduce energy costs and improve the energy efficiency of lighting systems.

Development of green corridors: Data on solar potential can support the development of green corridors that include integrated solar panels for energy production and vegetation planting to improve air quality and reduce urban temperatures.

Improving the safety and reliability of transportation infrastructure: Analysis of solar potential can help design transportation infrastructure more resilient to extreme weather conditions by providing sufficient lighting and power for signaling and safety systems even during disruptions to traditional power supply.

These applications illustrate the potential of using detailed solar data to improve transport infrastructure's resilience, efficiency, and safety in response to climate change.

4.2. Page for monthly and annual solar data maps (2005-2020 averages)

The monthly and annual solar resource maps page for the period 2005-2020 provides valuable information that can be used for various purposes, including adapting transport infrastructure to the needs of sustainable development and efficient use of solar energy. Here are some possible applications of this data in this area:

Integration of solar energy into transport infrastructure: Data on inclined solar irradiation can support the design and installation of photovoltaic panels on buildings and transport-related facilities, such as railway stations, bus and tram stops, and parking lots, to generate clean energy for the facilities themselves or for charging electric vehicles.

Planning for sustainable urban mobility: Solar maps can help plan sustainable urban transport systems by identifying optimal locations for deploying solar-powered electric vehicle charging infrastructure, promoting clean transport.

Adaptation to climate change: Analysis of solar resources can support the development of adaptive strategies for transport infrastructure to account for climate change and the potential increase in temperatures, using solar energy to cool and ventilate transport facilities.

Optimizing energy efficiency: Information on solar potential can be used to optimize the energy efficiency of existing transport infrastructure by identifying opportunities to reduce energy consumption using solar energy for lighting, signaling, and other needs.

Improving safety and reliability: Solar data can improve transportation infrastructure's safety and reliability by providing alternative energy sources for emergency lighting, signaling, and communication systems in the event of power outages.

Research and Education: The data provided on solar potential can be used in research and education programs to understand renewable energy resources and develop innovative solutions for their efficient use in transportation and other areas.

These applications highlight the importance of access to accurate and detailed data on solar resources in the planning and developing sustainable and adaptive transport infrastructure that can meet the challenges of climate change and the need to reduce carbon emissions.

5. Expected changes in solar radiation

Climate change is associated with several expected changes in solar radiation that can significantly impact the efficiency of solar applications. As global temperatures increase, changes in cloud cover, atmospheric transparency, and precipitation patterns are predicted to directly affect the quantity and quality of solar radiation reaching the Earth's surface. Some regions may experience an increase in sunny days and a reduction in cloud cover, resulting in higher direct solar irradiation. In other areas, an increase in the intensity and frequency of precipitation may increase diffuse irradiation and decrease direct irradiation. These changes require adaptation of solar systems and modeling of the new conditions to optimize the use of solar energy and achieve maximum efficiency. Understanding and predicting these dynamic changes in solar radiation is essential for future planning and development of sustainable solar applications in transportation infrastructure.

In the context of climate change, Bulgaria is also facing significant changes in solar irradiation that may impact the efficiency of solar applications. Average annual temperatures are expected to increase, which may lead to changes in cloud cover and atmospheric transparency. As a result, several key changes are foreseen:

Increase in direct solar irradiation: In some regions of Bulgaria, especially in the southern and eastern parts, a decrease in cloudiness is expected during the summer months, leading to an increase in direct solar irradiation. This could increase the efficiency of solar panels and increase electricity production from solar energy.

Changes in diffuse solar irradiation: In other regions, especially in the mountainous and northwestern parts, an increase in cloudiness and precipitation frequency can be observed. This will lead to an increase in diffuse solar irradiation, which can also be effectively used by solar panels, especially those with a higher sensitivity to diffuse light.

Seasonal variations: Changes in climatic conditions will also affect seasonal variations in solar radiation. During the winter months, an increase in cloudiness and precipitation can be expected, reducing the amount of available solar energy. However, when solar radiation is strongest in spring and summer, a more extended period of clear weather and higher radiation levels are predicted.

Effects on long-term sustainability: These changes require adaptation and optimization of solar systems in Bulgaria. Accurate modeling and forecasting of solar radiation are becoming increasingly important to maximize the efficiency and sustainability of solar applications in the context of changing climate conditions.

These expectations emphasize the need for detailed research and adaptive strategies for integrating solar technologies into Bulgaria's transport infrastructure to fully utilize the opportunities solar energy offers in the face of climate change.

6. Conclusions

The present study demonstrates the effective application of artificial neural networks for modeling solar data essential for solar applications in transportation infrastructure. The main findings of the study include:

- **Performance of ANNs:** Artificial neural networks show high accuracy and reliability in modeling various climate parameters, including horizontal, inclined, and reflected solar irradiation, diffuse fraction and terrain albedo. These models provide detailed and accurate data essential for the planning and optimizing solar applications in transport infrastructure.
- **Improving resilience:** The use of solar technologies in transport infrastructure contributes to increasing systems' resilience and adaptability to climate change. By integrating solar panels for street lighting and powering signal systems, carbon emissions are reduced, and energy efficiency is increased.
- **Practical applications:** The models developed through ANN provide valuable data for generating monthly and annual solar irradiation maps for Bulgaria's territory. These maps are helpful for various applications, such as designing sustainable road surfaces, optimizing street lighting, and planning charging stations for electric vehicles.
- **Improving transport infrastructure:** Accurate modeling of solar data provides reliable predictions that can help improve transport infrastructure. This includes adapting to extreme weather conditions, reducing maintenance costs, and increasing the safety and functionality of transport systems.

The research highlights the importance of applying artificial neural networks to plan and improve the sustainability of transport infrastructure. In the future, these methods can be extended and improved by incorporating additional climate parameters and larger data volumes, leading to even more accurate and reliable models for predicting and optimizing solar applications.

Acknowledgments

The development was financed under the "GIS-based analyses of climate change impacts on road infrastructure to build adaptive capacity" project at the Research, Consultancy, and Design Center of UACEG.

7. References

1. Republic of Bulgaria, National Strategy for Adaptation to Climate Change and Action Plan, 2019. <https://www.moew.government.bg/bg/klimat/mejdunarodni-pregovori-i-adaptaciya/adaptaciya/>
2. St. Ivanova. Artificial neural networks for energy efficiency, UACEG - Sofia, 2022.
3. St. Ivanova. Determination of cumulative monthly solar radiation on building surfaces with different slope and exposure. Solution by artificial neural networks based on satellite data. // First departmental scientific conference at the Department of "Industrial and Agrarian Buildings" at UACEG, Sofia, February 2022.
4. Photovoltaic Geographical Information System (PVGIS) https://joint-research-centre.ec.europa.eu/photovoltaic-geographical-information-system-pvgis_en
5. NASA POWER (Prediction Of Worldwide Energy Resources) <https://power.larc.nasa.gov/>
6. NASA Earth Observations <https://neo.gsfc.nasa.gov/>
7. N. Igawa. Improving the All Sky Model for the luminance and radiance distributions of the sky. // Solar Energy 105 (2014), pp. 354-372.
8. St. Ivanova, C. Gueymard. Simulation and applications of cumulative anisotropic sky radiance patterns. // Solar Energy 178 (2019), pp. 278-294.
9. Home page: <https://ait.uacg.bg/Transport/>
10. Page for solar potential by geographic positions: <https://ait.uacg.bg/Transport/solar-location.html>
11. Page for monthly and annual solar data maps: <https://ait.uacg.bg/Transport/solar-map.html>

ТЕХНОЛОГИИТЕ МЕЖДУ ИНДУСТРИЯ 4.0 И 5.0

TECHNOLOGIES BETWEEN INDUSTRY 4.0 AND 5.0

Д-р Личева, Теодора
Нов български университет, София, България
teodora_licheffa@abv.bg

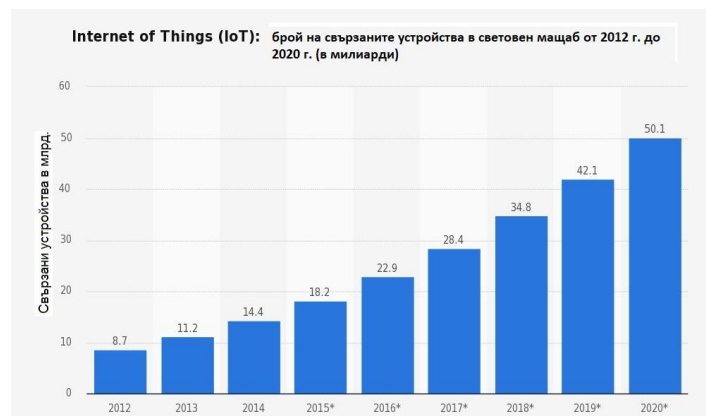
Abstract: In today's uncertain times, one thing is certain – new technological opportunities will change our way of life. The benefits are wide-ranging – energy consumption, supply chain traceability, healthy lifestyles and good healthcare. Technologies develop very dynamically and enter every sphere of our daily life.

Keywords: TECHNOLOGIES, INDUSTRY 4.0, INDUSTRY 5.0, BLOCKCHAIN

1. Увод

Технологиите се развиват все по-бързо и ако не се използват в държавния и частния сектор ще доведе до тяхното изолиране и изоставане. Свърхобемът от данни, които се получава трябва да се събере, обработи и анализира, за да може да се използва и съхранява.

Все още не може да се определи точно обхвата и сложността на промяната, която ще настъпи с бързото развитие на технологиите, но е сигурно, че ще промени света и миогледа ни.



Фиг. 1 Брой свързани устройства¹

2. Индустрия 4.0

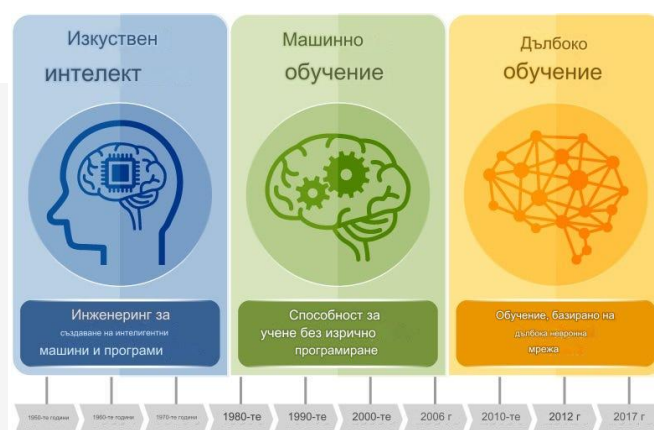
След механизацията (Индустрия 1.0), масово производство (Индустрия 2.0) и автоматизацията (Индустрия 3.0), се полага основата на „интернетът на предметите и услугите“. Интернет се превръща в неразделна част от производството. В индустриалното развитие на 4.0 се създават изключителни възможности за растеж и конкурентоспособност.

Интелигентните машини споделят информация едни с други, и дори могат да се организират, за да координират процеси и крайните срокове. Производството става все по-гъвкаво и ефективно.

Човешкият фактор играе важна роля, като контролира и проследява производствените процеси. Информационно-технологично базираните системи за видеонаблюдение имат възможност да се настроят към индивидуален подход и да се персонализират спрямо работоспособността на служителите.

¹ <https://xn--80aahddubcb0awc4bnhip4t.bg/tema/sklad-badeshteto/etikata-chetvartata-industrialna-revolyuitsiya>

Индустриалното развитие 4.0 не се ограничава до само производствени мощности и постижения. Невижданият бум на разработване и внедряване на изкуствен интелект (Artificial intelligence), Интернет на нещата (Internet of Things), 3D принтиране, блокчейн (Blockchain), Големите данни (Big data), роботиката, нано и биотехнологиите, променят из основи начина ни на живот, работа и дори отношението един към друг.



Фиг. 2 Развитие на алгоритмите на Изкуствения интелект²

Множеството интелигентни продукти се свързват с интернет и събраните данни се използват за оптимизиране на продуктите или да се използват за генериране на нови услуги.

Четвъртата промишлена революция е възможност да се внедрят новите технологични възможности в бизнеса, обществото и правителството. Истинската промяната промяна е в социалните и икономическите системи, които градят живота ни.

3. Индустрия 5.0

Началото на Индустрия 5.0 започва с разумни стратегии, добро планиране, отворена и приобщаваща комуникация и силната, стабилна основа на Индустрия 4.0. Концепцията е приета през 2020 г.³ от Европейския съюз и за основна причина е изтъкнато по-доброто интегриране на социалните и

² <https://medium.com/@mail2princeyadav/machine-learning-vs-deep-learning-b5c5a4fc5c>

³ Европейска комисия, Генерална дирекция „Научни изследвания и иновации“, *Industry 5.0 : human-centric, sustainable and resilient*, Publications Office, 2020, <https://data.europa.eu/doi/10.2777/073781>

екологически потребности в новите технологични възможности.

Индустрията трябва да съчетава цифров и екологичен преход и това насочва към ново поколение промишлена революция.

Петата по ред промишлена революция е синхрон между човека и технологиите. Новият производствен модел се фокусира именно върху това сътрудничество. Работната сила е обединяваща за човешкия и цифровия ресурс, за разработване на интелигентна фабрика на бъдещето.

В модела на Индустрия 5.0 се предвиждат и инвестиции за образование и обучение на работниците, за да могат да се оптимизират уменията им и да намерят място в съвременния бизнес модел⁴.



Фиг. 3 Цели на Индустрия 5.0

Индустрия 5.0 ще стъпи на подобрения и нововъведения от предишните промишлени революции.

■ Екологични решения – опазване на околната среда е от ключово значение на Индустрия 5.0. Ще се въведат нови технологии и корпоративна чувствителност, за бъде производството по-ефективно и устойчиво. Минимизиране на количеството ресурси и отпадъци ще направи производствения процес по-ефективен;

■ Оптимизиране на разходите – стремежът е насочен към създаване на нови бизнес модели, където да се инвестират по-малко ресурси, за да се получат по-големи ползи;

■ Креативност и индивидуалност – индивидуалността не може да се постигне с пълната автоматизация на процесите. В Индустрия 5.0 се цели работниците да се концентрират върху нови стратегии или да прилагат креативност, за сметка на рутинни и повтарящи се задачи.

⁴ <https://www.momenta.one/industry5.0>

4. Между Индустрия 4.0 и 5.0

В Индустрия 4.0 основна цел е да се увеличи броят на иновационните технологии в промишлеността, дори до момент на състезание между човек и робот за работно място. В Индустрия 5.0 акцента се измества и се търси сътрудничество между човек и робот, вместо да има заменяемост. Хората трябва да си сътрудничат с технологиите, като има използват силните страни и да компенсират съответните им слабости. Другата разлика е, че в Петата промишлена революция се търси постигане на благосъстояние на широк кръг от обществото, а в Индустрия 4.0 – движещата сила е печалбата.

В изследователски статии⁵ от 2020 г. се появява идеята за настъпването на Петата индустриална революция и тя се определя като сътрудничеството между човек и машина.

В основата на концепцията за Индустрия 5.0 са следните технологични категории:

- Сътрудничество между човека и машината
- Изкуствен интелект;
- Блокчейн;
- Дигитални близнаци и симулации;
- Технология за анализ, мигриране и съхранение на данни;
- Технологии за енергийна ефективност, възобновяема енергия, съхранение и автономност;
- Технологии, вдъхновени от природата, и умни материали.

Индустрия 5.0 се фокусира върху взаимодействието човек-машина и създава по-безопасни работни места, в която работници с подкрепата на роботи/коботи си партнират и изпълняват опасни производствени процеси.

На световния икономически форум Давос 2019 г. се популяризира идеята⁶:

„Blockchain+AI+Human=Magic“

С помощта на изкуственият интелект се създава нова производителност на човешкия труд, блокчейнът⁷ ще помогне на онези, които не са обхванати от банкиране, роботиката ще помогне на хората да поддържат възвръщаемостта на инвестициите целенасочено и много печелившо⁸.

Таблица 1. Разлика между Индустрия 4.0 и Индустрия 5.0⁹

Индустрия 4.0	Индустрия 5.0
---------------	---------------

⁵ <https://europeansting.com/2019/05/16/what-the-fifth-industrial-revolution-is-and-why-it-matters/>

⁶ <https://aryunetworks.com/what-will-the-5th-industrial-revolution-look-like/>

⁷ Личева, Т. „Модерна сигурност в управлението“, Изд. НТС по машиностроене „Индустрия-4.0“ София, 2023, стр. 208

⁸ <https://aryunetworks.com/what-will-the-5th-industrial-revolution-look-like/>

⁹ Собствено представяне на разликите между Индустрия 4.0 и Индустрия 5.0 с помощта на:

<https://www.sciencedirect.com/science/article/pii/S0022435922000288?via%3Dihub#bib0016>

Фокус: Хората и технологиите	Максимализиране	Търси максимален брой и обхват на технологиите и тяхната взаимосвързаност	Търси се синхрон на силните страни както на технологиите, така и на хората; Допълват се.
	Конкуренция / Сътрудничество	Хора и машини се конкурират за работни места	Хора и машини си сътрудничат
Фокус: Благо състоянието	Акцент върху околната среда	○ Няма; ○ Преследва се печалба	○ Търси благосъстояние на цялото общество и планета; ○ Акцент върху устойчиви и възобновяеми ресурси; ○ Печалбата да е с цел.
	Изместване на технологичните граници	Необходимост от доверие в технологиите	Хуманно използване на технологиите

Литература:

1. Европейска комисия, Генерална дирекция „Научни изследвания и иновации“, *Industry 5.0 : human-centric, sustainable and resilient*, Publications Office, 2020,
2. Личева, Т. „Модерна сигурност в управлението“, Изд. НТС по машиностроене „Индуктрия-4.0“ София , 2023, ISBN 978-619-7383-30-0
3. <https://aryunetworks.com/what-will-the-5th-industrial-revolution-look-like/>
4. <https://aryunetworks.com/what-will-the-5th-industrial-revolution-look-like/>
5. <https://data.europa.eu/doi/10.2777/073781>
6. <https://europeansting.com/2019/05/16/what-the-fifth-industrial-revolution-is-and-why-it-matters/>
7. <https://medium.com/@mail2princeyadav/machine-learning-vs-deep-learning-b5c5a4fc5c>
8. <https://www.momenta.one/industry5.0>
9. <https://www.sciencedirect.com/science/article/pii/S0022435922000288?via%3Dihub#bib0016>
10. <https://xn--80aahddubcb0awc4bnhip4t.bg/tema/sklad-badeshteto/etikata-chetvartata-industrialna-revolutsiya>

3. Заключение

С въвеждане на нови технологии, количеството и качеството на данните се подобрява, което води и до подобряване на тяхното използване. То е ключово за решаване на предизвикателствата, свързани с обществото, климатичните промени и околна среда, енергийна устойчивост и просперитет на обществото. Това е и предпоставката за преминаване от Индустрия 4.0 към Индустрия 5.0.

Индустрия 4.0 и Индустрия 5.0 се преплитат и съчетават, за да променят новите технологии по начин, за да се създадат нови възможности да се подобри живота и напредъкът в обществото.

Индустрия 5.0 е отлична възможност за сътрудничество между Големите данни, ИИ, блокчейн, IoT, облачни услуги, роботи и коботи за креативност и въображение в промишлеността. Тази индустриална революция ще създаде нови работни места с по-висока квалификация, но със свобода на изразяването и иновативност в процесите.

Роботите и коботите поемат част от производствените процеси, но решенията са отговорност на хората и се вземат от тях.

Ключово предизвикателство е решаване на проблемите с киберсигурността в индустриалните линии и системи.

С напредъка на новите технологични възможности се предвижда следващите революции в индустрията да се случват все по-бързо и в по-кратки срокове.

Новата индустриална революция цели да направи света по-добър, а не как да подобри ефективността и производителността.

Cybersecurity of the autonomous vehicles as cyber-physical systems

Magdalena Puncheva Marina, Ninoslav Marina

Extended Abstract:

Autonomous vehicles (AV) are expected to have a great impact on the current organization of the transportation systems. Moreover many autonomous vehicles are also used as support intelligent moving objects that carry various items in industrial environments as a part of the future smart manufacturing (Industrie 4.0, MIC25, etc.), hospitals, logistics, and delivery. The autonomous vehicles and robots may be also used in extreme environments.

These types of vehicles are equipped with various sensors, in order to sense the environment and plan their routes and are connected to an infrastructure for sending and receiving data. Such systems are also part of the broader group known as cyber-physical systems (CPS) for the interaction with the humans and since they connect the physical (real) the cyber (online) world. Because of their connection to multiple types of networks, they are vulnerable to various cyberattacks. Compared to conventional cyberattacks, which may cause various damages, both technical and financial, the attacks on the cyber-physical systems and especially the autonomous vehicles, may have catastrophic and fatal consequences. Therefore, the cybersecurity of the CPS requires necessary attention and additional security related actions. In this work, we summarize and discuss the most frequent types of known cyberattacks on the autonomous vehicles. We also identify some existing applications, including machine learning based methods, that aim at improving the cyber-security into a more trustworthy system.

Autonomous vehicles are expected to profoundly influence the way people perceive and use transportation today. There are several classes of autonomous vehicles, depending on the level of automation and the need of the human interaction. The class with the highest level of automation and no human interaction depends mostly on the data collected from the environment via its embedded systems called sensors. Since there is no infrastructure build along the roads, currently the AV is connected to the Internet through the wireless cellular networks such as fourth generation (4G) and fifth generation (5G) of mobile cellular network. Besides the first generation (1G) that was offering mobile voice service in a non-cellular network, all the other generations second (2G), third (3G), and fourth (4G), are considering voice and messaging (2G, 3G, 4G), emails and internet (3G), broadband and high-speed internet (4G). Besides the evolution in the services there was a significant evolution of the mobile terminals which modernized a great deal in less than three decades. A fundamental leap forward was made by the fifth generation (5G). 5G does not only improve the look and the performance of the mobile terminal, but also allows connection of various mobile devices to the cellular network. Of course this will be possible for all the generations beyond 5G. Hence, for the first time we have a possibility to connect various embedded sensors (or things) that can be as tiny as few millimeters to as large as different autonomous machines, including industrial machinery, vehicles, flying objects and underwater objects. The whole concept is known as Internet-of-Things (IoT) and gained prominence in the past decade.

Because of the broadcast nature of the wireless connections such a setting especially with numerous sensors is quite vulnerable to cyber attacks of various nature. A step further is the lack of solid cyber security protection of the 5G network. Traditionally, the features that are discuss as important in the 2/3/4/5/5+ generations are rarely taking the security challenges seriously. The reasons are multiple: not a big need (for 2G and 3G), focus on the data rates, frequencies, quality of service (QoS), roaming, latency, and similar. Probably the biggest reason of not addressing the security problem seriously is the lack of competence and knowledge that will be implemented on a wireless network with moving sensors. Traditionally, the cryptographic algorithms are made for point-to-point communication and even that one in a fixed and non-dynamic setting. With the rise of necessity for involving multiple sensors the whole security challenge changes since the connections are not anymore point-to-point but maybe multipoint to multipoint. This fact and the broadcast nature of the wireless channel raise the serious question on how to define the security. One fundamental solution that is not related to how long should be the key until the next technology improvements is the so-called information-theoretic secrecy or security. This is a concept defined by Claude Shannon in the 1940's and researches the reliable communication between one of more parties while the adversaries get no information. The concept applies the security on the physical layer (first layer) and we do not need to deploy further security algorithms that are applied typically in the higher layer, including the second (data link layer), the third (network layer), the fourth (transport layer) or the highest application layer. The problem with the information-theoretic secrecy is that besides the fact that it is well researched, there is a lack of practical secrecy capacity approaching codes that may be implemented in practice.

Since the autonomous vehicles are constantly connected and moving their vulnerability is quite high. In addition to the connectivity vehicle to infrastructure (V2I), i.e., to the network access points, such as base stations, hotspots, backbone network or intelligent roadside units, AVs have also a possibility for the so-called vehicle-to-vehicle (V2V) communications or vehicle-to-everything (V2X). In addition to the connectivity to the wireless networks AVs of the future will have a necessity to be connected to the power grid for the electric vehicles or to any device allowing machine-to-machine interaction to provide fuel (e.g., fossil, ethanol, hydrogen) or other services (maintenance, assistance, tire pressure, cleaning, etc.).

Ignoring the cybersecurity of autonomous vehicles or any autonomous machine can lead to serious damages, injuries or fatalities. The hacker may take the control of the steering wheel, block the engine or transmission, block the wheels, mislead the camera system, interfere with the maps or disturb the whole AV system in various ways. Hence, an early detection and prevention is more crucial than in the other networking systems.

The most common cyber-attacks on autonomous and non-autonomous vehicles could be summarized below:

- Injection attack – injecting messages to the controller area network (CAN) of the vehicle
- Remote hacking – to take control of the infotainment system
- Remote attack on the engine
- Remote access on the brakes
- Stealing the remote key data (e.g., by replay attack)
- Spoofing
- Jamming
- Meaconing attacks
- Propagating malware to other vehicles (V2V)
- or propagating malware to any other connected “thing”, so-called vehicle-to-everything (V2X).

Artificial Intelligence (AI) has recently become very popular in solving various problems. We propose the use of AI in protection of autonomous vehicles. AI methods include predictive analysis, intelligent agents, case-based reasoning, natural language processing and machine learning among others. Deep learning and neural networks are special types of machine learning. According to the objective, there exist descriptive AI, predictive and generative AI. In this work we conclude that since the cyber-security of the autonomous vehicles has so many weak points, AI might be a good solution to identify, to protect, to detect, to respond, and finally to recover if needed.

As future work it might be of interest to discuss the advantages and drawbacks of machine learning and artificial intelligence in mitigating all cyber-security threats.

ИССЛЕДОВАНИЕ НАПРЯЖЕННО-ДЕФОРМИРОВАННОГО СОСТОЯНИЯ МЕТАЛЛА ПРИ РЕАЛИЗАЦИИ НОВОГО ИННОВАЦИОННОГО СПОСОБА ПОЛУЧЕНИЯ УПРОЧНЕННОЙ ВИНТОВОЙ АРМАТУРЫ

INVESTIGATION OF THE STRESS-STRAIN STATE IN THE IMPLEMENTATION OF A NEW INNOVATIVE METHOD FOR OBTAINING HARDENED SCREW FITTINGS

Сергей Лежнев¹, Абдрахман Найзабеков², Евгений Панин³, Елена Ширияева⁴, Андрей Толкушкин¹, Дмитрий Куис⁵, Александр Арбуз⁶

¹ Рудненский индустриальный университет, Рудный, Казахстан

² Инновационный Евразийский университет, Павлодар, Казахстан

³ Карагандинский индустриальный университет, Темиртау, Казахстан

⁴ Магнитогорский государственный технический университет им. Г.И. Носова, Магнитогорск, Россия

⁵ Белорусский государственный технологический университет, Минск, Беларусь

⁶ Назарбаев Университет, Астана, Казахстан

Abstract: In recent years, there has been a new trend in the metallurgical industry to develop new innovative processes that combine various metalworking methods into a single technological process. At the same time, each such new development requires in-depth theoretical and experimental studies to determine the optimal geometric and technological parameters of the combined process being developed and its subsequent introduction into production. This work is devoted to computer modeling of a new method for obtaining hardened screw fittings, combining radial-shear rolling and pressing through a screw matrix and studying the stress-strain state of the metal during its implementation.

Keywords: RADIAL-SHEAR ROLLING, STRESS-STRAIN STATE, SIMULATION, SCREW ROD.

Введение

Получение различных черных и цветных металлов и сплавов с ультрамелкозернистой (УМЗ) и нанокристаллической (НК) структурами и, соответственно, высоким уровнем механических свойств и эксплуатационных характеристик является одной из главных задач обработки металлов давлением уже не одно десятилетие. Одним из основных путей решения данной задачи является применение для обработки металлов давлением различных способов, реализующих в процессе деформирования интенсивную пластическую деформацию (ИПД) [1] за счет которой осуществляется равномерная интенсивная проработка литой структуры во всем объеме деформируемой заготовки. Но, к сожалению, использование большинства способов ИПД в реальном производстве металлоизделий затруднительно по ряду причин, в том числе и по причине того что большинство из данных способов не позволяют обрабатывать длинномерные заготовки.

Альтернативным вариантом решения проблемы использования ИПД при обработке металлов давлением в промышленных условиях является разработка совмещенных способов обработки металлов давлением с реализацией в процессе деформирования ИПД. В настоящее время уже разработано целый ряд совмещенных процессов [2-6], которые позволяют обрабатывать длинномерные заготовки с достаточной степенью проработки литой структуры металла для получения в нем ультрамелкозернистой структуры. При этом многие из данных совмещенных способов деформирования позволяют получать в процессе их реализации не только полуфабрикаты, но и готовые изделия. Одной из последних таких разработок является процесс, предложенный в работе [7], который совмещает редуцирующую прокатку стержня круглого или квадратного поперечного сечения в квадратном калибре со скручиванием в формовочной матрице (скручивающем механизме). Данный совмещенный способ позволяет получать упрочненный арматурный профиль с градиентной ультрамелкозернистой структурой [8]. Несмотря на достаточно простую конструкцию и эффективную проработку исходной заготовки (за один цикл деформирования происходит развитие деформации от 0,8 в центральных слоях заготовки до 1,4 в поверхностных слоях), данный способ имеет

существенный недостаток, связанный с его технологичностью. При переводе производства на сортамент другого размера понадобится не только новая матрица, но и новая пара валков с калибром заданных размеров.

С целью повышения технологичности процесса получения упрочненного арматурного профиля, нами была предложена новая схема его формирования (Рисунок 1), включающая деформирование заготовки круглого поперечного сечения на стане радиально-сдвиговой прокатки (РСП) и последующее скручивание заготовки в формовочной матрице специальной конструкции (аналогичной представленной в работе [7]). В данном случае при переналадке производства на новый типоразмер арматурного профиля понадобится установка лишь новой матрицы нужного диаметра, так как для деформирования заготовки другого диаметра на стане РСП необходимо просто свести или развести валки сводится лишь к изменению величины обжатия. Кроме этого в данном случае будет идти более интенсивная проработка структуры металла, так как из работ [9] известно, что при радиально-сдвиговой прокатке, в очаге деформации реализуется схема напряженного состояния близкая к всестороннему сжатию с большими сдвиговыми деформациями, т.е. оптимальная для формирования УМЗ структуры.

Целью данной работы является компьютерное моделирование предложенного совмещенного способа получения винтовой арматуры и изучение напряженно-деформированного состояния металла при его реализации.

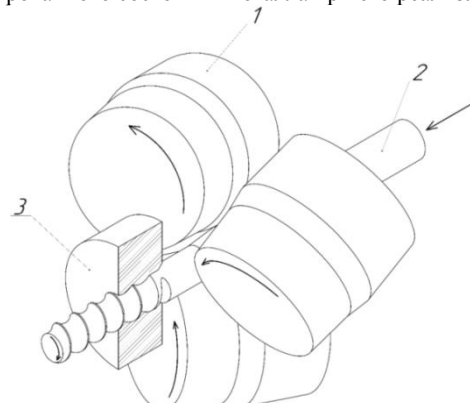


Рис. 1. Совмещенный способ получения арматурного профиля, 1 – прокатные валки, 2 – заготовка, 3 – матрица

Материал и методы

При создании компьютерной модели нового совмещенного процесса, в качестве основного оборудования были использованы параметры стана радиально-сдвиговой прокатки 10-30. В качестве исходной заготовки был взят пруток диаметром 22 мм и длиной 120 мм. Раствор валков был установлен на получение заготовки диаметром 19 мм. В качестве материала заготовки была выбрана сталь AISI 1015, аналог стали 15.

В качестве рабочего инструмента для компьютерного моделирования был использован программный комплекс Deform. При создании модели в программе Deform были приняты следующие допущения: материал заготовки является изотропным, в нем отсутствуют начальные деформации, несплошности, включения и т.д.; тип заготовки – пластичный, тип инструментов – жесткий; материал заготовки – сталь AISI 1015, соответствующая стали марки 15; начальная температура нагрева заготовки – 1100 °С; коэффициент теплопередачи между заготовкой и инструментом был принят равным 5 кВт/(м²·°С), как рекомендуемое значение программой Deform для процессов деформирования; теплопередача между заготовкой и окружающей средой была активирована; скорость вращения валков была принята равной 100 об/мин; на контакте заготовки и инструмента был установлен тип трения по Зигелю с коэффициентом трения 0,7 на валках и 0,1 в матрице.

После расчета была получена модель совмещенного процесса, в котором заготовка сперва прокатывается на стане радиально-сдвиговой прокатки, при этом на выходе из очага деформации на поверхности заготовки образуется винтовая реборда. При попадании в матрицу заготовка подвергается скручиванию вокруг продольной оси на угол, соответствующий конфигурации винтового канала.

При изучении любого процесса обработки давлением одним из важнейших этапов является исследование напряженно-деформированного состояния. Это дает возможность оценить возникающие значения деформаций и напряжений, выявить места возникновения их критических значений, что, в свою очередь, позволяет проанализировать прочность инструмента, вероятность возникновения дефектов и т.д.

Учитывая тот факт, что разрабатываемый способ деформирования имеет достаточно сложную схему течения металла, которая меняется по мере продвижения заготовки вдоль продольной оси, наиболее целесообразно рассматривать исследуемые параметры в продольном направлении заготовки в двух вариантах – на поверхности и в центральной части образца. Это позволит не только установить численные значения параметров НДС, но и оценить возникаемое градиентное распределение аккумулируемых факторов (эквивалентная деформация).

Результаты и обсуждение

Эквивалентная деформация

Для определения значений деформации необходимо находить значения компонентов соответствующего тензора, которые для трехмерного течения металла очень сложно визуализировать. Поэтому обычно при рассмотрении параметров НДС используют простой показатель интенсивности деформации, или так называемую эквивалентную деформацию.

При рассмотрении параметра «эквивалентная деформация» было установлено, что на стадии радиально-сдвиговой прокатки накопление данного параметра происходит преимущественно в поверхностных слоях заготовки (Рисунок 2). После выхода из валков на поверхности уровень деформации достигает 6,75÷8,5, постепенно снижаясь до 5 в периферийной зоне.

При попадании в винтовой канал матрицы заготовка подвергается интенсивной деформации за счет скручивания. В результате происходит значительный прирост деформации по всему сечению заготовки. Так, в поверхностных слоях заготовки, которые приобретают объемную винтовую форму, величина деформации достигает 11,0, плавно снижаясь до 9,0 в периферийной зоне. Центральная зона на данной стадии деформирования прорабатывается максимально интенсивно – здесь величина деформации колеблется в диапазоне 8÷9. Причем, после прохождения винтового канала можно отметить существенное снижение разброса значений деформации.

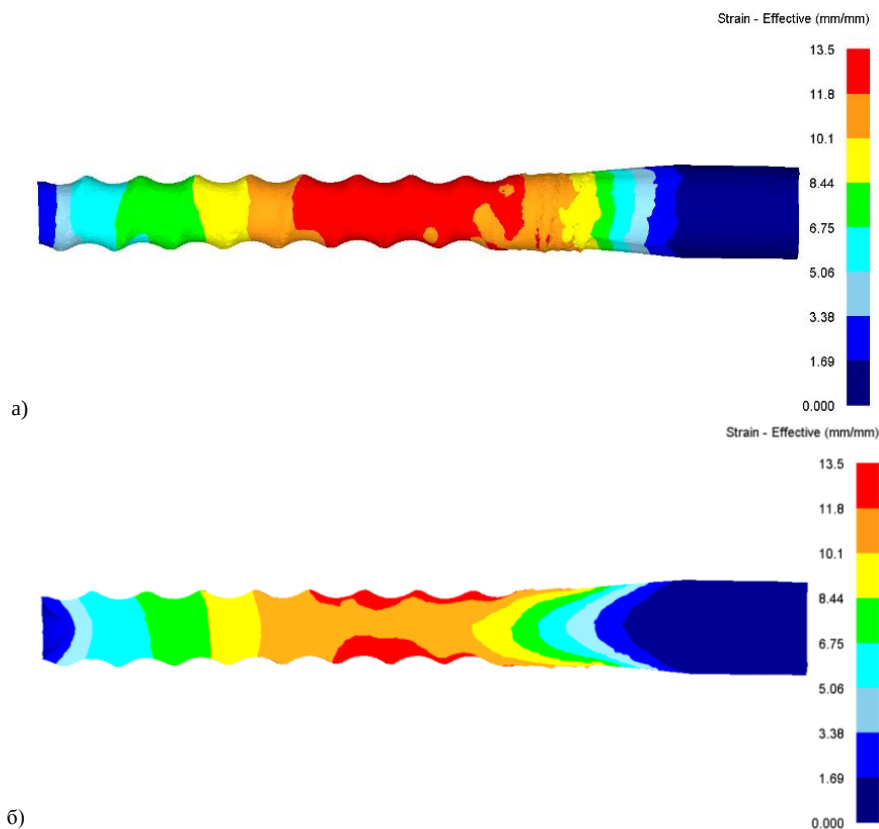


Рис. 2. Распределение эквивалентной деформации на поверхности (а) и в центре заготовки (б)

Напряженное состояние

Для оценки природы напряженного состояния был использован коэффициент Лоде [10]. Данный коэффициент позволяет провести оценку природы возникающей деформации в заготовке, т.е. определить, какой тип деформации реализуется в конкретной точке – растяжение, сжатие или сдвиг. Величина коэффициента варьируется от -1 до 1. Значение от 0 до 1 соответствует сжатию; от 0 до -1 соответствует растяжению; значение коэффициента, стремящееся к 0, соответствует сдвигу.

Для расчета данного коэффициента вдоль начальной длины заготовки, в центральном сечении были сделаны 120 точек с шагом 1 мм. При изучении продольной прокатки было установлено, что для расчета данного параметра достаточно 1 измерения [11]. Однако в случае изучаемого совмещенного процесса, заготовка наряду с поступательным, в каждый момент времени совершает еще и вращательное движение. С учетом того, что распределение параметров НДС, рассмотренных выше, является неоднородным, было решено провести расчеты коэффициента Лоде в трех продольных зонах, соответствующих классификации С.П. Галкина [12]:

- первое измерение было сделано строго по продольной оси вращения заготовки;

- второе измерение было сделано в продольном направлении со сдвигом 5 мм от оси вращения, что соответствует периферийной области;

- третье измерение было сделано в продольном направлении со сдвигом 10 мм от оси вращения (на расстоянии 1 мм от поверхности), что соответствует поверхностной области.

На рисунке 3 представлены картины изменения траектории течения металла во всех трех областях по мере продвижения через очаг деформации в валках и винтовой канал матрицы. Траектория течения в центральной зоне по мере движения через валки и матрицу практически не меняется. В периферийной зоне после прохождения очага деформации в валках, на промежуточном участке между данным очагом деформации и первым винтовым каналом в матрице образуется зона противодавления, вызывающая закручивание образца, что видно на рисунке 3б. По мере продвижения металла по винтовому каналу закручивание продолжается уровень течения металла по высоте остается практически неизменным, что отражается на уровне точек на конце заготовки.

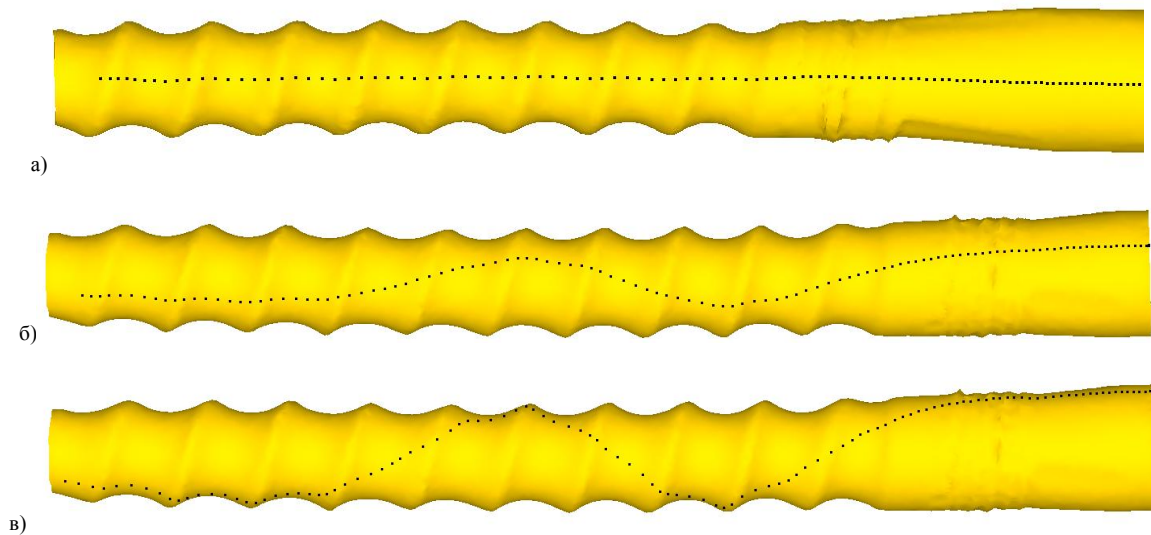


Рис. 3. Траектории течения металла в центральной (а), периферийной (б) и поверхностной (в) зонах

В поверхностной зоне (рисунок 3в) происходит наиболее интенсивное изменение как формы траектории движения металла, так и его скорости по абсолютному значению, что отчетливо видно по изменению расстояния между соседними точками в начале и в конце процесса деформирования. На промежуточном участке между данным очагом деформации и первым винтовым каналом в матрице наблюдается интенсивное закручивание образца на 90 градусов, после попадания этой области металла в винтовой канал уровень скручивания растёт с 90 до 270 градусов. При этом, помимо обжатия в поперечном сечении, заготовка получает некоторое растяжение в продольном направлении.

Поскольку при деформации в валках и формовке в матрице в каждый момент времени находится строго определенная область заготовки, необходимо определить примерную протяженность данных зон. Область контакта металла с валками находится в диапазоне точек № 72÷85 (Рисунок 4а). Область формовки в первом винтовом канале матрицы находится в диапазоне точек № 102÷109 (Рисунок 4б).

На рисунке 5 представлены результаты расчета коэффициента Лоде для всех трех областей. При анализе зоны прокатки в валках видно, что на данном участке природа деформации меняется в поперечном направлении от оси вращения к поверхности. Центральная часть заготовки находится под действием растягивающих напряжений ($\mu = -0,6 \div -0,7$), что соответствует данным С.П. Галкина о характере

течения металла при радиально-сдвиговой прокатке. По мере удаления от центра, в периферийной части возникают сжимающие напряжения ($\mu = 0,6 \div 0,7$), что является следствием действия давления от валков в очаге деформации.

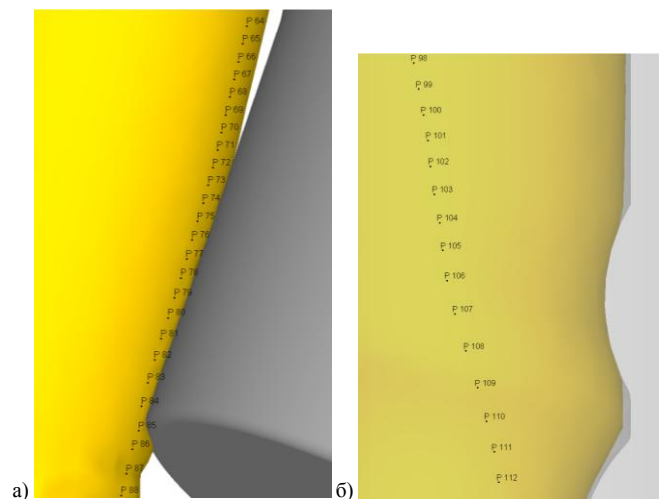


Рис. 4. Протяженности зон прокатки (а) и формовки (б)

Поверхностная зона представляет наибольший интерес, поскольку в данной области при продольной прокатке величина коэффициента Лоде составляет $0,95 \div 1,0$, что, по сути, является действием лишь сжимающих напряжений. Однако, в данном случае величина коэффициента Лоде составляет около 0,1. Такая величина коэффициента на определенной протяженности говорит о наличии деформации сдвига, что также подтверждается полученными при моделировании результатами изучения деформированного состояния, где видно, что наибольшая проработка металла осуществляется именно в поверхностных слоях.

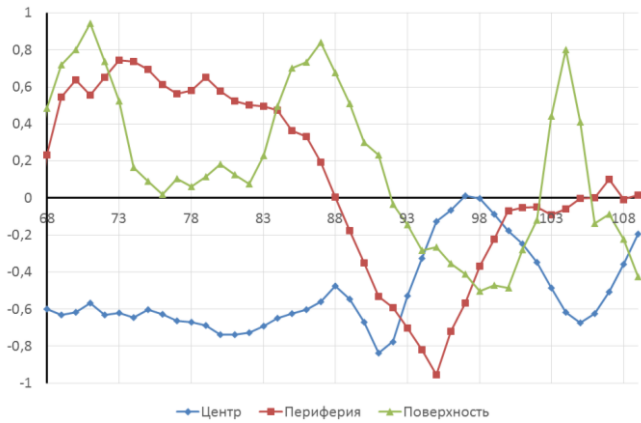


Рис. 5. Коэффициент Лоде

Для дальнейшего анализа напряженного состояния были рассмотрены следующие параметры: эквивалентное напряжение (или интенсивность напряжений) и среднее гидростатическое давление. При РСП максимальные значения эквивалентного напряжения возникают в зонах контакта металла с вальками, достигая величины 200 МПа (Рисунок 6). На свободных от контакта зонах значение эквивалентного напряжения колеблется в пределах $80 \div 100$ МПа. При попадании в матрицу максимальный уровень напряжений развивается лишь на первом витке, где формируется винтовой профиль. Здесь значение эквивалентного напряжения находится в диапазоне $140 \div 150$ МПа в поверхностных слоях, подвергающихся формовке, и $90 \div 100$ МПа в центральной части заготовки. В последующих витках величина напряжения сравнительно невелика (около 30 МПа), поскольку в них металл движется уже с готовой винтовой формой.

При анализе зоны формовки было выявлено, что при прохождении участка винтового канала в матрице, во всех трех зонах, как и при радиально-сдвиговой прокатке, создаются деформации разных типов. Причем центральная и поверхностная зоны имеют примерно одинаковый, пикообразный характер деформаций, но разный по знаку. Вершины пиков соответствуют моменту прохождения через середину ширины винтового канала. Центральная зона подвергается действию растягивающих напряжений, здесь коэффициент Лоде составляет -0,2 на входе в винтовой канал, в

центре винтового канала достигает максимального значения -0,67, затем плавно снижаясь снова до -0,2.

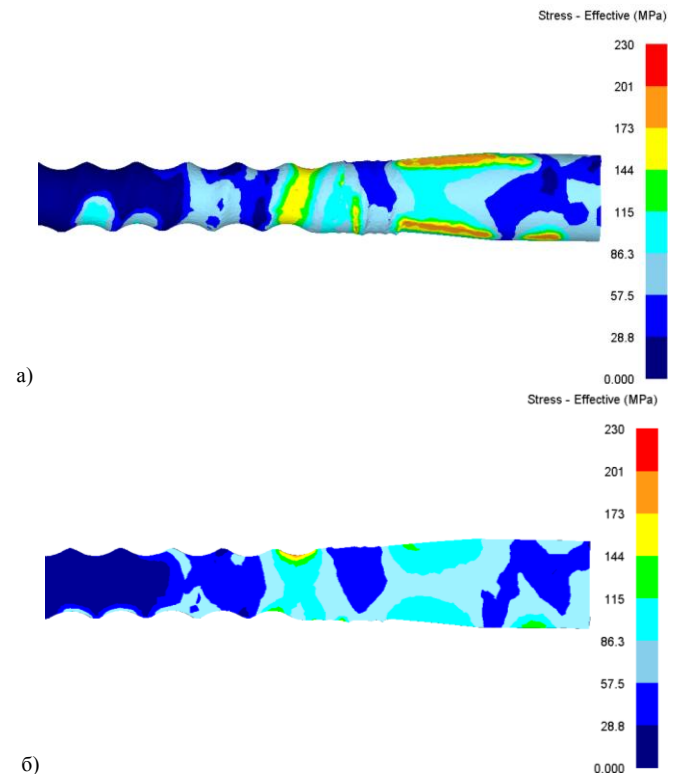
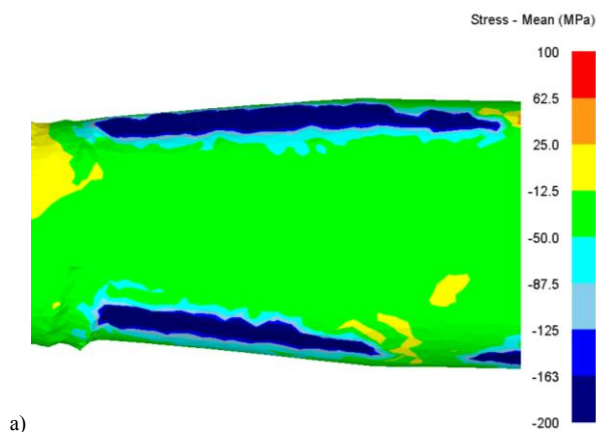


Рис. 6. Распределение эквивалентного напряжения на поверхности (а) и в центре заготовки (б)

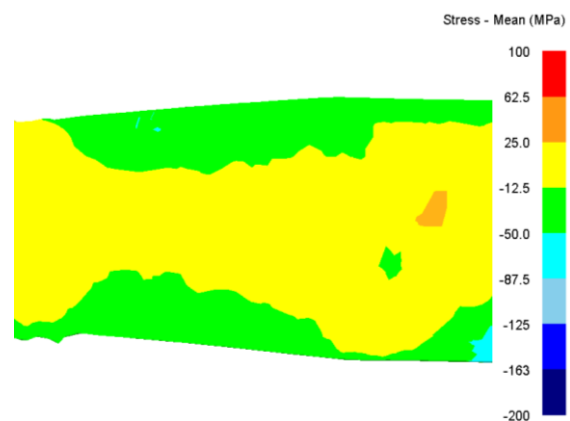
При рассмотрении эквивалентного напряжения необходимо понимать, что, являясь подкоренным выражением, его величина всегда положительна. Для оценки величины напряжения с учетом знака целесообразно рассматривать среднее гидростатическое давление или среднее напряжение (Рисунок 7).

Для наиболее эффективной оценки данного параметра в Deform рекомендуется выставить диапазон шкалы таким образом, чтобы нулевая отметка примерно попадала на какую-либо границу цветовой шкалы. Это позволит легко выявить зоны растягивающих и сжимающих напряжений.

При радиально-сдвиговой прокатке (Рисунок 7а) в зонах контакта металла с вальками развиваются преимущественно сжимающие напряжения достигая величины -200 МПа на поверхности и снижаясь до -12,5 ближе к центральной части заготовки. На свободных от контакта зонах величина данного параметра достигает -50 МПа. При формовке в матрице (Рисунок 7б) в первом витке развиваются сжимающие напряжения около -80 МПа в поверхностных слоях, подвергающихся формовке, и -60 ÷ -40 МПа в центральной части заготовки.



а)



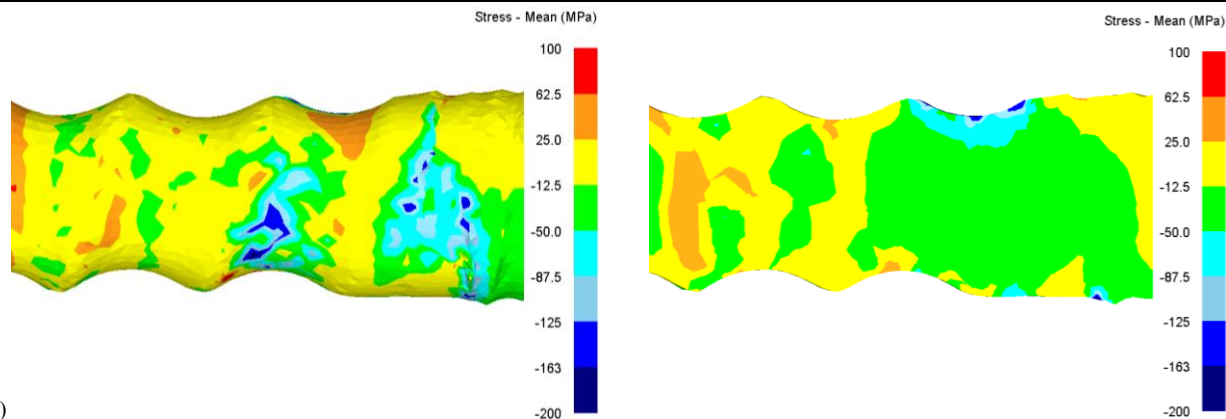


Рис. 7. Среднее гидростатическое давление на поверхности и в центре заготовки при радиально-сдвиговой прокатке (а) и формовке в матрице (б)

Выводы

В работе была поставлена задача моделирования совмещенного процесса радиально-сдвиговой прокатки – прессования для получения арматурного профиля из круглой заготовки. По результатам компьютерного моделирования был выполнен анализ параметров напряженно-деформированного состояния в зонах радиально-сдвиговой прокатки и формирования винтового профиля в матрице на поверхности и осевой частях заготовки. В качестве параметров НДС были выбраны эквивалентная деформация, показатель Лоде, эквивалентное напряжение и среднее нормальное напряжение. Максимальное значение эквивалентной деформации достигается при формировании винтового профиля на заготовке в матрице, что свидетельствует об интенсивной проработке начальной структуры заготовки. Также за счет сжимающих напряжений в матрице снижается неравномерность распределения эквивалентной деформации по диаметру заготовки в сравнении с зоной радиально-сдвиговой прокатки.

В периферийной зоне, за счет скручивающего действия винтового канала, по всей протяженности действуют деформации сдвига, что отражается на коэффициенте Лоде, величина которого находится здесь в пределах $-0,05 \div 0,1$. Поверхностная зона помимо скручивания получает также и некоторое обжатие по высоте за счет внедрения винтовой реборды. В результате величина коэффициента Лоде в этой области существенно изменяет свое численное значение. На входе в винтовой канал металл подвергается давлению со стороны реборды матрицы, за счет чего до середины ширины канала здесь возникает сжатие. Это отражается на величине коэффициента, возрастающего до 0,8. После преодоления середины канала сжимающее действие прекращается, поскольку контур заготовки уже соответствует конфигурации канала. В результате значение коэффициента резко снижается до $-0,1$.

Данное исследование финансировалось Комитетом науки Министерства науки и высшего образования Республики Казахстан (Грант № AP14869135).

Литература

1. Утяшев, Ф.З. Современные методы интенсивной пластической деформации. Уфа: РИК УГАТУ, 2008. - 313 с.
2. Shimov G.V., Bogatov A.A., Kovin D.S. FEM Simulation of Copper Busbar Pressing on the Continuous Extrusion Line "CONFORM". Solid State Phenomena. 2018. Vol. 278. 158-165.
3. I.P. Semenova, A.V. Polyakov, G.I. Raab, R.Z. Valiev, T.C. Lowe, Enhanced fatigue properties of ultrafine-grained Ti rods processed by ECAP-Conform, Journal of materials science 47 (22) (2012) 7777-7781.

4. Ghaforian, N.H., Khalili, K., Gerdooei, M. Theoretical and Numerical Investigation of Required Torque in ECAP-Conform Process // Metallurgical and Materials Transactions B. 2020, Vol. 51. pp.519.

5. E. Panin, I. Volokitina, A. Volokitin, A. Naizabekov, G. Akhmetova, S. Lezhnev, A. Tolkushkin, A. Esbolat, R. Yordanova. Finite Element Modeling of ECAP-Linex Combined Process of Severe Plastic Deformation. Modelling and Simulation in Engineering, Volume 2023, Article ID 1573884.

6. Naizabekov, A., Lezhnev, S., Panin, E., Volokitina, I., Arbutz, A., Koinov, T., Mazur, I. Effect of combined rolling - ECAP on ultrafine-grained structure and properties in 6063 Al alloy. // Journal of Materials Engineering and Performance. 2019. №28. pp.200.

7. Г.И. Рааб, А.Г. Рааб, Способ упрочнения и формирования винтового арматурного стержня, Патент РФ 2640705.

8. G.I. Raab, A.G. Raab, A new combined method of intensive plastic deformation with the formation of the screw profile of the rod, Machines. Technologies. Materials (2018) 320-322.

9. Galkin S.P. Radial shear rolling as an optimal technology for lean production // Steel in Translation. – 2014. - №44 (1). – P. 61-64.

10. W. Lode, Trial on the influence of the middle principle voltage on the flow of metal iron, copper and nickel, Zeits. Phys. 36 (1926) 913-939.

11. А.Б. Найзабеков, С.Н. Лежнев, М.В. Чукин, Разработка и теоретическое исследование совмещенного процесса «прокатка-прессование», позволяющего получать металл с субультрамелкозернистой структурой, Обработка сплошных и слоистых материалов 37 (2011) 3-12.

12. S.P. Galkin, Regulating radial-shear and screw rolling on the basis of the metal trajectory, Steel in translation 34 (7) (2004) 57-60.

ПРОЕКТИРОВАНИЕ ПОДЪЕМНОЙ МАШИНЫ ДЛЯ КОНТЕЙНЕРНОЙ ТЕХНОЛОГИИ НА ОТВАЛАХ

DESIGN OF A LIFTING MACHINE FOR CONTAINER TECHNOLOGY ON LANDFILLS

Sergey Kuzmin, Andrey Turbit, Rustem Iskakov

Rudny Industrial University, Rudny, Kazakhstan

Abstract: An analysis of open-pit mining technologies and the current state of mining production indicates an urgent need to develop resource-saving and environmentally friendly technology for creating external dumps. We propose a new container technology for transporting rock mass in containers without construction on the surface of the dump of technological and transport communications. For its implementation, it is proposed to use a rope excavator from the Liebherr company. The most important part of the excavator is the boom, which must withstand the load from the weight of the loaded container. In addition, inertial and wind loads on the working equipment must be taken into account when designing the boom. The design of the boom was carried out using computer modeling methods. To identify the performance of the dragline, a dynamic calculation of the excavator drives was carried out and a static calculation was performed with the machine checking for stability and balance of the rotary platform. All calculations prove the operability of the proposed option.

Keywords: WALKING EXCAVATOR, BOOM, END LOAD, MINING TECHNOLOGY, COMPUTER MODELING, DUMP, ROCK.

Введение

При разработке рудной залежи с помощью открытых горных работ, приходится вынимать большое количество вскрыши, объём которой во много раз превосходит объём вынимаемой руды. Добытая из карьера горная масса складывается во внешние отвалы, которые занимают большие площади и сильно ухудшают экологию региона [1]. Горнодобывающая промышленность подошла к новому этапу развития и способна поставлять на рынок не только добытое полезное ископаемое, но и сформированные в процессе ведения горных работ техногенные георесурсы, обладающие определенной ценностью. Основными направлениями по формированию и использованию георесурсов являются:

- целенаправленное селективное складирование вскрышных пород в виде техногенных месторождений с целью их дальнейшего освоения;

- целенаправленное формирование карьерного пространства и отвалов вскрышных пород в качестве самостоятельных объектов различных назначений, в том числе место для отдыха, строительная площадка, ёмкость под складирование опасных отходов производства.

Технологиям отвалообразования в настоящий момент не уделяется большого внимания, и они значительно отстают по производительности и себестоимости от технологий в карьере. В настоящее время известны несколько технологий отвалообразования, зависящие от средств транспортирования горной массы и физико-механических свойств горных пород.

Отвалообразование хорошо дробленых скальных пород при доставке их на отвалы железнодорожным транспортом может производиться отвальными плугами. Отвальный плуг представляет собой самоходный или прицепной вагон специальной конструкции, оборудованный с одной или с двух сторон отвальными лемехами. Технология отвалообразования заключается в перемещении во время своего движения по путям разгруженной из думпкарв породы отвальными лемехами в сторону откоса отвала. При этом отвалообразование осуществляется по всей длине отвального тупика [2].

Укладка крепких горных пород в отвалы при доставке автомобильным транспортом производится мощными бульдозерами. Автосамосвалы разгружаются на некотором расстоянии от бровки по периферии отвала. Бульдозеры перемещают ее под откос, оставляя на бровке предохранительный вал.

При разработке месторождений с крепкими горными породами и перевозке железнодорожным транспортом применяется отвалообразование механическими лопатами. Технология отвалообразования заключается в экскавации

разгружаемой из думпкарв породы и укладке ее в отвал одноковшовым экскаватором. Экскаватор располагается ниже уровня железнодорожных путей на высоту разгрузки экскаватора. Порода из думпкарв разгружается в приямок, сооружаемый самим экскаватором. После разгрузки породы экскаватор производит укладку ее сначала в нижний подступ, а затем в верхний [3].

Для хорошо раздробленных пород и при применении в карьере конвейерного транспорта возможна организация отвалообразования с помощью транспортно-отвальных мостов [4].

Для всех известных технологий отвалообразования необходимо организовать движение транспортных средств по поверхности отвала, поэтому необходимо создать хорошие дороги, так как отсыпанная порода имеет низкую устойчивость, то за автомобильными и железнодорожными дорогами надо осуществлять постоянный контроль.

Работа горнотранспортного оборудования на отвалах при существующих технологиях вызывает ряд проблем, главной из которых является движение техники возле призмы возможного обрушения породы [5]. Необходимо разрабатывать специальные мероприятия, способствующие предотвратить обрушение горных пород.

Материалы и методы

Для решения всех обозначенных вопросов предлагается для транспортировки укладки пород в отвал применить контейнерную технологию. Главным элементом технологии является мобильная подъемная машина, при этом не возникает необходимость в строительстве автомобильных дорог или железнодорожных путей на борту и поверхности отвала.

Доставленная из карьера руда загружается в контейнеры, которые подъёмными машинами доставляются сразу к месту разгрузки. Контейнеры оборудованы возможностью саморазгрузки, в результате чего горная масса высыпается непосредственно в место укладки на отвал. Контейнеры с породой устанавливаются перед фронтом отвальной заходки, а подъёмная машина размещается на специальной площадке предотвала.

При использовании контейнерной технологии происходит снижение себестоимости транспортирования из-за уменьшения времени разгрузки транспортных средств, следовательно, растет производительность всего комплекса оборудования подъёма горной массы.

Технологическая схема контейнерного подъёма на отвалах обладает достаточной надёжностью и гибкостью. Она позволяет постепенно наращивать высоту подъёма, распределить оборудование подъёма по всему отвалу,

распределить горную породу по разным видам, что даст возможность дальнейшего её использования.

Контейнерный подъём позволяет решить некоторые проблемы отвалообразования горных пород:

во-первых, подъем горной массы выполняется с использованием подъемных машин, что снижает энергозатраты, по сравнению с транспортированием автотранспортом;

во-вторых, в отличие от железнодорожного транспорта при укладке породы в отвал не выполняется процесс перевалки горной массы из прицепа;

в-третьих, большая длина стрелы подъемной машины позволяет значительно увеличить высоту отвального яруса и ширину заходки.

За счет применения нового комплекса оборудования может быть увеличена высота отвала, что приведет к сокращению его площади. Это позволит получить значительный экологический и экономический эффект.

Контейнерная технология изменит порядок работы, при этом снижается высота подъема горной массы за счет ее послонной горизонтальной укладки. Большие размеры стрелы подъемной машины позволяют увеличить количество укладываемой горной массы в пределах отвальной заходки.

Безопасность горных работ будет достигнута отсутствием транспортных коммуникаций на поверхности отвала. На земной поверхности рядом с отвалом будет создан мобильный перегрузочный пункт, где будет осуществляться перецепка контейнеров. Порода будет автоматически выгружаться в необходимой точке на поверхности отвала, после чего пустой контейнер доставляется назад в зону загрузки подъемной машиной.

Заполнение контейнеров горной массой на отвалах карьера будет осуществляться через промежуточный бункер пластинчатым или ленточным питателем. Данная схема отвалообразования представлена на рисунке 1.

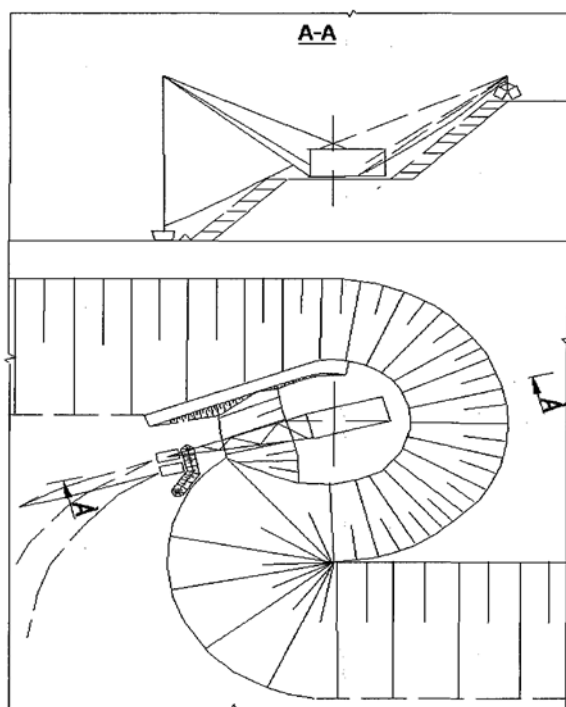


Рис. 1. Схема отвалообразования при контейнерной технологии с послонной укладкой породы

По этой схеме горную массу доставляют от забоев автосамосвалами любой грузоподъемности и выгружают в бункер пластинчатого питателя, который заполняет контейнеры по мере их подачи карьерной подъемной машиной.

В качестве подъемной машины предлагается применить кран на базе канатного экскаватора фирмы Liebherr HS-8300 (Рисунок 2). Экскаватор HS 8300 разработан для тяжелых

условий горных работ и впечатляет своей высокой производительностью. Гидроаккумулятор обеспечивает возможность рекуперации энергии во время работы, при этом снижается расход топлива и износ лебёдки свободного сброса.



Рис. 2. Канатный экскаватор HS-8300 с грейфером

Результаты и обсуждение

Для определения высоты отвала воспользуемся схемой, изображенной на рисунке 3.

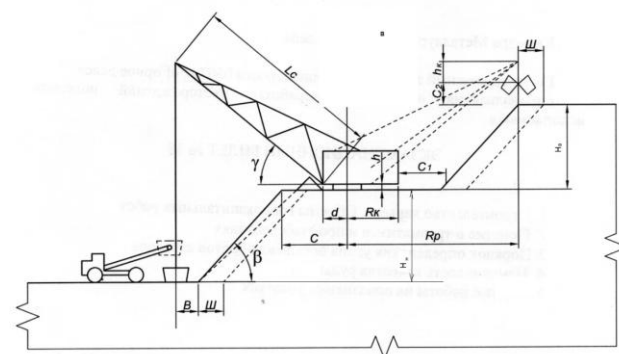


Рис. 3. Схема к определению высоты подъёма

Радиус разгрузки подъемной машины R_p , м, определяется по формуле:

$$R_p = L_c \times \cos(\gamma) + d, \quad (1)$$

где L_c – длина стрелы, м;

γ – угол наклона стрелы, град.;

d – расстояние от оси вращения подъемной машины до пяты его стрелы, м.

Высота нижнего уступа H_{ny} , м, определяется по формуле:

$$H_{ny} = (R_p - B - C) \times \tan(\beta), \quad (2)$$

где B – расстояние от оси контейнера до нижней бровки нижнего уступа, м;

C – безопасное расстояние от оси вращения драглайна до верхней бровки нижнего уступа, м;

β – угол откоса уступа, град.

Высота верхнего уступа H_{vy} , м, по условию безопасности вращения подъемной машины определяется по формуле:

$$H_{\text{вы1}} = (R_p - C_1 - R_k) \times \text{tg}(\beta), \quad (3)$$

где C_1 – безопасное расстояние между кузовом подъёмной машины и откосом верхнего уступа, м;

R_k – радиус вращения кузова подъёмной машины, м.

Высота верхнего уступа $H_{\text{вы2}}$, м, по условию подъёма контейнера определяется по формуле:

$$H_{\text{вы2}} = L_c \times \sin(\beta) + h - h_k, \quad (4)$$

где h – высота оси пяты стрелы над уровнем земли, м;

h_k – высота переподъёма контейнера при установке на верхний уступ, м.

Высота верхнего уступа (по условию минимума) определяется в результате сравнения $H_{\text{вы1}}$ с $H_{\text{вы2}}$. За значение уступа принимается меньшее значение. Высота подъёма определяется по формуле:

$$H_{\text{п}} = H_{\text{ну}} + H_{\text{вы}}, \quad (5)$$

Расчеты были проведены для длины стрелы подъёмной машины 50 метров. На основании выполненных вычислений построен график зависимости высоты подъёма от угла наклона стрелы подъёмной машины (в соответствии с рисунком 4).

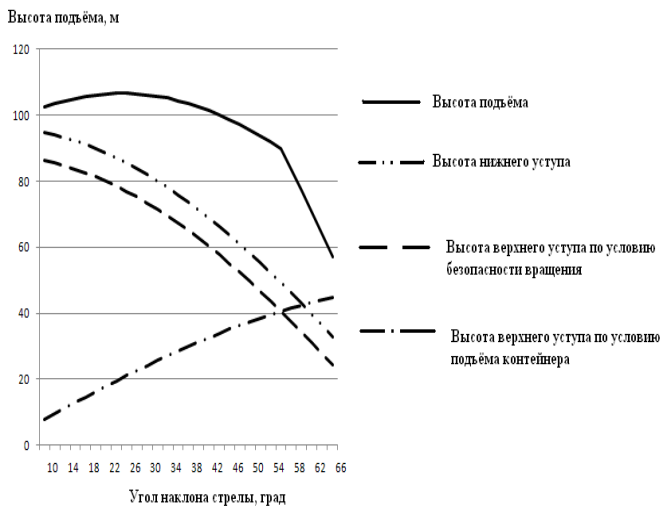


Рис. 4. Зависимости высоты подъёма от угла наклона стрелы подъёмной машины

Учитывая необходимость подъёма на нижнем и верхнем ярусах отвала на 40 м, принимается угол установки стрелы подъёмной машины 56° . При этом общая высота подъёма достигает 80 м за один подъём

Преимуществом контейнерной технологии на отвале является возможность снижения средней высоты подъёма породы. Для этого породу укладывают горизонтальными слоями в пределах шага перемещения подъёмной машины. По мере отсыпки породы между передвижками подъёмной машины высота подъёма породы постепенно повышается до максимальной, равной высоте отвального яруса. Таким образом, средневзвешенная высота подъёма породы равна половине высоты отвального яруса. После заполнения части отвальной заходки с одной точки установки подъёмной машины ее перемещают по предотвалу и повторяют заполнение отвальной заходки в пределах шага передвижки горизонтальными слоями снизу вверх.

Увеличение высоты отвала от 40 метров при традиционной технологии до 80 метров при новом способе отсыпки уменьшит площадь занятую отвалами на 175 тысяч м^2 в год.

Выводы

Применение контейнерной технологии при отвалообразовании шагающими экскаваторами показывает, что при выгрузке породы из контейнеров в отвал снижается энергоёмкость отвалообразования и повышается экономическая выгода отвала. Это позволит сократить площади, занятые отвалами и смягчить вредное воздействие горных работ на окружающую среду.

Данное исследование финансировалось Комитетом науки Министерства науки и высшего образования Республики Казахстан (Грант № AP19675410).

Литература

1. Rakhmangulov, A., Burmistrov, K., Osintsev, N. (2021). Sustainable open pit mining and technical systems: concept, principles, and indicators. Sustainability. 13(3), 1101. <https://doi.org/10.3390/su13031101>
2. Bogusz, K., Sulich, A. (2019). The sustainable development strategies in mining industry, In: Soliman, K.S. (ed.) Education Excellence and Innovation Management through Vision 2020. International Business Information Management Association (IBIMA). 6893-6911. ISBN 9780999855126.
3. Stevens, R. (2011). Mineral Exploration and Mining Essentials. Pakawau Geomanagement, Inc. 317 p. ISBN: 9780986722103.
4. Koptev, V.Yu., Kopteva, A.V., Ivanova, T.S. (2021). Directions for the development of transport machines for open-pit mining. Journal of Applied Engineering Science. 19(1), 137-141. DOI:10.5937/jaes0-28708
5. Oggeri, C., Fenoglio, T., Godio, A., Vinai, R. (2019). Overburden management in open pits: Options and limits in large limestone quarries. International Journal of Mining Science and Technology. 29(2), 217-228.

ЭВОЛЮЦИЯ МИКРОСТРУКТУРЫ ПРИ ПРОКАТКЕ В РЕЛЬЕФНЫХ ВАЛКАХ С ФАКТОРОМ АСИММЕТРИИ ПО РАЗЛИЧНЫМ ТЕХНОЛОГИЧЕСКИМ СХЕМАМ

MICROSTRUCTURE EVOLUTION DURING ROLLING IN RELIEF ROLLS WITH AN ASYMMETRY FACTOR ACCORDING TO VARIOUS TECHNOLOGICAL SCHEMES

Evgeniy Panin¹, Alexandr Arbuz², Almas Yerzhanov¹, Dmitry Kuis³, Aibol Esbolat¹

¹ Karaganda Industrial University, 30 Republic ave., Temirtau, 101400, Kazakhstan

² Nazarbayev University, 53 Kabanbay Batyr ave., Astana, 010000, Kazakhstan

³ Belarusian State Technological University, 13a Sverdlova str., Minsk, 220006, Belarus

e-mail: cooper802@mail.ru

Abstract: In this work, finite element modeling of the microstructure evolution of a copper billet during asymmetric rolling in relief rolls according to various technological schemes was performed. The following technological schemes of deformation were proposed: 1) the workpiece after the first pass of deformation is re-fed into relief rolls with 180° edging along the axis; 2) the workpiece after the first pass of deformation is re-fed into relief rolls without any changes; 3) the workpiece after the first pass of deformation is re-fed into relief rolls with a transverse shift for the relief period. A comparison of the simulation results showed that deformation with a 180° rotation of the workpiece between passes significantly reduces the influence of the asymmetry factor. This is reflected in the level of grain grinding and its more uniform distribution. Deformation without changing the position of the workpiece between the passes has the opposite effect, such a scheme significantly increases the influence of the asymmetry factor. Deformation with transverse displacement of the workpiece during the unloading period between passes has the effect of a "golden mean".

Keywords: ROLLING, RELIEF ROLLS, ASYMMETRY, SIMULATION, MICROSTRUCTURE.

Разрабатываемый процесс интенсивной пластической деформации, представляющий собой комбинацию прокатки в рельефных валках с наложением фактора асимметрии, хорошо зарекомендовал себя на теоретическом этапе исследований. Так, в работах [1-2] были изучены различные параметры процесса при использовании кинематической и геометрической асимметрии. Также в работе [3] были рассмотрены энергосиловые параметры данной схемы деформирования.

Проведенный анализ параметров НДС показал, что одного цикла деформирования (прокатка заготовки в рельефных валках и выравнивание в гладких валках до первоначальной формы) недостаточно для качественной проработки металла. При многопроходном деформировании обычно исследуется влияние числа проходов на уровень проработки и изменение начального размера зерна. Однако данный подход при прокатке обоснован в случае симметричной нагрузки от обоих валков. В рассматриваемой схеме прокатки целесообразно рассмотреть различные технологические схемы деформирования, которые будут учитывать влияние фактора асимметрии. В соответствии с этим были предложены следующие технологические схемы деформирования:

1) заготовка после первого прохода деформирования заново подается в рельефные валки с кантовкой на 180 градусов вдоль оси;

2) заготовка после первого прохода деформирования заново подается в рельефные валки без каких-либо изменений;

3) заготовка после первого прохода деформирования заново подается в рельефные валки с поперечным сдвигом на период рельефа.

Для оценки эффективности технологических схем деформирования было рассмотрено изменение начального размера зерна, заданного равным 100 мкм. В системе Deform традиционным методом моделирования эволюции микроструктуры является классический метод Джонсона-Мейла-Аврами-Колмогорова (Johnson – Mehl – Avrami – Kolmogorov или JMAK). В методе JMAK окончательный средний размер зерна определяется на основе начального среднего размера зерна, констант материала и полей переменных данных, таких как температура, деформация, скорость деформации и время.

На рисунках 1-3 показаны картины изменения микроструктуры после двух проходов для трех рассматриваемых технологических схем деформирования. При кантовке заготовки на 180 градусов происходит последовательный контакт обеих горизонтальных граней с

каждым валком. В результате влияние фактора асимметрии является минимальным, поскольку каждая грань поочередно имеет контакт с валком большего диаметра. Рассматривая эволюцию микроструктуры после двух проходов с кантовкой на 180 градусов вдоль оси (Рисунок 1), было выявлено, что распределение размера зерна имеет характерный полосчатый вид, чередующийся в поперечном сечении заготовки. Средний размер зерен в наиболее проработанных областях составляет 35 мкм, а в наименее проработанных областях примерно 65 мкм.

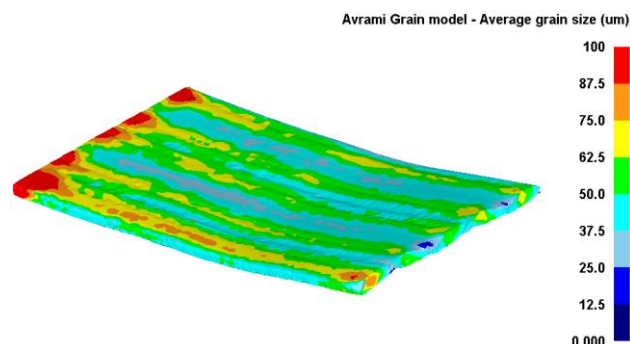


Рис. 1. Эволюция микроструктуры после двух проходов с кантовкой заготовки на 180 градусов между проходами

При деформировании без изменения позиционирования заготовки происходит двойной идентичный контакт заготовки и валков. В результате влияние фактора асимметрии является максимальным, поскольку каждая грань имеет контакт только с одним валком и только с определенной зоной рельефа. Рассматривая эволюцию микроструктуры после двух проходов без изменения позиционирования заготовки (Рисунок 2), было выявлено, что такое деформирование приводит к существенной разнице размера зерна по прорабатываемым областям. Средний размер зерен в наиболее проработанных областях составляет 25 мкм, а в наименее проработанных областях примерно 85 мкм.

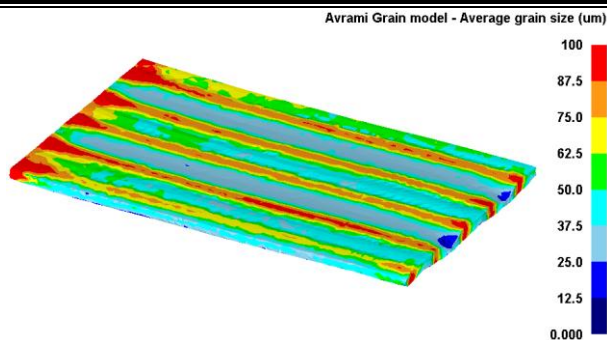


Рис. 2. Эволюция микроструктуры после двух проходов без изменения позиционирования заготовки

При деформировании со сдвигом заготовки на период рельефа происходит двойной контакт заготовки и валков, однако здесь каждая грань имеет контакт с разными зонами рельефа. Рассматривая эволюцию микроструктуры после двух проходов с поперечным сдвигом заготовки на период рельефа (Рисунок 3), было выявлено, что такая схема деформирования приводит к более однородному распределению размера зерна по ширине заготовки. Средний размер зерен в наиболее проработанных областях составляет 30 мкм, а в наименее проработанных областях примерно 50 мкм.

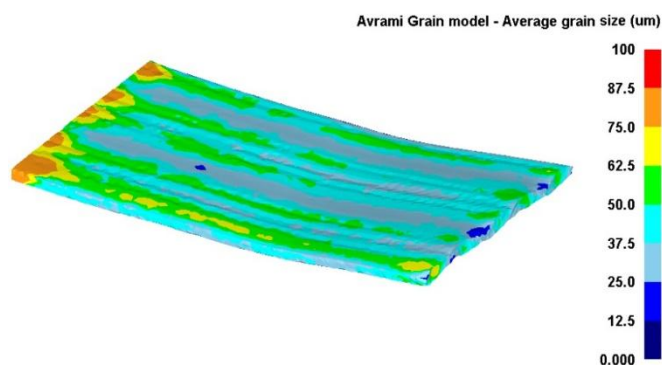


Рис. 3. Эволюция микроструктуры после двух проходов с поперечным сдвигом заготовки на период рельефа

Выводы

В данной работе было выполнено конечно-элементное моделирование эволюции микроструктуры медной заготовки при асимметричной прокатке в рельефных валках по различным технологическим схемам. Были предложены следующие технологические схемы деформирования: 1) заготовка после первого прохода деформирования заново подается в рельефные валки с кантовкой на 180 градусов вдоль оси; 2) заготовка после первого прохода деформирования заново подается в рельефные валки без каких-либо изменений; 3) заготовка после первого прохода деформирования заново подается в рельефные валки с поперечным сдвигом на период рельефа. Сравнение результатов моделирования показало, что деформация при повороте заготовки на 180° между проходами значительно снижает влияние фактора асимметрии. Это отражается на уровне измельчения зерна и его более равномерном распределении. Деформация без изменения положения заготовки между проходами имеет обратный эффект, такая схема значительно увеличивает влияние фактора асимметрии. Деформация с поперечным смещением заготовки в период разгрузки между проходами имеет эффект "золотой середины".

Данное исследование финансировалось Комитетом науки Министерства науки и высшего образования Республики Казахстан (Грант № AP14869080).

Литература

1. A. Esbolat, E. Panin, A. Arbuz, D. Kuis, S. Lezhnev, I. Krupenkin, M. Erpalov. Concepts analysis of asymmetry factor implementation during rolling in relief rolls. *Journal of Chemical Technology and Metallurgy*, 2022, Vol. 57, Iss. 6, P. 1243-1250.
2. A.B. Naizabekov, S.N. Lezhnev, E.A. Panin, A.A. Tymchenko, A.B. Esbolat, Improvement of the deformation technology in relief rolls by asymmetric rolling realization, *Ferrous Metallurgy. Bulletin of Scientific, Technical and Economical Information*, 2021, Vol. 77, Iss. 4, P. 445-454.
3. A. Esbolat, E. Panin, A. Arbuz, D. Kuis, A. Naizabekov, S. Lezhnev, A. Yerzhanov, I. Krupenkin, A. Tolkushkin, A. Kawalek, D. Voroshilov. Investigation of force parameters during rolling in relief rolls. *Journal of Chemical Technology and Metallurgy*, 2023, Vol. 58, Iss. 5, P. 961-968.

Prototype and characteristic measurement of led-lit pet bottle handmade lantern with baby milk dropped into water

Mayumi Tanaka¹, Keiko Ishihara², Akira Okamoto², Yuko Iwahori³, Satoru Tada⁴, Masayuki Yamauchi⁴, Koji Kakugawa⁴, Takeshi Tanaka^{4,*}, and Katia Vutova⁵

Hiroshima Prefectural Hiroshima Minami High School, Hiroshima, Japan¹

Hiroshima International University, Kure, Japan²

TED Ltd. Tokyo, Japan³

Hiroshima Institute of Technology, Hiroshima, Japan⁴

Institute of Electronics, Bulgarian Academy of Sciences, Sofia, Bulgaria⁵

tanaka@cc.it-hiroshima.ac.jp

Abstract: In recent years, natural disasters have frequently occurred in Japan, almost yearly, causing enormous damage. In Hiroshima Prefecture, heavy rains in August 2014 and July 2018 caused large-scale landslides. Power outages continued in the affected areas, and some survivors had been lighting campfires outdoors and using candles indoors. In times of disaster in which bad weather persists, it is difficult to safely secure lighting, which can adversely affect health. Lanterns can be created by putting a few drops of milk into a PET bottle filled with water and shining light from the bottle's bottom. In this study, we used a light emitting diode (LED) as a light source to evaluate the luminance of handmade lanterns. Results suggested that when using tap water with one LED lamp and a 500 ml PET bottle, 1–2 drops of milk are suitable to achieve the optimal light scattering effect.

Keywords: LIGHT EMITTING DIODE, LANTERN, PET BOTTLE, MILK

1. Introduction

In recent years, natural disasters have frequently occurred in various parts of Japan, almost yearly, causing enormous damage. In Hiroshima Prefecture, heavy rains in August 2014 and July 2018 caused large-scale landslides. Power remained out in the affected areas, and some survivors were lighting campfires outdoors and using candles indoors. During disasters where bad weather persists, it is extremely difficult to ensure safe lighting, which can have a negative impact on health. The shelter also has PET bottles of water for adults and children and milk for infants.

In milk, casein exists as colloidal particles with a diameter of 30–600 nm, which are called casein micelles[1]. Scattering by particles with a size similar to the wavelength of light does not depend on the wavelength; all wavelengths of light are scattered, and this scattering is called Mie scattering[2]. Approximately 15 trillion casein micelles (macroaggregates of casein, phosphorus, and calcium) and 2–6 billion fat globules are suspended in 1 mL milk. Milk appears white because each of these particles reflects light[3].

A lantern can be fabricated by putting a few drops of milk into a PET bottle filled with water and shining a flashlight from the bottom of the bottle[4].

In this study, we used a light emitting diode (LED) as the light source and made handmade lanterns using milk and PET bottles. We also considered the shape of the PET bottles used for the lanterns and the water inside them.

2. Experimental

2-1 PET bottles used in the experiment

Figure 1 shows a PET bottle used in the experiment.

Focusing on commercially available PET drinking bottles that may be readily available at evacuation centers, we used carbonated water (soft water, carbon dioxide concentration 3.6 GV) and soft water PET bottles as Samples 1 and 2, respectively (Fig. 1).



Fig. 1 Bottles used

2-2 Infant liquid milk used in the experiment

Generally, liquid milk for infants and infant formula are available at evacuation shelters during disasters, as shown in Fig. 2.



Fig. 2 Example of infant liquid milk for infants at a disaster evacuation center[5]

We dropped the infant liquid milk shown in Fig. 2 into a PET bottle filled with soft or carbonated water. The mass of one drop of milk was 0.05 g, and the deposit was 40 μ L. The amount of casein micelles in 1 mL milk is approximately 15 trillion; thus, one drop of milk (40 mL, 0.05 g) is approximately 60 million.

To understand the overall picture of casein micelles, a stereo pair of scanning electron micrographs[6] and a transmission

electron micrograph[6] of casein micelles are shown in Figs. 3 and 4, respectively.

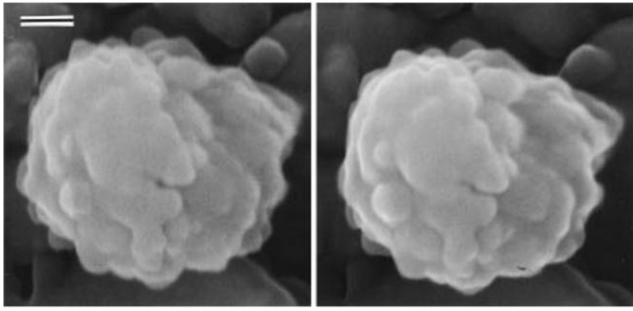


Fig. 3 Stereo pair of scanning electron micrographs of casein micelle adsorbed on a ceramic membrane. The sample was coated with a 2 nm layer of iridium before imaging at a magnification of 70,000 $\times$. Bar = 100 nm[6]

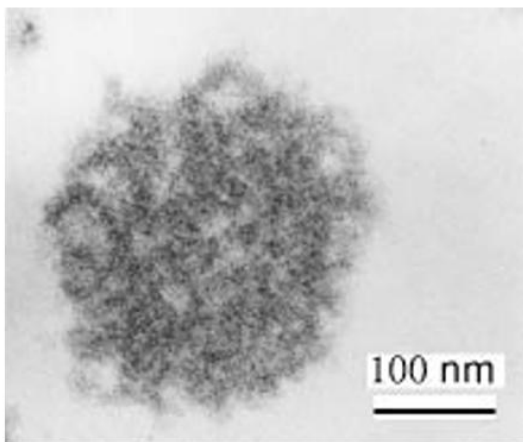


Fig. 4 Transmission electron micrograph of casein micelles in raw milk.

Sample preparation involved solidifying milk by mixing with molten agar, cutting the milk and agar mass into 1 mm³ cubes and fixing in 2.5% glutaraldehyde in phosphate buffer, pH 7.0, postfixing in 1% OsO₄, dehydrating in ethanol series, infiltrating with Spurr's epoxy resin (Electron Microscopy Sciences, Fort Washington, PA), hardening the resin, cutting 70 nm thick sections, and then imaging at 85,000 $\times$ magnification[6]

The size of casein micelles is estimated to be several hundred nm from Figs. 3 and 4, which is comparable to the wavelength of visible light. From the above, the scattering by casein micelles is considered to be Mie scattering, which is the scattering by particles with a size comparable to the wavelength of light. This scattering is wavelength-independent and appears white because all wavelengths of light are similarly scattered. Therefore, in creating lanterns, it may be useful to put milk in water and use the LED light scattering phenomenon.

2-3 LED

In previous studies using flashlights[4], it seems difficult to obtain and analyze illuminance and luminance data due to various light sources. For this reason, in lantern production, we used an LED as the light source, for which data were provided by the manufacturer.

The LED used is Tricon's TUWT6D-5A10-A[7], which has a diameter of 5 mm and a white illumination angle of 10°. Data regarding the device are shown in Table 1 and Figs. 5 and 6.

Table 1 LED characteristics[7]

1. 品名/Name : TUWT6D-5A10-A
2. 発光色/Illuminated Color : 白/White
3. 樹脂色/Resin Color : 無色透明/Clear
4. 最大定格/Absolute Maximum Rating (Ta=25°C)

項目 ITEM	記号 SYMBOL	最大定格 MAXIMUM RATING	単位 UNIT
直流順電流 Forward Current	IF	30	mA
パルス順電流 * Pulse Forward Current	IFP	100	mA
直流逆電圧 Reverse Voltage	VR	7	V
許容損失 Power Dissipation	PD	105	mW
動作温度 Operating Temperature	Topr	-30~85	°C
保存温度 Storage Temperature	Tstg	-40~100	°C
半田温度 Lead Soldering Temperature	Tsol	260°C, 5秒以内 (within 5sec.) (樹脂部から2mm/2mm from resin)	

* 条件/Condition: パルス幅/Pulse width $\leq 1\text{msec.}$, デューティ比/Duty $\leq 1/10$

5. 電気・光学的特性/Electrical-Optical Characteristics (Ta=25°C)

項目 ITEM	記号 SYMBOL	測定条件 CONDITION	特性値/CHARACTERISTIC			単位 UNIT
			MIN.	TYP.	MAX.	
直流順電圧 Forward Voltage	VF	IF=20mA		3.1	3.5	V
直流逆電流 Reverse Current	IR	VR=7V			5	μA
光度 Luminous Intensity	Iv	IF=20mA	20000	35000		mc
色度座標 Chromaticity Coordinates	x	IF=20mA		0.32		
	y			0.32		
指向角度 Directive Angle	2 θ 1/2	参考値 Reference		10		°

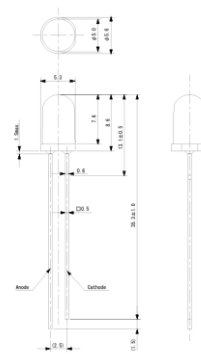


Fig. 5 LED external dimensions (unit:mm)[7]

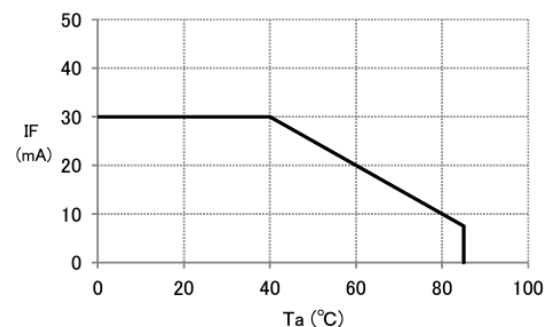


Fig. 6 IF-Ta derating curve[7]

2-4 Luminance measurement of PET bottle lantern

Fig. 7 provides an overview of how to place a PET bottle with milk dropped into water and the angle at which the luminance was measured.

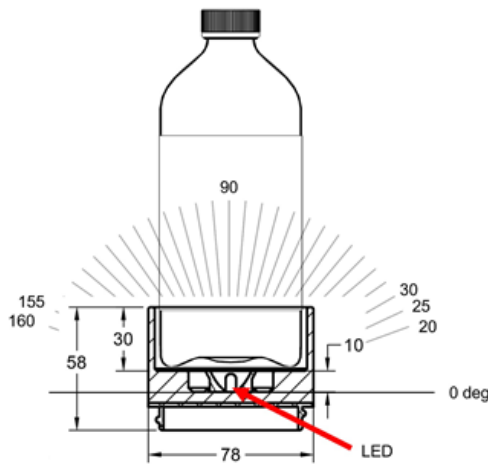


Fig. 7 Placement of plastic bottles and measuring angles

Figure 8 shows a rear view of the device used to measure the luminance of a PET bottle lantern. A protractor for measuring angles (the scale is in units of 1 degree) was placed on top of the LED, and a luminance meter alignment guide was attached to the protractor.

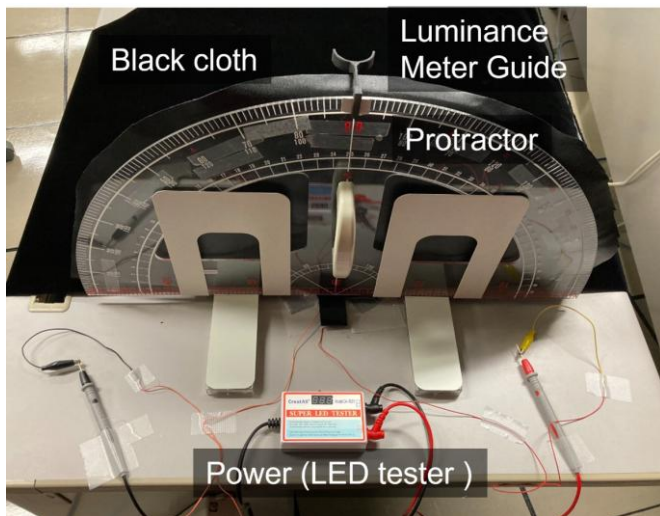


Fig. 8 Rear view of the measurement device (LED). A protractor was placed on top of the LED, and a luminance meter alignment guide was attached to the protractor

An LED tester CA-300V (fangzi) was used as a constant current power supply. The power supply is 100 V for household use; the manufacturer's information states that the output current is 28 mA and the output voltage is automatically adjusted. A BM-9M (TOPCON) was used as a luminance meter. The camera was a LUMIX GH5 (Panasonic), and the lens was a Voigtlander 28 mm F 0.95. The settings used were F = 8, ISO = 640, and shutter speed = 1/10 seconds. As a measurement environment, we installed an LED facing upward on a desk and used a three-dimensional (3D) printer to create a case to place a PET bottle on top of it.

Luminance was measured at 5° increments from a distance of 30 cm along the vertical plane around the LED. However, we avoided the range where there is a case between the LED and luminance meter and set the range to 20°–160°. To maintain a constant

distance from the LED to the luminance meter, a 3D printer was used to create a guide for determining the lens position of the luminance meter on a large semicircular protractor with a 30 cm radius. To reduce the reflected light from the desk and protractor, a black flocked cloth that absorbs 99.9% of visible light was pasted on and the luminance meter guide was coated with matte black paint.

To measure the luminance, the measurer used a guide attached to the protractor to position the lens of the luminance meter so that it was equidistant from the LED and then measured the luminance (Fig. 9).



Fig. 9 Measurement. Using a guide attached to a protractor, the measurer positions the lens of the luminance meter so that it is equidistant from the LED and measures the luminance

3. Results and Discussion

Figure 10 shows the luminance distribution at maximum luminance when adjusted milk was added and mixed immediately after opening the PET bottle.

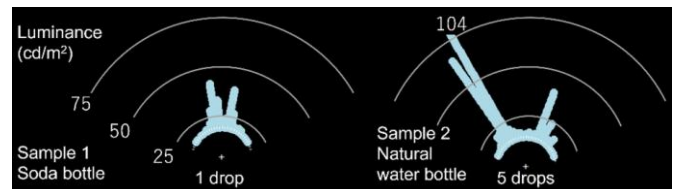


Fig. 10 Luminance distribution at the highest luminance when milk is mixed with the contents immediately after opening

In Sample 1, when there are zero drops, the light is bright only in the upward direction, reflecting the LED directivity. Reflecting the bottle shape, the graph is almost symmetrical around 90°. Because the light scattering effect can be seen with up to two drops, the data for one drop are shown on the left side of Fig. 10.

In Sample 2, the scattered light increased up to five drops of milk. The data for five drops of milk are shown on the right side of Fig. 10.

The above results suggest that adding carbonic acid to soft water increases the scattering effect when milk is added.

To investigate the shape effect of the bottles of Samples 1 and 2, both bottles were filled with water. Further, to investigate the scattering intensity depending on the amount of milk dropped into a PET bottle, Fig. 11 shows photographs of each bottle with respect to the amount of milk dropped.

The top of Fig. 11 shows the bottle of Sample 1, and the bottom shows the bottle of Sample 2. In the horizontal direction of Fig. 11, photographs of zero, one, two, three, and four drops of milk are shown.

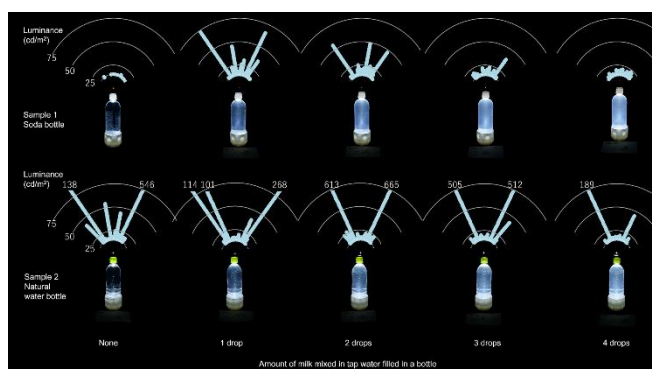


Fig. 11 Luminance distribution for the amount of milk added to bottles of Samples 1 and 2 filled with tap water

The results in Fig. 11 show that for both bottles, although there are some shape differences, the light scattering effect tends to be brighter when there are 1–2 drops of milk.

The above results suggest that adding carbonic acid to soft water increases the scattering effect when milk is added. Further, when bottles were reused as lanterns, the PET bottle shape did not significantly influence the light scattering effect. It has also been suggested that the optimal light scattering effect is 1–2 drops of milk per 500 ml PET bottle or approximately 0.01%–0.02% by volume/weight ratio to water.

4. Conclusion

In recent years, natural disasters have frequently occurred in Japan, almost yearly, causing enormous damage. In Hiroshima Prefecture, heavy rains in August 2014 and July 2018 caused large-scale landslides. Power outages continued in the affected areas, and some survivors had been lighting campfires outdoors and using candles indoors. In times of disaster in which bad weather persists, it is really difficult to safely secure lighting, which can deteriorate health. Lanterns can be created by putting a few drops of milk into PET bottles filled with water and shining light from the bottle's bottom. In this study, we used an LED as a light source to study the luminance of the handmade lanterns. Results suggest that adding carbonic acid to soft water increases the scattering effect when milk is dropped. Further, when bottles were reused as lanterns, the PET bottle shape did not significantly influence the light scattering effect. The optimal light scattering effect is suggested to be 1–2 drops of milk per 500 ml PET bottle or approximately 0.01%–0.02% by volume/weight ratio to water.

Acknowledgments

We thank Enago (www.enago.jp) for the English language review.

References

- 1) Aoki, T., R. Mizuno, T. Kimura, S. Dousako. Structural model of casein micelles and milk processing. *Milk Sci.* 66, 2017, 125–143. (in Japanese)
- 2) Mie, G. Contributions to the optics of turbid media, particularly of colloidal metal solutions. *Annal. Phys.* 25, 1908, 377–445.
- 3) Japan Dairy Industry Association. Milk and dairy products Q&A. https://nyukyoku.jp/dairyqa/2107_140_395/ 2024/05/18 viewed (in Japanese)
- 4) M. Tanaka, et al., Can handmade lanterns be used as lighting during disasters? ~ In search of light that illuminates space and

heart ~. *Journal of the Illuminating Engineering Institute of Japan*. 7, 2024, to be published. (in Japanese)

5) Meiji. Meiji smile easy milk: Meiji Hohoemi easy milk 120ml. <https://www.meiji.co.jp/products/milkpowder/4902705095489.html> 2024/05/18 viewed

6) McMahon, D. J., W. R. Mcmanus. Rethinking casein micelle structure using electron microscopy. *J. Dairy Sci.* 81, 1998, 2985–2993.

7) Toricon. Details of TUWT6D-5A10-A. <https://www.toricon.co.jp/files/20120620083418.Pdf> 2024/05/18 viewed

Extracted Handal seed oil (Citrullus Seeds) used as biodiesel fuel in internal combustion engines

Ahmad O. Hasan

Department of Mechanical Engineering, AL-Hussein Bin Talal University, Maan P.O.Box20, Jordan

E-mail: ahmad_almaany@yahoo.com

Abstract: Most important of fuel economy research and vehicle industry today is to find a unique fuel with high calorific value and low tail pipe emissions . in this research a good effort was spent to find a solution for automotive industry by presenting a new fuel which was extracted from Citrullus and presented as handle seeds oil to be used as bio fuel in internal combustion engine, which is very important issue for automotive industry . The importance of extracting Handel oil has led to investigate the characteristics of chemical and physical properties of new fuel, such as density, specific gravity, kinematic viscosity, flash point, pour point, cloud point, which will help of controlling the composition of biofuel and improve the fuel quality which can be used in internal combustion engines . The two important processes were used for biodiesel production are Esterification and Transesterification.

KEYWORDS: , HANDAL SEED OIL , BIOFUEL

1. Introduction

Biodiesel as promising fuel in the future and replacing diesel fuel as diesel is described as fatty acid methyl or ethyl esters (FAME) from vegetable oils or animal fats. One of the most important properties to characterize fuel is its energy content (1) . Biodiesel is fuel made of plant or animal oils. There are several non-edible oils, animal fat and vegetable oil sources for biodiesel production. Jatropha plant/oil seed containing 27-40% oil while the residue (press cake) can also be processed and used as biomass feedstock to power electricity plants or used as fertilizer as it contains nitrogen, phosphorus, and potassium (2) .

Biofuel can be blend with diesel and be used in compression engines leading to lower exhaust emissions with similar power (3) ; biodiesel has a technical advantage such as less exhaust emission and toxicity, biodegradability, a derivation from a renewable and domestic feedstock, negligible sulfur content, superior flash point, higher combustion efficiency and they would not require a significant modification of existing engine hardware (4) .

Therefore, nonedible plant oils become very promising alternative feedstocks for biodiesel production because of the large demand for edible oils as food, the higher prices of edible oils than that of fossil fuels, and the lower cost of non-edible oil plant cultivation . Biodiesel production has derived from forestry, agriculture, and agro-industrial wastes that exhibit good properties as feedstocks (5)

The transesterification process depends upon a number of process parameters that are required to be optimized in order to maximize the biodiesel yield, such as molar ratio, catalyst concentration, reaction time, and reaction temperature (6) . But before converted the oil to biodiesel the main problem faced vegetables crude oil is particles, impurities, moisture, and high content of free fatty acids, so they must be removed from the crude Hndal oil before converted to biodiesel. Gums should be separated from the oils to make the oils suitable for use as fuel. The use of crude oil without pretreatment can cause problems such as emulsification in the transesterification process (7) , it becomes imperative to analyzed the sunflower methyl ester (SFME) and its blends as an alternate source of fuel for diesel engines. Biodiesel was prepared from sunflower oil by base transesterification. Four- cylinder Deutz F4L912 diesel engine was used to perform the tests on various blends of sunflower biodiesel. The emissions of CO, HC were lower than diesel fuel for all blendstested. The NOx emissions were higher due to the high volatility and high viscosity of biodiesel.

Others blended biodiesel from Citrullus lanatus with petro-diesel, thereby forming blends of different percentages, 5%, 10%, 15%, 20%, and 25%. The calorific values of the biodiesel and blends were all within the ASTM standard with B20 value slightly below that of pure diesel; similarly, the flash point of the pure biodiesel and its blends were above the pure diesel value but falls within the ASTM range. The fuel was safe to handle during storage because it cannot easily spark when exposed to flame and hence they were

recommended for use in the CI engines. The pour points of all the fuel samples conform to the ASTM standards. The cetane numbers are all greater than that of pure diesel and conform to ASTM standards. The biodiesel will be having the shortest possible ignition delay when they burn in the CI engines. (8).

2. Work process

2.1. Collection

The samples were collected from different area in Jordan such as Wadi Ram, Al Baydah and Al Deseh of Maan city south of Jordan as well as Azraq east of Jordan .

2.2. Preparation

Plenty of handal were collected and reserve 20 kg to be used, Hndal Samples were smashed with taking the seeds out and cleaned. Then seeds were dried for three days at room temperature, and then were ground using an electric blender .

2.3. Extraction of oil

First sample was 2 kg of grind seed to be used for extraction purposes. Hundred gram each of grinned seed was packed into the thimble of the soxhlet extractor that incorporates a condensation system . Hexane (extraction agent) is heated to reflux in the distillation flask, the vapors are condensed in the reflux condenser and drop into the chamber containing the thimble with the material to be extracted. The chamber slowly fills with the warm solvent and some of the material is dissolved in the solvent. When the Soxhlet chamber is almost full it is automatically emptied by a siphon sidearm, with the solvent running back down to the distillation flask. This cycle is repeated several times to enrich the solvent in the distillation flask with the extracted material. The mantle heater was set at 60oC close to the evaporation temperature of hexane (64oC).

2.4. Samples Production

The amount of NaOH required for the transesterification of the oil. ca 1 ml oil was dissolved in 10 ml isopropanol and titrated with 0.1M NaOH solution (standardized against KHP). A suitable mass of the catalyst-based on above titration is dissolved in methanol (NaOH: Methanol: Oil; 0.0016 g:0.2ml:1ml) with 20 min vigorous stirring until dissolve all NaOH to prepare methoxide (CH₃O⁻). Then, this mixture is transferred to the reactor (R-flask) containing treated moisture-free oil.

2.5. Esterification Process

In this work, the procedure of known researcher for the esterification process was followed (9) . Titration solution was prepared by adding 1 gram of potassium hydroxide to 1 liter of distilled water. In a small beaker, 1 ml of oil was added to 10 ml of ethanol and was stirred well to get a mixture. 2 drops of phenolphthalein were added as an indicator to the mixture. KOH solution was poured drop by drop by a burette to the mixture until it turns into a pink solution.

It is the most commonly used and important method to reduce the viscosity of vegetable oils. In this process, triglyceride react with three molecules of alcohol in the presence of a catalyst producing a mixture of fatty acids, alkyl ester, and glycerol. The process of

removal of all the glycerol and the fatty acid from vegetable oil in the presence of a catalyst is called esterification (10). glycerol from the bottom of the funnel and pour the top layer (crude biodiesel) back into the beaker.

2.7. Washing process

In this process, water is used for purification, thus, warm distilled water was added to the separator funnel that contains crude biodiesel. Then the separator funnel was shaken gently to mix water with crude biodiesel. Allow the water to settle down of the biodiesel for one day. The water was drained out and the washing process was repeated for 3 to 4 more times until the water has completely become transparent. To remove undesired components such as the excess methanol or existing water, the washed biodiesel samples were dried at 110 °C for 2 hrs

3. Results and Discussion

The results for the comparison of the properties of the biodiesel produced from used vegetable oils and those from diesel fuel are discussed as follows:

3.1. Density

The density of Handal oil is 899 kg/m³, whereas in biodiesel was measured is 888kg/m³. It is higher than that of standard diesel (820-870 kg/m³).

The limit value proposed in pre-standard Fuels for vegetable oil compatible combustion engines - Fuel from rapeseed oil ranged between 900 to 960 kg/m³, so as not to exclude vegetable oils such as castor, jatropha, mahua, neem, or babassu oil, it was mentioned that, indicating density of Handal agrees with pre-standard Fuels for vegetable oil compatible combustion engines- Fuel. (11)

3.2. Specific gravity

The specific gravity of Handal oil is 0.898 at 20°C, while in biodiesel was observed is 0.886. The Specific gravity of Thumba oil slightly variable between 0.880 and 0.905.

3.3. Kinematic viscosity

The significant physical properties of Handal oils indicate that the kinematic viscosity of Handal oils is 16.49 cSt at 40°C, which is an order of magnitude almost 4 times higher than that of diesel.

Viscosity is a rapid indicator of fuel quality before use, especially if the nature of the feedstock is not well known, or if the oil could have been deteriorated or polymerized during storage. The kinematic viscosity of Handal oil is higher than that of conventional diesel fuel but lower than that of heavy fuel oil. (12)

3.4 Flash point

The flash point of Handal oil is very high (above 180 °C) and the heating values are in the range of 39.603 MJ/kg, while biodiesel was measured is 39.5, as compared to diesel fuels which is about 66 MJ/kg. The used Handal biodiesel oil had a flash point (OC) of 49.5 and was lower than that, Petroleum diesel, Diesel fuel, and fossil diesel, which had values of, 68.3, 68, and 47 °C, respectively.

3.5. Pour point

The pour point of Handal oil value is -9 °C, whereas biodiesel value is +3. The results observed that the vegetable oil shows the least values of pour point ranging from -9 °C to 40°C, while the Pour point of the used Thumba oil was -8, and lower than that of Thumba biodiesel which is -12. The corn oil had the pour points is -40°C while the Jatropha oil and Karanja oil have to pour point values 8°C and 6 °C, respectively (Table 3). Vegetable oils can be mixed with conventional diesel in any proportion and blends can be used successfully in engines (13). It can be concluded that all the chemically modified oil products have the pour points below 0°C which is a desirable property for lubricating oil. The pour point obtained (12°C) is still much higher than that is required for any lubricant application (-6°C for two- stroke engine lubricant as per IS14234) (14).

4. Conclusion

4.1.A lower viscosity and density was observed by dilution of handal oils with diesel oil., flash points of handal oils blends were higher than that of diesel oil.

4.2. The Handal oils exhibit relatively different flash point, cloud point, pour point, and sulfur content.

4.3. Handal oil showed have high energy content and obviously good fuel properties,

4.4. The treated handal oils, could be converted into biodiesel, which has properties suitable for application in diesel engines

4.5. The transesterification process has been the most suitable and acceptable method for biodiesel production

4.6. The density of Handl seed oil agrees with pre-standard Fuels for vegetable oil compatible combustion engines- Fuel.

4.7. The kinematic viscosity of Handal oil is higher than that of conventional diesel fuel but lower than that of heavy fuel oil

REFERENCES

- (1) HOEKMANA, S., BROCHA, A., ROBBINS, C., CENICEROS, E. & NATARAJAN, M. 2012. Review Of Biodiesel Composition, Properties, And Specifications. *Renewable And Sustainable Energy Reviews*, 16, 143–169.
- (2) SANTORI, G., DI NICOLA, G., MOGLIE, M. & POLONARA, F. 2012. A Review Analyzing The Industrial Biodiesel Production Practice Starting From Vegetable Oil Refining. *Applied Energy*, 92, 109-132.
- (3) SIDIBÉ, S. S., BLIN, J., VAITILINGOM, G. & AZOUMAH, Y. 2010. Use Of Crude Filtered Vegetable Oil As A Fuel In Diesel Engines State Of The Art: Literature Review. *Renewable And Sustainable Energy Reviews*, 14, 2748-2759.
- (4) ELANGO VAN, T. & JERYRAJKUMAR, L. 2015. Biodiesel And Its Properties From Various Feedstocks. *International Journal For Trends In Engineering & Technology*, 1(2).
- ERHAN, S.Z. & ASADAUSKAS S. 2000. Lubricant Basestocks From Vegetable Oils. *Ind. Crops Prod.* 11, 277–282. [http://dx.doi.org/10.1016/S0926-6690\(99\)00061-8](http://dx.doi.org/10.1016/S0926-6690(99)00061-8).
- (5) HASSAN, S. S., WILLIAMS, G. A. & JAISWAL, A. K. 2019. Lignocellulosic Biorefineries In Europe: Current State And Prospects. *Trends Biotechnol.* 37, 231- 234.
- (6) USTA, S., M., MATHUR, Y. B. & SAMAD, A. 2016. Optimization Of Transesterification Process For Production Of Thumba Methyl Ester. *International Journal Of Innovative Research In Science, Engineering And Technology (IJIRSET)*, 5(1), 311-316.
- (7) THEIN, T. M. L., JINDAL, V. K., JINDAL, R. & YOSWATHANA, N. 2019. AcidCatalyzed Esterification Pretreatment Of High Free Fatty Acid Crude Rice Bran Oil For Biodiesel Production. *Environment And Natural Resources Journal*, 17(3), 68- 79.
- (8) SANTORI, G., DI NICOLA, G., MOGLIE, M. & POLONARA, F. 2012. A Review Analyzing The Industrial Biodiesel Production Practice Starting From Vegetable Oil Refining. *Applied Energy*, 92, 109-132.
- (9) ALAM, N., HUSSAINA, M., MUSHTAQA, F. & SHAHIDA, M. 2016. Production Of Biodiesel From Waste Cooking Oil By Transesterification Process Using Potassium Hydroxide As Catalyst. *International Journal Of Chemical And Environmental Engineering*, 7(4), 235-239.
- (10) NG, J.-H.; NG, H. K.; GAN, S. Advances in biodiesel fuel for application in compression ignition engines. *Clean technologies and environmental policy*, v. 12, n. 5, p. 459-493, 2010.
- (11) ESPADAFOR F. J., GARCÍA M. T., VILLANUEVA J. B. & GUTIÉRREZ J. M. 2009. The Viability Of Pure Vegetable Oil As An Alternative Fuel For Large Ships. *Transportation Research Part D: Transport And Environment*, 14(7), 461-469.
- (12) KNOTHE G. 2006. Analyzing Biodiesel: Standards And Other Methods. *Journal of the American Oil Chemists' Society*, 83, 823-833.
- (13) DEMIRBAS A. 2003. Biodiesel Fuels From Vegetable Oils Via Catalytic And Non- Catalytic Supercritical Alcohol Transesterifications And Other Methods: A Survey. *Energy Convers Manage*, 44, 2093-109.
- (14) TOSCANO, G., RIVA, G., FOPPA PEDRETTI E. & DUCA, D. 2012. Vegetable Oil And Fat Viscosity Forecast Models Based On Iodine Number And Saponification Number. *Biomass And Bioenergy*, 46, 511-516.

From 3D scanning, reverse engineering to real part production of personalized car parts

Erald Piperi^{1,*}, Ilo Bodi¹, Artan Hoxha¹
Polytechnic University of Tirana, Albania¹
epiperi@fim.edu.al

Abstract: In personalized mechanical parts production, an important issue is the functionality of the part itself in the chain of multi assembled parts. There are cases when same custom parts need to be produced by fulfilling the solution of connecting two or more elements together and creating a functional mechanical system. Personalized components that are going to feet in between two existing and functional element or systems are difficult to be produced. These difficulty in terms of mechanical solutions means, dimensional accuracy tolerances of the new part and real existing features parts. By creating these personalized middle nodes which will connect two existing elements the precision accuracy feature, like tolerances, is very important. The challenges are bigger when dealing with maximum lifespan performance while withstanding mechanical shock and minimizing fatigue destruction problems. Reducing such a challenging task the 3D scanners and reverse engineering workflow fill the gap by increasing precision and accuracy tolerances digitalization's of existing parts, creating the reverse counterpart, and reducing time to market. In our work we show a full workflow from non-contact 3D scans, reverse engineering process, for CNC aluminium milling of the digitized part in a real application sector such as modified car application. This attempt will show the benefits of active 3D scanner by surpassing some challenges such as metallic surface reflection problem that arises during scan. In the end the entire process gives to the user all the detailed technical documentation including drawings, materials, and manufacturing instructions.

Keywords: 3D SCANN, REVERSE ENGINEERING, CNC

1. Introduction

In today's globally competitive and rapidly changing production landscape, the need for agile and adaptable manufacturing processes is paramount. Reverse engineering (RE) emerges as a critical solution, offering advantages in design time reduction, rapid prototyping, and commercial efficiency. This technique is vital for reconstructing existing parts, repairing damaged ones, or innovating based on existing designs across diverse sectors such as, machinery components [1], [2], cultural heritage [3], healthcare [4], etc. By detailing the process from 3D scanning to prototype development, this research aims to illustrate the broad applicability and benefits of reverse engineering (RE), by illustrating a full workflow of the RE process particularly in producing personalized car components. Sometimes the need of personalized part by modifying existing production is difficult and time consuming by classical metrology. By addressing common challenges associated with traditional measurement and 3D data capture methods, 3D active scanners provide a more efficient, accurate, and user-friendly solution. There are many application when different scanning technology are used and it's an extended list continuing to be enlarged nowadays [5] and in Industry 4.0 related to quality assurance topics [6]. There are some important consideration attributes when choosing the 3D scanning system as:

- Speed and Efficiency,
- Accuracy and Precision,
- Ease of Use,
- Flexibility and Portability,
- Automation and Integration,
- Versatility,
- Cost.

In this work we have used the commercial active 3D structured light scanner RangeVision Spectrum © [7]. The scanner itself can be used in wide range of applications and is a high-resolution 3D scanner Table 1.

Table 1. Technical specification of the 3d scanner.

RangeVision Spectrum 3d scanner			
Scanning technology	structured light		
Resolution	1280 x 800		
FOV size, [mm]	520x400x400	320x250x220	135x100x100
Accuracy, up to [mm]	120	60	40
3D resolution, up to [mm]	0.25	0.15	0.06

2. Experiment

3D scanning

Every 3D scanning technology procedure start with the calibration of the internal and external extrinsic parameters of the scanner and the desired area wanted for digitalization. In our case the procedure is conducted using the marker calibration plate and following the software ScanCenter NG © software calibration procedure offered by the scanner package. The other steps of the workflow are showed in the Fig. 1 below.

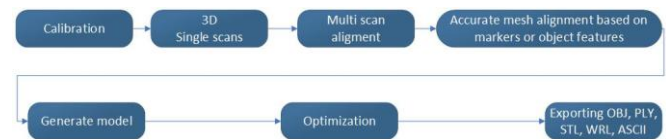


Fig. 1 Workflow principle of 3D scanner

In our case the need to cover the outer surface was obligatory (Fig. 2) to achieve high quality scans with precise dimension accuracy. As showed earlier by the author the thickness of the coating spray is also related to the user experience and in our case, it was measured as an average layer thickness of max 18 μm [8].



Fig. 2 Single 3D scan of the car part.

Car part digitalization

The main objective of the study was to produce a personalized aluminum part that will connect the two existing car parts (Part A and Part B) for modifying and updated solution for a real racing car by the final user.

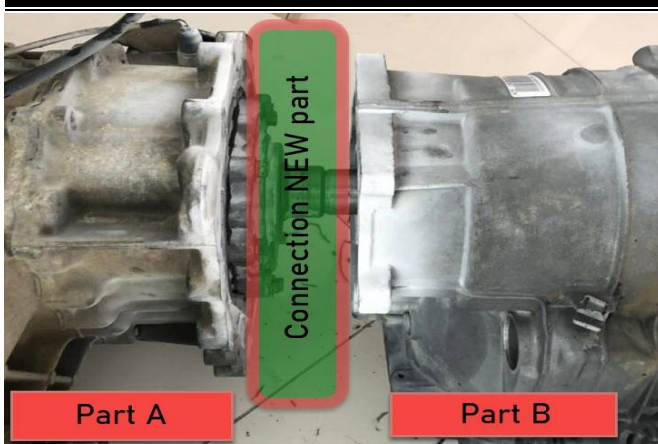


Fig. 3 New part connector functional solution for connecting part A and part B with reverse engineering process.

In fact, as it can be understood the new proposed solution will be one counterpart of the respective outer surfaces of part A and part B carried out by the reverse engineering process workflow. After capturing all the single scans of each part and fused the scans in a single 3D mesh (Fig. 4 (b)) the user can export in OBJ or STL format the files and can start the reverse engineering process.

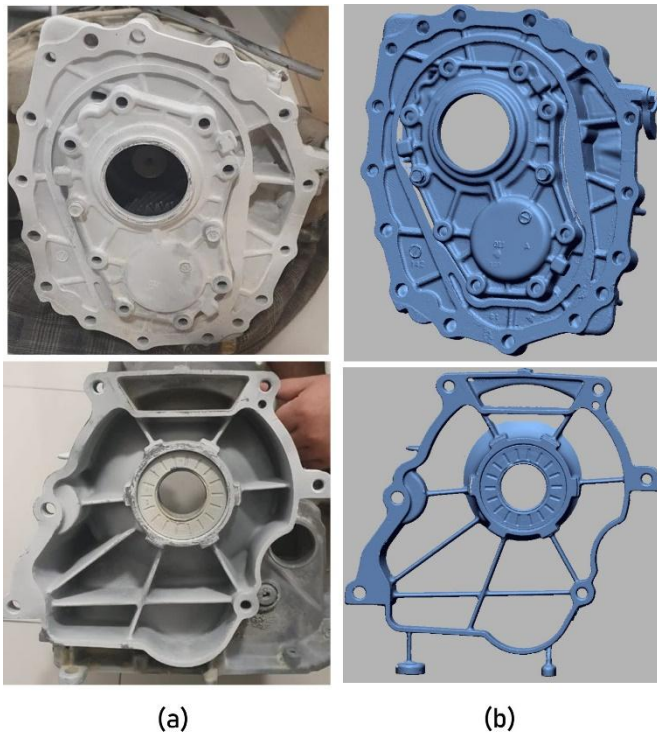


Fig. 4 Final 3D mesh (on the right (b)) file of real part A and Part B (on the left (a)).

Reverse engineering process

The RE process, also known as Reverse Engineering process, carried out by the operator before proceeding with feature object recognition shapes typically involves several key steps. The process generally starts after the physical object has been digitized into a 3D model, which is essential for further analysis and modification in various engineering applications. These steps consist in part of alignment and mesh optimization. Once the mesh is optimized (reducing polygon numbers, cleaning and refining mesh quality, smoothing process, etc.,) then the process of RE may continuing in both manual and automated directions. Taking in consideration the importance of the dimensional precision we used the manual

procedure offered by the Geomagic Design X © [9] by creating simple sketch from mesh data. As it can be seen in Fig. 5 the main features as holes, surface planes and contour shape have been maximized within a range of ± 0.1 mm.

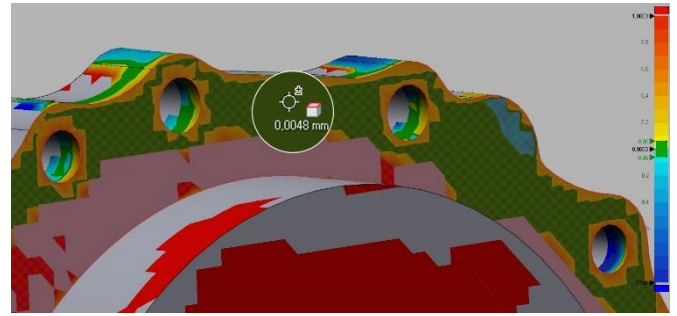


Fig. 5 RE process for feature recognitions.

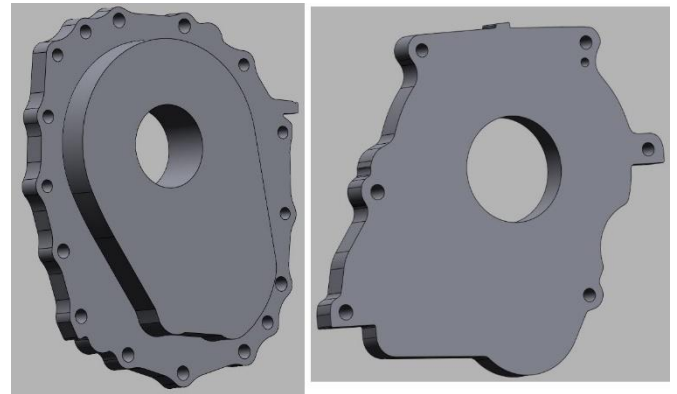


Fig. 6 Final outputs of the RE process for outside surfaces of Part A and Part B.

This process is critical because the final product that is the “counterpart” must interact with an existing product of both outer surfaces of part A and part B. The final step of the RE is to document with detailed drawings, specifications, and manufacturing documentation for part production. The file can be exported also in different formats as STEP or IGES for CNC machining.

CNC production of the new counterpart

The new counterpart will be a new functional solution for racing car by updating the transmission and increasing performance of the modified system. These counterparts will be produced by AL7075-T6 alloy by CNC machining. This is a high-strength alloy often referred to as “aircraft grade” aluminum due to its use in the aerospace industry and other high-performance applications. The material is mainly used in aviation, in the manufacturing of racing cars etc. It is not the scope of these research work to show all the steps of the CNC production workflow, but we are showing only some of the stages of produced counterpart by CNC machining in Fig. 7.

3. Results

When original designs are not available, or when a new product must be produced and must also interact with an existing component in the assembly the process of RE in part design plays a crucial role. One important role in these workflows play the 3D scanning process itself. In this work we have shown a real workflow application requested for a missing market solution for the racing cars to increase the car performance. The work showed the entire process from 3D scanning to RE process and part manufactured by CNC machining. The advanced technologies as 3D scanning are an important tool for designers and engineers to facilitate daily task faster with very accurate results. In this case the total time needed for 3d scan process and the RE procedure was about 1.5 hours and the results were very accurate depending on the complexity and size

of the part being scanned and the detail required for the reverse engineering. The produced part was assembled and very functional by completing the challenges required by the clients.

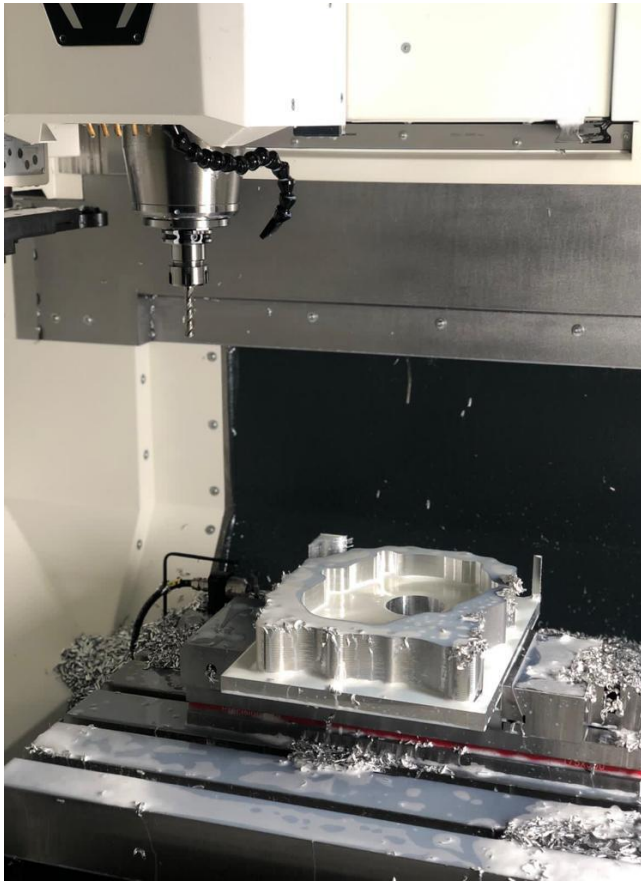


Fig. 7 CNC machining of the counterpart (up), Final counterpart produced by AL 7075-T6 (down)

4. References

- A. V. Dongaonkar and R. M. Metkar , "Reconstruction of
- 1] Damaged Parts by Integration Reverse Engineering (RE) and Rapid Prototyping (RP)," in *3D Printing and Additive Manufacturing Technologies*, Singapore, Springer, 2018, pp. 159-171.
- F. Buonomici, M. Carfagni, R. Furferi, L. Governi, A.
- 2] Lapini and Y. Volpe, "Reverse engineering of mechanical parts: A template-based approach," *ournal of Computational Design and Engineering*, vol. 5, no. 2, pp. 145-159, 2018.
- T. Segreto , . A. Bottillo, R. Teti , . L. M. Galantucci , F.
- 3] Lavecchia and . M. . B. Galantucci , "Non-contact Reverse Engineering Modeling for Additive Manufacturing of Down Scaled Cultural Artefacts," *Procedia CIRP*, vol. 62, pp. 481-486, 2017.
- M. F. Noor, F. Hasan, S. Bhardwaj and S. Hasan, "Reverse
- 4] Engineering in customization of products: Review and Case Study," in *International Conference on Humanizing Work and Work Environment (HWWE)*, India, 2017.
- D. Lazarević, B. Nedić, Ž. Šarkoćević and M. Blagojević,
- 5] "Optical inspection of cutting parts by 3D scanning," *Physica A*, 2019.
- R. Rachel HUGHES, R. HUGHES, M. A. A. RAHMAN,,
- 6] C. Lin, P. G. MAROPOULOS, . E. MOHAMAD, . M. . R. SALLEH and A. A. ABDUL RAHMAN, "Implementation of a 3D scanning system for quality inspection in manufacturing," *Proceedings of International Conference on Design and Concurrent Engineering & Manufacturing Systems Conference*, 2023.
- RangeVision. [Online]. Available:
- 7] <https://rangevision.com/en/products/spectrum/#>. [Accessed 5 March 2024].
- E. Piperi, I. Bodi, P. DOUMALIN and J. KAÇANI,
- 8] "Scanning, 3D GD&T Inspection for automotive plastic elements," in *9th International Textile Conference & 3rd International Conference on Engineering and Entrepreneurship* , 2021.
- OQTON. [Online]. Available: <https://oqton.com/geomagic-designx/>. [Accessed 6 March 2024].
- 9]

Exergy analysis of a two-cylinder steam turbine from combined cycle power plant at three operating regimes

Mrzljak Vedran¹, Anđelić Nikola¹, Baressi Šegota Sandi¹, Prpić-Oršić Jasna¹

¹Faculty of Engineering, University of Rijeka, Vukovarska 58, 51000 Rijeka, Croatia

E-mail: vedran.mrzljak@riteh.uniri.hr, nikola.andjelic@riteh.uniri.hr, sandi.segota@riteh.uniri.hr, jasna.prpic.orsic@riteh.uniri.hr

Abstract: This paper presents an exergy analysis of a two-cylinder low power steam turbine from combined cycle power plant at three operating regimes. The highest mechanical power produced in the whole turbine is 6807.24 kW in Operating regime 1. Cylinders of the observed turbine did not have the same operation dynamics in relation to produced mechanical power in all operating regimes. In each operating regime High Pressure Cylinder (HPC) has lower exergy destruction and higher exergy efficiency in comparison to Low Pressure Cylinder (LPC) due to the influence of wet steam which expands through the last LPC stages (water droplets in wet steam increases LPC exergy destruction and decreases LPC exergy efficiency). Whole turbine exergy efficiency is between 51.62% (in Operating regime 2) and 64.98% (in Operating regime 1). This range of exergy efficiencies can be expected for a low power steam turbine. An increase in the ambient temperature decreases exergy efficiency of the whole turbine and both turbine cylinders, regardless of the observed operating regime. The exergy efficiency of the LPC is low in all operating regimes, so any improvements should be based on this cylinder first.

KEYWORDS: EXERGY ANALYSIS, TWO-CYLINDER STEAM TURBINE, COMBINED CYCLE POWER PLANT, VARIOUS OPERATING REGIMES

1. Introduction

Steam turbines are nowadays constituent components of many power plants worldwide: conventional [1], nuclear [2, 3], cogeneration [4], marine [5, 6] and various others. In addition, steam turbines can also be found in the combined cycle power plants where steam is produced by using combustion gases from gas turbine or more of them [7, 8]. Utilizing heat from gas turbine(s) combustion gases in combined cycle power plants results in very high plant overall efficiency [9].

Medium and high power steam turbines are composed of several cylinders [10, 11], while low power steam turbines can also be composed of one cylinder only [12, 13]. Each turbine cylinder is usually composed of several turbine stages [14]. Steam turbine cylinders can be single flow and dual flow cylinders, while dual flow cylinders can be symmetrical and dissymmetrical [15-17].

In this paper is analyzed, from the exergy viewpoint, a two-cylinder low power steam turbine which operates in a combined cycle power plant. Whole turbine and both of her cylinders are analyzed in three operating regimes. It should be highlighted that for the observed low power steam turbine and her cylinders can be expected much lower exergy efficiencies in comparison to medium and high power steam turbines [18]. According to the obtained exergy analysis results, it will be presented guidelines for possible improvement of the observed turbine and her cylinders.

2. Description and operating characteristics of the analyzed steam turbine and her cylinders

The analyzed steam turbine is a constituent component of a combined cycle power plant [19]. General scheme of the observed steam turbine, along with operating points required for the exergy analysis at all loads are presented in Fig. 1.

The observed steam turbine is composed of two cylinders: High Pressure Cylinder (HPC) and Low Pressure Cylinder (LPC) connected to the same shaft which drives an Electric Generator (EG). HPC get steam from Heat Recovery Steam Generator (HRSG). The whole steam mass flow rate from HRSG expands through HPC (which means that HPC did not possess any steam extraction). A certain steam mass flow rate is taken between turbine cylinders and delivered to Deaerator (operating point 3, Fig. 1). As can be seen from Fig. 1, analyzed steam turbine did not possess steam re-heater between cylinders [20]. Whole steam mass flow rate which enters in the LPC (operating point 4, Fig. 1) expands through it, which means that also LPC did not possess any steam extraction). After LPC, the remaining steam mass flow rate is delivered to the Main condenser for condensation. Both cylinders of the analyzed steam turbine are single flow cylinders which means that steam expands in one direction from the cylinder inlet until the cylinder outlet.

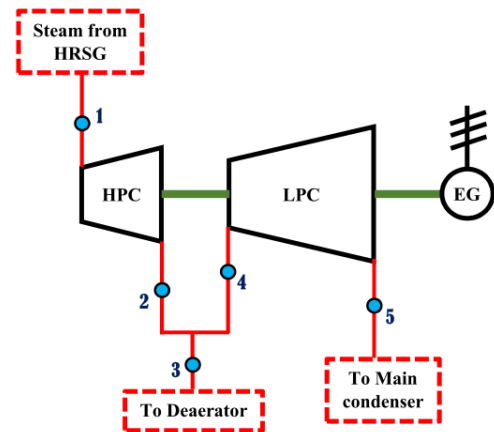


Fig. 1. General scheme of the analyzed steam turbine along with operating points necessary for the exergy analysis

Unlike isentropic analysis [21], exergy analysis of any steam turbine and her cylinders is based on the real (polytropic) steam expansion process only (along with fluid exergy flows at each cylinder inlet, outlet and in steam extractions if the extractions exist). Real (polytropic) steam expansion process through both cylinders of the analyzed turbine is presented in Fig. 2. General steam expansion process presented in Fig. 2 is valid for all three observed steam turbine operating regimes.

Fig. 2 shows that HPC operates by using superheated steam only, while the LPC partially operates by using superheated steam and partially by using wet steam (wet steam operation is evident at the last few LPC stages, under the saturation line). The operating points of real (polytropic) steam expansion process presented in Fig. 2 are marked in accordance to operating points from Fig. 1.

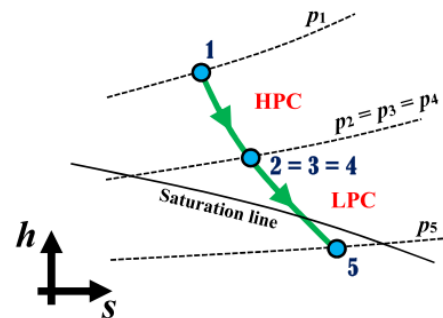


Fig. 2. Specific enthalpy/specific entropy (h/s) diagram of real (polytropic) steam expansion process throughout the analyzed turbine

3. Equations for the exergy analysis

Unlike isentropic or energy analyses, exergy analysis takes into consideration the parameters of the ambient (ambient pressure and temperature) in which observed component and component parts operate [22]. Exergy analysis enables change in the ambient parameters (especially the change in the ambient temperature) to observe the influence of such change on component and component parts exergy destructions and efficiencies [23]. Therefore, in any exergy analysis the base ambient state must be defined.

3.1. Overall exergy equations and balances

The overall exergy balance equation for any component (or component part) at steady state is [24]:

$$\dot{X} - P = \sum \dot{E}x_{out} - \sum \dot{E}x_{in} + \dot{E}x_D. \quad (1)$$

In the above overall exergy balance equation, P is real (polytropic) mechanical power, index in denotes input (inlet), index out denotes output (outlet) and index D denotes destruction. $\dot{E}x$ is a fluid total exergy flow which can be defined by the equation [25]:

$$\dot{E}x = \dot{m} \cdot \varepsilon, \quad (2)$$

while the exergy transfer by heat at the temperature T ($\dot{X}$) is defined by the equation [25]:

$$\dot{X} = \sum (1 - \frac{T_0}{T}) \cdot \dot{Q}. \quad (3)$$

In the above two equations (Eq. 2 and Eq. 3), T is a temperature, $\dot{Q}$ is the energy transfer by heat, $\dot{m}$ is fluid mass flow rate, index 0 denotes the base ambient state, while ε is fluid specific exergy which is defined, according to the literature [26], by an equation:

$$\varepsilon = (h - h_0) - T_0 \cdot (s - s_0), \quad (4)$$

where h is fluid specific enthalpy and s is fluid specific entropy. Always valid mass flow rate balance for any component (or component part) is:

$$\sum \dot{m}_{in} = \sum \dot{m}_{out}. \quad (5)$$

3.2. Equations for the exergy analysis of whole observed steam turbine and her cylinders

Equations for the exergy analysis of each cylinder (HPC and LPC) and whole observed steam turbine are defined according to the recommendations from the literature [27, 28]. Exergy analysis equations presented in this Subsection remain always the same, regardless of the observed operating regime. The indexes used in the equations are related to operating points from Fig. 1 and Fig. 2. Index WT denotes the Whole Turbine.

Real (polytropic) mechanical power

$$P_{HPC} = \dot{m}_1 \cdot (h_1 - h_2), \quad (6)$$

$$P_{LPC} = \dot{m}_4 \cdot (h_4 - h_5), \quad (7)$$

$$P_{WT} = P_{HPC} + P_{LPC}. \quad (8)$$

Exergy inlet

$$\dot{E}x_{in,HPC} = \dot{m}_1 \cdot \varepsilon_1, \quad (9)$$

$$\dot{E}x_{in,LPC} = \dot{m}_4 \cdot \varepsilon_4, \quad (10)$$

$$\dot{E}x_{in,WT} = \dot{m}_1 \cdot \varepsilon_1. \quad (11)$$

Exergy outlet

$$\dot{E}x_{out,HPC} = \dot{m}_2 \cdot \varepsilon_2 + P_{HPC}, \quad (12)$$

$$\dot{E}x_{out,LPC} = \dot{m}_5 \cdot \varepsilon_5 + P_{LPC}, \quad (13)$$

$$\dot{E}x_{out,WT} = \dot{m}_3 \cdot \varepsilon_3 + \dot{m}_5 \cdot \varepsilon_5 + P_{WT}. \quad (14)$$

Exergy destruction

$$\dot{E}x_{D,HPC} = \dot{E}x_{in,HPC} - \dot{E}x_{out,HPC}, \quad (15)$$

$$\dot{E}x_{D,LPC} = \dot{E}x_{in,LPC} - \dot{E}x_{out,LPC}, \quad (16)$$

$$\dot{E}x_{D,WT} = \dot{E}x_{in,WT} - \dot{E}x_{out,WT} = \dot{E}x_{D,HPC} + \dot{E}x_{D,LPC}. \quad (17)$$

Exergy efficiency

$$\eta_{ex,HPC} = \frac{P_{HPC}}{\dot{E}x_{in,HPC} - \dot{E}x_{out,HPC} + P_{HPC}} = \frac{P_{HPC}}{\dot{E}x_{D,HPC} + P_{HPC}}, \quad (18)$$

$$\eta_{ex,LPC} = \frac{P_{LPC}}{\dot{E}x_{in,LPC} - \dot{E}x_{out,LPC} + P_{LPC}} = \frac{P_{LPC}}{\dot{E}x_{D,LPC} + P_{LPC}}, \quad (19)$$

$$\eta_{ex,WT} = \frac{P_{WT}}{\dot{E}x_{in,WT} - \dot{E}x_{out,WT} + P_{WT}} = \frac{P_{WT}}{\dot{E}x_{D,WT} + P_{WT}}. \quad (20)$$

4. Steam operating parameters required for the exergy analysis at three operating regimes

Steam operating parameters of the analyzed steam turbine and her cylinders, at each operating regime, are found in the literature [19]. In the mentioned literature are found steam temperature, pressure and mass flow rate in each operating point and in each operating regime, other steam operating parameters are calculated by using NIST-Refprop 9.0 software [29].

The base ambient state is defined by the ambient temperature of 298.15 K and by the ambient pressure of 1 bar.

Steam operating parameters required for the whole observed turbine and her cylinders exergy analysis in Operating regime 1 are presented in Table 1, in Operating regime 2 steam operating parameters are presented in Table 2, while in Operating regime 3 steam operating parameters are presented in Table 3. From Table 1, Table 2 and Table 3 can be clearly seen that in different observed operating regimes the dominant change in each operating point is related to the change of steam temperature and steam mass flow rate.

Table 1. Steam operating parameters in Operating regime 1

Operating regime 1						
O. P.*	Temp. (K)	Pressure (bar)	Mass flow rate (kg/s)	Specific enthalpy (kJ/kg)	Specific entropy (kJ/kg·K)	Specific exergy (kJ/kg)**
1	673.15	95.0	13.94	3108.2	6.2504	1249.30
2	453.03	10.0	13.94	2777.1	6.5850	818.35
3	453.03	10.0	1.04	2777.1	6.5850	818.35
4	453.03	10.0	12.90	2777.1	6.5850	818.35
5	372.76	1.0	12.90	2607.2	7.1772	471.91

* O. P. = Operating Point (according to Fig. 1)

** Specific exergy at the base ambient state

Table 2. Steam operating parameters in Operating regime 2

Operating regime 2						
O. P.*	Temp. (K)	Pressure (bar)	Mass flow rate (kg/s)	Specific enthalpy (kJ/kg)	Specific entropy (kJ/kg·K)	Specific exergy (kJ/kg)**
1	628.15	95.0	6.33	2959.7	6.0218	1168.90
2	453.03	10.0	6.33	2777.1	6.5850	818.35
3	453.03	10.0	0.40	2777.1	6.5850	818.35
4	453.03	10.0	5.93	2777.1	6.5850	818.35
5	395.08	0.6	5.93	2561.1	7.2757	396.45

* O. P. = Operating Point (according to Fig. 1)

** Specific exergy at the base ambient state

Table 3. Steam operating parameters in Operating regime 3

Operating regime 3						
O. P.*	Temp. (K)	Pressure (bar)	Mass flow rate (kg/s)	Specific enthalpy (kJ/kg)	Specific entropy (kJ/kg·K)	Specific exergy (kJ/kg)**
1	653.15	95.0	9.71	3045.8	6.1562	1214.90
2	453.03	10.0	9.71	2777.1	6.5850	818.35
3	453.03	10.0	0.59	2777.1	6.5850	818.35
4	453.03	10.0	9.12	2777.1	6.5850	818.35
5	318.96	0.1	9.12	2488.2	7.8488	152.61

* O. P. = Operating Point (according to Fig. 1)

** Specific exergy at the base ambient state

5. Exergy analysis results and discussion

Real (polytropic) produced mechanical power in each cylinder and whole turbine for three observed operating regimes is presented in Fig. 3. Fig. 3 confirms that the observed steam turbine is a low power turbine, because the highest mechanical power produced in the whole turbine is 6807.24 kW in Operating regime 1. In Operating regime 2 whole turbine produces the lowest mechanical power equal to 2436.74 kW only.

It is also evident that cylinders of the observed turbine did not have the same operation dynamics in relation to produced mechanical power in all operating regimes. HPC produces the highest mechanical power in Operating regime 1 (equal to 4615.53 kW), while the LPC produces the highest mechanical power in Operating regime 3 (equal to 2634.77 kW). The only similarity between HPC and the LPC is that both cylinders produce the lowest mechanical power in Operating regime 2.

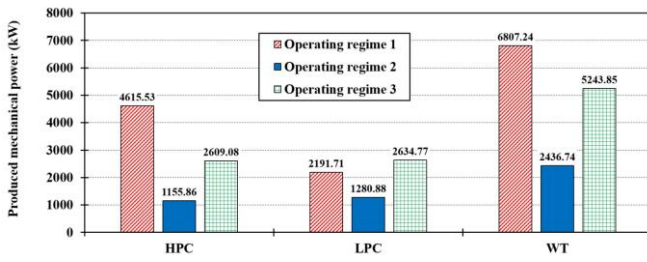


Fig. 3. Produced mechanical power in each cylinder and whole turbine for three observed operating regimes

Exergy destruction of the whole turbine and each of her cylinders at all three observed operating regimes is presented in Fig. 4.

The lowest exergy destruction in each cylinder and whole turbine is obtained in the Operating regime 2 in which whole turbine and each cylinder produce the lowest mechanical power, Fig. 3 and Fig. 4. Observation of the cylinders leads to conclusion that in each operating regime HPC has lower exergy destruction in comparison to LPC what can be explained by the influence of wet steam which expands through the last LPC stages (wet steam increases LPC exergy destruction).

Whole turbine has the highest exergy destruction in Operating regime 3 (equal to 4678.20 kW) and the lowest exergy destruction in Operating regime 2 (equal to 2284.11 kW).

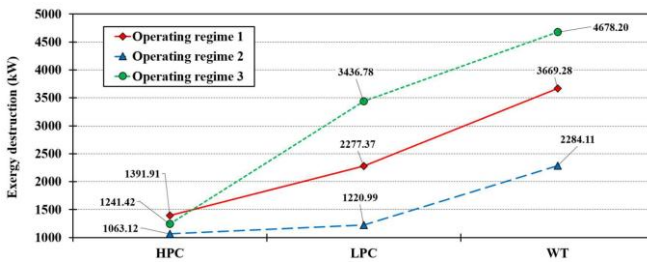


Fig. 4. Exergy destruction in each cylinder and whole turbine for three observed operating regimes

Exergy efficiency of the whole observed turbine and her cylinders at all three operating regimes is presented in Fig. 5.

Observing both cylinders, it can be concluded that HPC has higher exergy efficiency in comparison to the LPC in all operating regimes. This result related to the cylinders is expected due to much lower HPC exergy destruction in comparison to LPC, Fig. 4 and Fig. 5. The highest HPC exergy efficiency is obtained in Operating regime 1 (equal to 76.83%), while LPC has the highest exergy efficiency in Operating regime 2 (equal to 51.20%).

Whole observed turbine exergy efficiency is between 51.62% (in Operating regime 2) and 64.98% (in Operating regime 1). According to the literature, this range of exergy efficiencies can be expected for low power steam turbines [30, 31]. However, it should be highlighted that obtained exergy efficiencies (for the whole

turbine and each cylinder) are notably lower in comparison to medium and high power steam turbines from various power plants [32].

Finally, it can be concluded that exergy efficiency of LPC in all operating regimes is low and it surely can be improved.

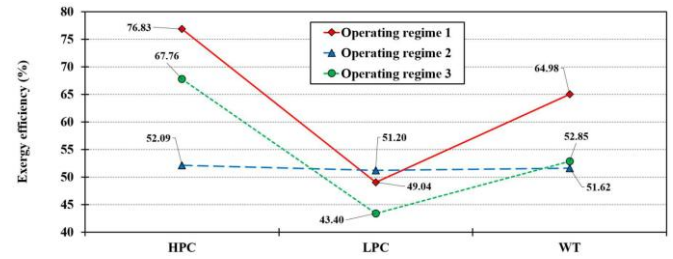


Fig. 5. Exergy efficiency in each cylinder and whole turbine for three observed operating regimes

As stated before, exergy analysis is dependable on the ambient parameters. In this paper is performed an analysis related to the ambient temperature change and its influence on the steam turbine exergy efficiency. In Fig. 6 is presented exergy efficiency of the whole turbine during the ambient temperature change for all three observed operating regimes. However, it should be highlighted that the same conclusions related to the ambient temperature change of the whole turbine are also valid for each cylinder (HPC and LPC).

From Fig. 6 is clear that an increase in the ambient temperature decreases exergy efficiency of the whole turbine, regardless of the observed operating regime (what is also valid for both turbine cylinders).

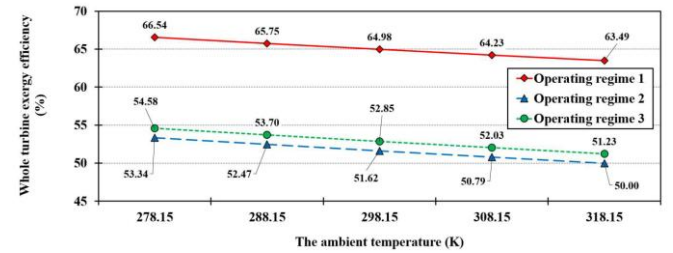


Fig. 6. Exergy efficiency of the whole turbine during the ambient temperature change for three observed operating regimes

Further analyses of the observed steam turbine and her cylinders can and will be based on the usage of many conventional [33] and artificial intelligence [34, 35] methods and processes. Results obtained in this analysis at all operating regimes will be used as a baseline in further analyses.

6. Conclusions

In this paper is performed exergy analysis of a two-cylinder low power steam turbine from combined cycle power plant at three operating regimes. The most important conclusions are:

- The highest mechanical power produced in the whole turbine is 6807.24 kW in Operating regime 1.
- Cylinders of the observed turbine did not have the same operation dynamics in relation to produced mechanical power in all operating regimes.
- In each operating regime HPC has lower exergy destruction and higher exergy efficiency in comparison to LPC what can be explained by the influence of wet steam which expands through the last LPC stages (water droplets increases LPC exergy destruction and decreases LPC exergy efficiency).
- Whole observed turbine exergy efficiency is between 51.62% (in Operating regime 2) and 64.98% (in Operating regime 1). This range of exergy efficiencies, according to the literature, can be expected for a low power steam turbine.
- Analysis shows that exergy efficiency of LPC is low in all operating regimes, so any improvements should be based on this cylinder first.

- An increase in the ambient temperature decreases exergy efficiency of the whole turbine and both turbine cylinders, regardless of the observed operating regime.

7. Acknowledgment

This research has been fully supported by the Croatian Science Foundation under the project IP-2022-10-2821.

8. References

- [1] Khanmohammadi, S., Azimian, A. R., & Khanmohammadi, S. (2013). Exergy and exergo-economic evaluation of Isfahan steam power plant. *International Journal of Exergy*, 12(2), 249-272. (doi:10.1504/IJEX.2013.053386)
- [2] Ebrahimgol, H., Aghaie, M., Zolfaghari, A., & Naserbegi, A. (2020). A novel approach in exergy optimization of a WWER1000 nuclear power plant using whale optimization algorithm. *Annals of Nuclear Energy*, 145, 107540. (doi:10.1016/j.anucene.2020.107540)
- [3] Talebi, S., & Norouzi, N. (2020). Entropy and exergy analysis and optimization of the VVER nuclear power plant with a capacity of 1000 MW using the firefly optimization algorithm. *Nuclear Engineering and Technology*, 52(12), 2928-2938. (doi:10.1016/j.net.2020.05.011)
- [4] Abbaspour, H., Ehyaei, M. A., Ahmadi, A., Panahi, M., Abdalisousan, A., & Mirzohosseini, A. (2021). Energy, exergy, economic, exergoenvironmental and environmental (5E) analyses of the cogeneration plant to produce electrical power and urea. *Energy Conversion and Management*, 235, 113951. (doi:10.1016/j.enconman.2021.113951)
- [5] Mrzljak, V., Poljak, I., & Medica-Viola, V. (2017). Dual fuel consumption and efficiency of marine steam generators for the propulsion of LNG carrier. *Applied Thermal Engineering*, 119, 331-346. (doi:10.1016/j.applthermaleng.2017.03.078)
- [6] Baressi Šegota, S., Mrzljak, V., Anđelić, N., Poljak, I., & Car, Z. (2023). Use of Synthetic Data in Maritime Applications for the Problem of Steam Turbine Exergy Analysis. *Journal of Marine Science and Engineering*, 11(8), 1595. (doi:10.3390/jmse11081595)
- [7] Najjar, Y. S., & Manaserh, Y. M. A. (2019). Aligning combined cycle power plant performance with field measurements. *Arabian Journal for Science and Engineering*, 44, 1657-1669. (doi:10.1007/s13369-018-3615-2)
- [8] Lorencin, I., Anđelić, N., Mrzljak, V., & Car, Z. (2019). Genetic algorithm approach to design of multi-layer perceptron for combined cycle power plant electrical power output estimation. *Energies*, 12(22), 4352. (doi:10.3390/en12224352)
- [9] Mohammed, M. K., Awad, O. I., Rahman, M. M., Najafi, G., Basrawi, F., Abd Alla, A. N., & Mamat, R. (2017). The optimum performance of the combined cycle power plant: A comprehensive review. *Renewable and Sustainable Energy Reviews*, 79, 459-474. (doi:10.1016/j.rser.2017.05.060)
- [10] Leyzerovich, A. S. (2021). *Steam turbines for modern fossil-fuel power plants*. River Publishers.
- [11] Elhelw, M., Al Dahma, K. S., & el Hamid Attia, A. (2019). Utilizing exergy analysis in studying the performance of steam power plant at two different operation mode. *Applied Thermal Engineering*, 150, 285-293. (doi:10.1016/j.applthermaleng.2019.01.003)
- [12] Behrendt, C., & Stoyanov, R. (2018). Operational characteristic of selected marine turbounits powered by steam from auxiliary oil-fired boilers. *New Trends in Production Engineering*, 1(1), 495-501. (doi:10.2478/ntpe-2018-0061)
- [13] Mrzljak, V., Prpić-Oršić, J., & Poljak, I. (2018). Energy power losses and efficiency of low power steam turbine for the main feed water pump drive in the marine steam propulsion system. *Pomorski zbornik*, 54(1), 37-51. (doi:10.18048/2018.54.03)
- [14] Drbal, L., Westra, K., & Boston, P. (Eds.). (2012). *Power plant engineering*. Springer Science & Business Media.
- [15] Vedran, M., Ivan, L., Nikola, A., & Zlatan, C. (2020). Dual-flow dissymmetrical low pressure steam turbine energy analysis-comparison of both turbine cylinders. *Machines. Technologies. Materials.*, 14(8), 336-339.
- [16] Tanuma, T. (Ed.). (2017). *Advances in steam turbines for modern power plants*. Woodhead Publishing.
- [17] Poljak, I., Mrzljak, V., Gospić, I., & Glavan, I. (2021). The Change in Low Power Steam Turbine Operating Parameters During Extractions Opening/Closing. *Pomorski zbornik*, 61(1), 57-78. (doi:10.18048/2021.61.05)
- [18] Elčić, Z. (1995). *Steam turbines*. ABB, Karlovac, National and University Library Zagreb.
- [19] Stradioto, D. A., Seelig, M. F., & Schneider, P. S. (2015). Reprint of: Performance analysis of a CCGT power plant integrated to a LNG regasification process. *Journal of Natural Gas Science and Engineering*, 27, 18-22. (doi:10.1016/j.jngse.2015.06.009)
- [20] Mrzljak, V., Šegota, S. B., Kocijel, L., & Prpić-Oršić, J. (2020). Energy (isentropic) analysis of three-cylinder steam turbine with re-heating. *Innovations*, 8(1), 37-40.
- [21] Poljak, I., & Mrzljak, V. (2023). Thermodynamic Analysis and Comparison of Two Marine Steam Propulsion Turbines. *NAŠE MORE: znanstveni časopis za more i pomorstvo*, 70(2), 0-0. (doi:10.17818/NM/2023/2.2)
- [22] Szargut, J. (2005). *Exergy method: technical and ecological applications* (Vol. 18). WIT press.
- [23] Kopac, M., & Hilalci, A. (2007). Effect of ambient temperature on the efficiency of the regenerative and reheat Çatalağzı power plant in Turkey. *Applied Thermal Engineering*, 27(8-9), 1377-1385. (doi:10.1016/j.applthermaleng.2006.10.029)
- [24] Kanoğlu, M., Çengel, Y. A., & Dinçer, İ. (2012). *Efficiency evaluation of energy systems*. Springer Science & Business Media.
- [25] Mrzljak, V., Poljak, I., Prpić-Oršić, J., & Jelić, M. (2020). Exergy analysis of marine waste heat recovery CO₂ closed-cycle gas turbine system. *Pomorstvo*, 34(2), 309-322. (doi:10.31217/p.34.2.12)
- [26] Tan, H., Shan, S., Nie, Y., & Zhao, Q. (2018). A new boil-off gas re-liquefaction system for LNG carriers based on dual mixed refrigerant cycle. *Cryogenics*, 92, 84-92. (doi:10.1016/j.cryogenics.2018.04.009)
- [27] Erdem, H. H., Akkaya, A. V., Cetin, B., Dagdas, A., Sevilgen, S. H., Sahin, B., ... & Atas, S. (2009). Comparative energetic and exergetic performance analyses for coal-fired thermal power plants in Turkey. *International Journal of Thermal Sciences*, 48(11), 2179-2186. (doi:10.1016/j.ijthermalsci.2009.03.007)
- [28] Koroglu, T., & Sogut, O. S. (2018). Conventional and advanced exergy analyses of a marine steam power plant. *Energy*, 163, 392-403. (doi:10.1016/j.energy.2018.08.119)
- [29] Lemmon, E. W., Huber, M. L., & McLinden, M. O. (2010). *NIST Standard Reference Database 23, Reference Fluid Thermodynamic and Transport Properties (REFPROP), version 9.0*, National Institute of Standards and Technology. R1234yf. fld file dated December, 22, 2010.
- [30] Mrzljak, V., Poljak, I., & Mrakovčić, T. (2017). Energy and exergy analysis of the turbo-generators and steam turbine for the main feed water pump drive on LNG carrier. *Energy conversion and management*, 140, 307-323. (doi:10.1016/j.enconman.2017.03.007)
- [31] Mrzljak, V., & Poljak, I. (2019). Energy analysis of main propulsion steam turbine from conventional LNG carrier at three different loads. *NAŠE MORE: znanstveni časopis za more i pomorstvo*, 66(1), 10-18. (doi:10.17818/NM/2019/1.2)
- [32] Mrzljak, V., Jelić, M., Poljak, I., & Prpić-Oršić, J. (2023). Analysis and Comparison of Main Steam Turbines from Four Different Thermal Power Plants. *Pomorstvo*, 37(1), 58-74. (doi:10.31217/p.37.1.6)
- [33] Mrzljak, V., Poljak, I., Jelić, M., & Prpić-Oršić, J. (2023). Thermodynamic Analysis and Improvement Potential of Helium Closed Cycle Gas Turbine Power Plant at Four Loads. *Energies*, 16(15), 5589. (doi:10.3390/en16155589)
- [34] Baressi Šegota, S., Lorencin, I., Anđelić, N., Mrzljak, V., & Car, Z. (2020). Improvement of marine steam turbine conventional exergy analysis by neural network application. *Journal of Marine Science and Engineering*, 8(11), 884. (doi:10.3390/jmse8110884)
- [35] Wilding, P. R., Murray, N. R., & Memmott, M. J. (2020). The use of multi-objective optimization to improve the design process of nuclear power plant systems. *Annals of Nuclear Energy*, 137, 107079. (doi:10.1016/j.anucene.2019.107079)

Pilot Project for an Integrated Complex for clean soil-based production, harvest and delivery (logistic operations) of Iceberg lettuce

Ivan Gramatikov¹, Teodor Georgiev², Krasimir Kumchev³, Kostadin Kostadinov⁴, Marin Midilev⁵

Freelancer and Association Agrorobotika founder¹, Robco Smart², Chairman at the Association of Agrorobotika³, Association Agrorobotika member, VSI Plovdiv⁴, Representative of the Cluster for Electric Mobility⁵

Abstract: In order to address multiple UN Sustainable Development Goals, we aim to revolutionize iceberg lettuce harvesting with intelligent robots. Our team, comprised of agricultural producers and engineers, has utilized TRIZ methodology, brainstorming, and other techniques to develop an agricultural technology solution: an Integrated Complex that automates harvesting and loading of produce onto transportation vehicles. Our design leverages state-of-the-art deep computer vision – a form of AI that analyzes camera images – to precisely locate mature lettuce heads. The robot identifies optimal grasping points for harvesting and carefully places the heads in a designated transport area. With deep learning at its core, our robots will optimize yield, reduce waste and costs, ushering in a new era of intelligent agriculture.

KEYWORDS: LETTUCE, HARVEST, ROBOT, COMPUTER VISION, LOGISTICS, INTEGRATED COMPLEX

1. Introduction

The amount of food waste, the issues of lacking workforce in the field of agriculture (especially in countries like Bulgaria and in periods such as the COVID-19 pandemic), the number of starving people, etc. are among the food-related issues which the United Nations and the European Union and individual States are trying to resolve.

There are specific goals and strategies such as the Sustainable Development Goal 12.3, which aims to (halve per capita food waste at the retail and consumer level by 2030, and reduce food losses along the food production and supply chains) [1,2].

According to the Food and Agriculture Organization of the United Nations, (roughly one-third of food produced for human consumption is lost or wasted globally, which amounts to about 1.3 billion tons per year. This inevitably also means that huge amounts of the resources used in food production are used in vain, and that the greenhouse gas emissions caused by production of food that gets lost or wasted are also emissions in vain. Food is lost or wasted throughout the supply chain, from initial agricultural production down to final household consumption) [3].

The European Commission itself (has already carried out important steps to prevent food losses and waste, in line with the actions put forward under the first [Circular Economy Action Plan](#), adopted in 2015) [4].

To these aims, the European Commission has suggested certain actions such as reforms aimed at increasing the awareness and usage of the food labels, the date of expiry, easing of the process of donating of foods to food banks and charity organizations, support of the initiatives for the increased usage of foods which have not been sold and other products which can be used as a resource in the production of animal feed [5].

During the COVID pandemic, a lot of the agricultural produce was left on the field, spoiled, due to the restrictions on the working conditions and the lack of workforce. The team, issuing this report, has witnessed first-hand the sheer number of rows of iceberg lettuce in Trilistnik, Plovdiv district, left unusable.

The resolution of such food related problems requires a complex cooperation of a country's administrative apparatus, producers, distributors, scientific institutes, associations, NGOs and consumers. Our team is mostly concentrated on the technological possibilities for solving these issues as it is our firm conviction that they can be addressed by the robotization of the agricultural lifecycle or at least, for the time being, the automation of the harvesting and delivery parts of the process. To this end, we suggest

the development of an Integrated Complex of robots, which are based on a system of simple and existing building blocks.

2. Prerequisites and means for solving the problem

2.1 Necessary preconditions

The agricultural producers' role in the team are is vital, as they give the necessary data depending on the type or sort of plant, which will be worked on by the robot – sizes of the plant, physical characteristics of the ripe produce, the planting schematic data, the timetable of all necessary steps in the lifecycle of the plant's development.

For the whole process of research on iceberg lettuce robotic harvesting, the necessary data are being obtained from the rich repository of technical expertise of the agricultural producer Petar Kumchev, with the place of the production of lettuce being the village of Trilistnik, Plovdiv district.

Further requirements include the creation of an industrial park (type C), an agricultural robotics laboratory, equipped with a FabLab and a team of mechatronic engineers and software developers

2.2 Logistic considerations

2.2.1 Energy Expenditure

One can observe a very clear rise in the research and development of robots for the harvest of fruits and vegetables all over the planet. And there are multiple reasons for this.

A discussion on the energetic requirements of food production, including our future efforts, is in order, so as to prove the above point.

According to economist Manfred Max-Neef of the CEPAAUR-Dag Hammarskjöld Foundation:

(In Mexico, according to information provided by the Xochkalli Foundation, 19,000 kcal have to be used in order to put 2,200 kcal of food on the table. From another angle, the amount of energy consumed in transporting foodstuffs in Mexico is almost equal to the total energy required by the primary sector for food production. The fact that such situations are considered to be positive is, undoubtedly, a conceptual aberration) [6]

Other than the human energetic expenditure, it should also be noted that human workers normally spend all of this energy in an 8, 10 or even 12-hour workdays, however they can realistically be productive for about 6 of those 10 hours, as opposed to a robot, which can productively work for almost the entire day.

Larger fields of about 100-500 decares and more (100-500 000 square meters) that require crop care, harvesting and logistical post-processing and delivery operations would be a good example of cumbersome, repetitive and error-prone labor in the human worker case. Such a scale would be handled better by a fleet of robots, including drones/UAVs, as they would employ much bigger precision, starting from the process of precision agriculture, e.g. recording and evaluating the problematic areas on the field (such as in Fig. 1) [7], the predictive algorithms and AI architectures for the volumes of produce, which the farmer will obtain, etc. [8].

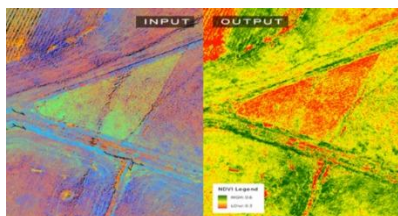


Fig. 1 Remote Sensing spectral images by drone to evaluate problematic areas⁷

Another useful feature would be evaluating the soluble solids (sugars) [9] and other contents present in the target plants for an overall picture, allowing estimation of the crop, the season and possible profits. Yet another one, which the team is planning, is a system for protection from hail.

2.2.2 Delivery and design considerations of the Integrated Complex

The logistics process after and other than the harvesting will now be discussed. Once a full crate has been filled by the harvesting robot, it is moved to a preset location, freeing up space for an empty crate. When there is a sufficient number of ready to deliver crates, they are loaded onto the transporting platform, which has already carried another set of empty crates ready to fill up. The transporting platform delivers the produce, which has been picked to an aggregating platform. The transporting platform carrying the produce is being unloaded and loaded with the necessary number of empty crates. The produce is carried then from the aggregating platform to the cooler by electric vehicles. On the way back, the said electric vehicles carry empty crates. Then the harvested produce is being processed (packaged) and is labeled with information for traceability of the quality, nutrients, pesticide content, etc.

On the basis of the information on the label and the overall aggregation of produce-related data, as discussed above, a subscription-based model for the delivery of specific products could be established, based on specific dietary, energy and health-related considerations (such as receiving fresh produce, which has been picked from the plant for up to 72 hours previously).

The aggregating platform provides the main energy for the 24-hour work cycles of the Integrated Complex by photovoltaics and air. It is situated outside of the crop area.

The electric vehicle is going to be supplemented with a robotic loading and unloading system and its task is to deliver empty storage crates to the robotic harvester mobile platform and carrying of the picked produce in the harvesting period. It may be used for carrying of materials during the rest of the production period.

An agrivoltaics system could be installed, which ensures the necessary electrical supply to the robotic harvester and the electric loading vehicle alike.

The delivery part of the logistics can be done to an end consumer, including shops, including fully automated ones, restaurants with or without robotic cooking systems (and robotic servers).

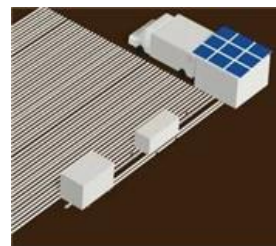


Fig 2 A schematic representation of a part of the Integrated Complex

3. Solution of the examined problem

3.1 Overall organization of the harvesting robot unit

The robots in the Integrated Complex, which will be used for plant care and harvesting will feature different (by function) types of attachable modules. Such modules could include: surface processing of the soil, soil sterilization, creation of plant beds and grooves, creation of a mulching and drip irrigation system, planting of seedling (growing of thinned out or pricked out seedling by means of robotic technology is being considered by our team as future work, as well).

There will also a system for grasping the lettuce and a system for removing the excess leaves prior to packaging so as to comply with the standards of the supermarket chains.

The changeable modules will have an application in the Integrated Complex system of machines for the lifecycle of the iceberg lettuce enterprise as a means to process said lettuce in outdoor conditions. It will also be able to perform in indoor conditions such as greenhouses, soilless (hydroponic and aeroponic) and vertical growing setups, some of which are situated in abandoned industrial buildings and rooftops (such as Gotham Greens, BrightFarms, Sky Vegetables and Brooklyn Grange in the US and the Lufa Farms in Canada). Naturally, due to the fact that the indoor growing methods rely on an artificial environment, the factors that influence the plant growth are less and the harvesting period for a robotic harvesting machine can be 24 hours a day, however the taste of an outdoor grown plant is better than an indoor grown one.

Figure 3 below presents an initial schematic 3D model for the prototype of the harvesting robot the team has created.



Fig. 3 Basic schematic initial prototype of a robotic harvester.

3.2 Sensors and perception

The harvesting robot module will feature the following main input or sensors: GPS, cameras for machine vision, an IMU, encoders and a laser scanner, etc., all of which will be gathering telemetry data for further processing.

For the machine vision, a picture dataset will be created, featuring all necessary aspects – photos of iceberg lettuce will be made in different conditions – lighting, occlusion, deformation, different scales, viewpoints, similarly-looking objects and one of the most

essential ones – background noise and clutter (which is most definitely going to be present in an outdoor setting with leaves of lettuce involved, weeds, etc.). Different additional techniques will be used to dramatically boost the accuracy of the machine vision model, such as data augmentation, transfer learning, etc. Annotation of the gathered and augmented dataset will be done and with the help of frameworks such as Detectron2, one can further improve the process of labeling by crafting (or reusing existing) semi-automatic annotation features [10].

Convolutional Neural Networks such as the Mask R-CNN [11] architecture will be used for multiple object detection (instance segmentation) and size measurement/evaluation in order to determine if the iceberg lettuce is ready for harvesting or not. This readiness will be a trigger for different operations, mostly the harvesting sequence if ready – including localization in space, grasping, cutting, lifting, placing in the respective container for transportation as discussed, etc.

An important fact to point out is that the research phase of such a robot clearly presents the opportunity to reuse certain concepts, technologies and implementations for other (especially similar) plants. This achieves speeding up the time for research, development and production of robots for other plant cultures as well as reduces their cost.

As noted above, the telemetry data gathered from the sensors, calculated, etc. can be analyzed and used for improvement and data pipelines can be utilized for that end, as well as for the deep computer vision model MLOPS lifecycle [12] (from dataset gathering to retraining for continuous model improvement).

While the annotation, experiment handling, deep computer vision model training, etc. can be done on a cloud platform (such as GCP, AWS, Azure, FloydHub, Paperspace, etc.), the model will have to fit on an embedded processing platform with AI acceleration capabilities such as the NVIDIA Jetson Xavier AGX, Orin, etc., where one could start with a prototype device at the experimental phases and then move on to professional carrier boards for real world model deployment on an embedded system (in this case, an agricultural robot).

4. Results and discussion

The integrated complex for the growing/care, harvesting and delivery to an end consumer gives a practical solution to the requirement of the amount of realized production being the basis of subsidies for the EU-based agricultural producers.

The harvesting robot is in its very infancy – a lot of work is yet to be done on it for the fruition of the first robot in a fleet that will automatically and optimally take care of the lifecycle of multiple types of plant produce due to its planned reusability and multifunctionality.

5. Conclusion:

The presented report attempts to lay out the principle of the Integrated Complex system, its most notable characteristics and uses.

6. References:

- [1]https://food.ec.europa.eu/safety/food-waste/eu-actions-against-food-waste_en
- [2]https://sdgs.un.org/goals/goal12#targets_and_indicators
- [3]<https://www.fao.org/4/mb060e/mb060e00.pdf>
- [4]https://food.ec.europa.eu/safety/food-waste/eu-actions-against-food-waste/eu-platform-food-losses-and-food-waste_en
- [5]https://wwwx.mzh.government.bg/media/filer_public/2021/10/15/natsionalna_programa_za_predotvrativane_i_namaliavane_na_zagubata_na_khrani.docx
- [6] King, Alexander, and Bertrand Schneider. The First Global Revolution. New York: Pantheon Books, 1991, 40p.
- [7] Meivel, S., Maheswari, S. Remote Sensing Analysis of Agricultural Drone. J Indian Soc Remote Sens 49, 689–701 (2021). <https://doi.org/10.1007/s12524-020-01244-y>
- [8] Rehash, Rushmi & Pavitra, A. Rehash Rushmi & Vadhani, Madhu. (2022). A SURVEY ON HARVEST PREDICTION USING MACHINE LEARNING AND INTERNET OF THINGS. 10. 85-87. 10.33144/23478578/IJCST-V10I3P16.
- [9]<https://www.sciencedirect.com/science/article/abs/pii/S1350449523000993>
- [10] Guillermo, Marielet & Billones, Robert Kerwin & Bandala, Argel & Vicerra, Ryan & Sybingco, Edwin & Dadios, Elmer & Fillone, Alexis. (2020). Implementation of Automated Annotation through Mask RCNN Object Detection model in CVAT using AWS EC2 Instance. 708-713. 10.1109/TENCON50793.2020.9293906.
- [11] Sumaiya Islam, Md Nasim Reza, Milon Chowdhury, Shahriar Ahmed, Kyu-Ho Lee, Mohammad Ali, Yeon Jin Cho, Dong Hee Noh, Sun-OK Chung, Detection and segmentation of lettuce seedlings from seedling-growing tray imagery using an improved mask R-CNN method, Smart Agricultural Technology, Volume 8, 2024, 100455, ISSN 2772-3755 <https://doi.org/10.1016/j.atech.2024.100455>.
- [12] <https://easyflow.tech/mlops-for-computer-vision/>

Implementing predictive analysis using self-learning digital twins and image analysis with GPT-4 turbo with vision for inspection and repair of construction

Mihail Mateev

Department of Computer-Aided Engineering – UACEG – Sofia, Bulgaria
mikeamm_fce@uacg.bg

Abstract: Nowadays, many structures should be inspected, analyzed, and repaired. This is a complex and expensive process that also includes predictive analytics to prevent possible construction failures.

One of the most used predictive analytics applications involves extracting necessary metadata from images and videos to evaluate the condition of real-world systems and recommend measures to sustain these systems. Image analysis is not a new concept - many solutions have been used for several decades.

The current paper mainly focuses on OpenAI-based capabilities to implement Image Analysis and Cognitive Digital Twins and proposes faster, cheaper implementation and more adaptive approaches to offering predictive analysis for constructions.

ChatGPT (Chat Generative Pre-Trained Transformer) is one of the trending technologies in modern Artificial Intelligence (AI), and experts in this area expect to have a very high impact on the industry shortly.

One of the latest versions - GPT-4 Turbo with Vision, developed by OpenAI, is a significant multimodal model (LMM) capable of interpreting images and providing text-based answers to queries regarding those images. It combines capabilities in natural language processing and visual comprehension.

The proposed approach considers using OpenAI LLM and Digital Twins for three different aspects of predictive analysis for Construction: image analysis, case decomposition, and creation of self-adaptive models to find possible trends to compromise structures and offer preventive actions. This research compares traditional methods for inspection and repair of Construction, including the time required for predictive analysis, the correctness of the proposed actions, and the cost of the methodology.

KEYWORDS: OPENAI, CHATGPT, GENERATIVE AI, COMPUTER VISION, GPT-4-TURBO, PREDICTIVE ANALYSIS IOT,

1. Introduction

Computer Vision is a branch of Artificial Intelligence (AI) that uses machine learning, neural networks, and other AI-related technologies to get specific, relevant information from images, videos, and other inputs. The computer Vision method can be applied to various kinds of analysis:

- Descriptive analysis
- Diagnostic analysis
- Predictive analysis
- Prescriptive analysis

Descriptive analysis provides data analysis that helps to describe the current status of a specific case. This analysis can describe, show, and summarize data, generating input for other types of analysis. Descriptive analysis is the best fit for Computer Vision, which describes the meaning of a specific image and/or video. It can be implemented on top of different technologies, but extracting metadata from different types of digital content is very easy to handle using AI-related technologies.

The diagnostic analysis is analyzing the root causes of something that happens (for example, construction failure). Descriptive metadata from images and videos can be critical in explaining why something (events, behaviors, outcomes) happened. It works together with descriptive analysis to get information describing the case and find the main reason for having a specific situation.

Predictive analysis mainly focuses on forecasting what will happen, using different data for input. This kind of analysis can be empowered by AI to have a pure analytic solution (without artificial intelligence elements) or to use a mixed approach as part of the solution. Predictive analysis can be implemented using different LLM – GPT, DALL-E, DAVINCI, and WHISPER. Computer Vision usually is important to extract the information needed as input for predictive analysis, and the prediction is done with some of the appropriate LLMs analyzing the metadata generated from the Computer Vision.

Prescriptive analysis gives recommendations about future actions based on results from predictive analysis. It uses statistical algorithms, machine learning technologies, and other AI-related

solutions to provide the best recommendations for the next steps in a specific case.

Generative AI (Gen AI) is not only one of the latest trends in the modern IT industry and Artificial intelligence but also a very powerful approach to solving a broad spectrum of cases. This concept allows for the generation of content based on different criteria. Nowadays, all prominent software vendors work on solutions for GenAI based on large-language models (LLM) and small-language models (SLM).

The ongoing research emphasizes the use of Generative AI and GPT models, accessible through the OpenAI API. OpenAI, a relatively new artificial intelligence research lab and company, was established in 2015 by tech visionaries such as Elon Musk and Sam Altman. OpenAI conducts research across various AI domains, including natural language processing (NLP), computer vision, reinforcement learning, and robotics.

This research is based on an analysis done with OpenAI, especially ChatGPT 4 Turbo with Vision, which is implemented on top of Microsoft Azure as SaaS. The prototype also uses other OpenAI models to optimize the cost and performance of predictive analytics solutions.

Generative AI platforms offer two different concepts of models and corresponding models for integration:

- Conversational mode (chat models)
- Text Completion mode (completion models)

A completion model attempts to extend the user's prompt with relevant text. Completion models do not have conversational capabilities. We send a single request and receive a response, and that's the end of the interaction.

Chat models are optimized for engaging in conversations. While they technically function as completion models, they can tell apart who said what, leading to a more natural conversation flow.

Both modes are essential for analysis. The conversational mode allows interaction and adjustment of requirements using chat agents, and the completion mode offers a complete analysis in one step with a clear description. Completions provide easier options to automate business processes. Completions are also used in ChatGPT-4 Turbo with Vision for Image analysis.

Table 1. below provides a sample comparison between chat and completion models.

Table 1. OpenAI Foundation GPT models

	Completion	Chat
Endpoint	/v1/completion	/v1/chat/completion
Models	davinci curie babbage ada text-davinci-003 text-curie-001 text-babbage-001 text-ada-001	All "gpt-4" models All "gpt-3.5-turbo" models

Completion vs Chat models

Some OpenAI LLM models work with chat and completion modes exposed via the corresponding API, but some models support only one of the modes.

The Fig. 1 shows the high-level taxonomy of OpenAI LLM models:

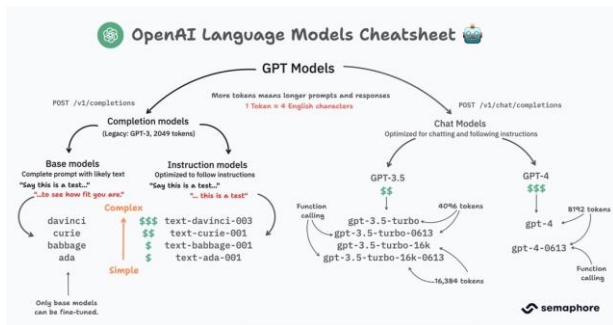


Fig. 1 Cheatsheet with the models offered by OpenAI [1]

At the end of 2023, it was released in preview GPT-4-Turbo with Vision, a GPT model that analyzes text and images and works in chat and completion modes.

ChatGPT-4 Turbo with Vision (ChatGPT 4 TV) provides high-speed and accurate analysis of images and video, allowing the analysis of digital content (descriptive and diagnostic analysis) as the first step of more complex predictive and prescriptive analysis. OpenAI LLMs allow for several deep and complex inquiries related to possible system failures.

LLM models such as ChatGPT 4 TV allow all steps to be implemented using the same model (suitable for prototyping). Still, for production, the most appropriate LLM for each step is usually used regarding cost and performance optimization.

The current research is focused on creating a reference framework to implement with a short time to market solutions, able to analyze cases from images and video and provide predictive and predictive analysis to implement predictive Maintenance and decrease the overall maintenance cost for different structures in the construction industry.

The overall schema explaining relations between different types of analysis using Computer Vision and Open AI is demonstrated in Fig.2

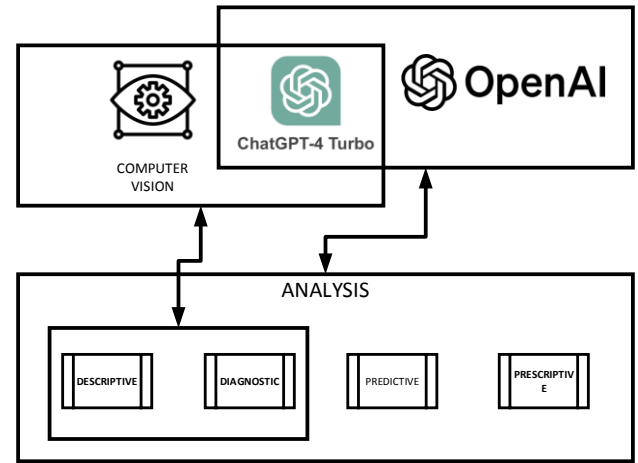


Fig. 2. Implementation of different types of analysis using Computer Vision, Open AI – general dependencies

The next chapter explains the prerequisites to design and implement the PoC used for the experimental research.

2. Prerequisites and means for solving the problem

This research is based on several pillars as a prerequisite:

- General knowledge of OpenAI and relevant LLM that can be used for predictive analytics (ChatGPT-4-TV) in case there should be analyzed digital assets such as images and videos
- Microsoft Azure subscription. The PoC is based on the Microsoft Azure cloud, which is an advantage because OpenAI is hosted in Azure.

This paper covers the following areas:

- Options for the implementation of Computer Vision analysis based on OpenAI GPT models (ChatGPT-4-TV)
- Design and implementation of Predictive Analysis based on Computer Vision using OpenAI
- Democratization of the implementation of flows, providing Predictive Analysis with OpenAI and Power Automate (Low Code/No Code)
- Analyze the results (preciseness and correctness) based on sample data (photos of structures and construction schema).
- Optimization of the cost and performance

2.1 Areas of Implementation of ChatGPT in Construction

The Construction has many options for the implementation of Generative AI. Some of the most popular domains are:

- Project Planning and Design
- Troubleshooting and Maintenance
- Resource Management
- Virtual Assistants for Clients (support)
- Training and Protocols
- Smart Building Interfaces
- Documentation and Reporting
- Energy Efficiency and Sustainability
- Troubleshooting and Maintenance (predictive and regular).

Generative AI in the context of predictive analysis can use all these domains, but current research is focused on Troubleshooting and Maintenance, analyzing possible failures of structures.

3. Solution of the examined problem

The main focus is on how to solve similar cases for Predictive Analysis for Construction using OpenAI with Computer Vision. Generative AI needs a very well-described case to generate a valuable answer. In the experimental solutions,

3.1 Implementation of Analysis based on Computer Vision and OpenAI/ChatGPT

Implementation of a Computer Vision Concept:

For image analysis, the experimental environment uses ChatGPT 4-Turbo with Vision (ChatGPT-4-TV). This model has the following specifics:

- Open AI-based
- LLM based
- Good at solving not defined requirements in both conversational and completion modes
- Option to use Retrieval Augmented Generation (RAG)

In the case related to analyzing the video stream (input stream from drones and robots' cameras during construction observation), ChatGPT-4-TV is empowered by Azure AI Vision.

Azure AI Vision is a different solution for computer vision analysis, which uses cognitive analysis to understand videos and images. It has the following characteristics:

- Not LLM based
- Good in clearly defined tasks, object recognition, etc.
- Working with video streaming.

AI Vision extracts frames from the stream and sends them to the OpenAI service using a ChatGPT-4-TV deployment.

One high-level schema of the solution is presented in Fig. 3 [6]

The coordination service has a very important role in supporting and generating the list of questions needed to clarify the exact search and allow the OpenAI service to provide a correct solution.

Digital Twins (DT) is an optional part of the solution, used for two reasons:

- Decomposition of complex cases
- Simulations in Predictive Analysis

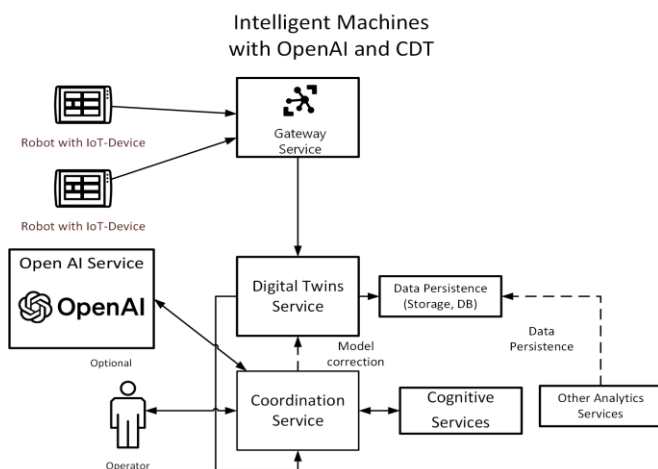


Fig.3. Experimental Setup: Predictive analysis using Computer Vision Open AI based on Microsoft Azure [6]

For analysis, sample data of rusty construction components from the USA, like the one in Fig. 4 :



Fig.4. Sample image of steel construction with corrosion

3.2 Democratization of the implementation of flows, providing Predictive Analysis with OpenAI and Power Automate (Low Code/No Code)

One of the leading cases using computer vision is to prototype the PoC fast, proving that the concept is feasible.

For PoC, Power Automate with Azure Open AI using deployment of ChatGPT-4-TV – Fig. 5

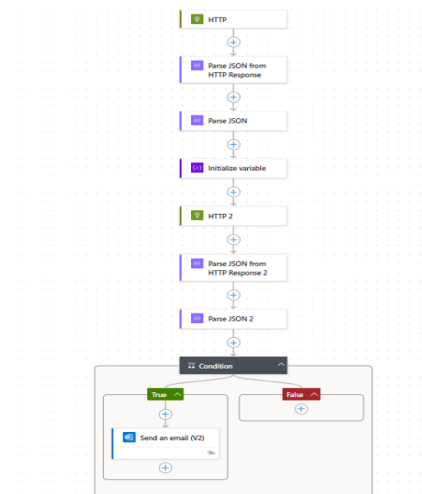


Fig.5. Power Automate flow, implementing Predictive Analysis.

3.3 Optimization of the cost and performance

One important goal when solving complex problems with Generative AI is to find a good solution to optimize the cost. In this research, steps are considered assuming the successful Prompt Engineering approach has already been implemented.

There are several options for LLM optimization – Fig. 6

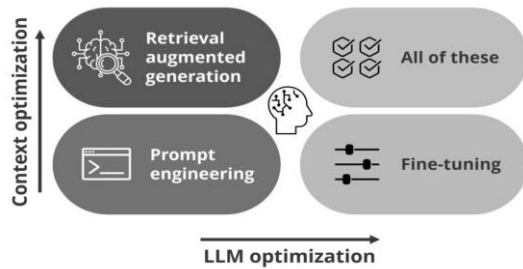


Fig.6. LLM Optimization stages.

RAG is used to collect embeddings from already performed requests and implement vector search to semantically compare the requests comparing vectors. Azure AI Search is currently used for the Vector Search implementation – Fig.7. [7]

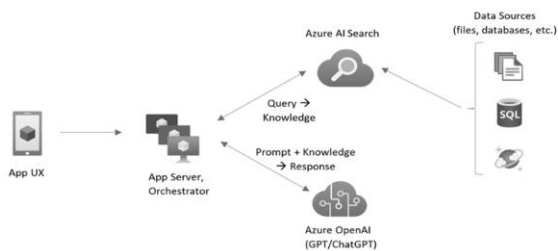


Fig.7. RAG Implementation with Azure AI Search.

4. Results and discussion

In this paper, there are added metrics related to the experimental project PoC, realized using Microsoft Azure, OpenAI (ChatGPT-4-TV), and 1000 experimental images. The solution is implemented using Custom Solution – Fig.3 and Low-Code/No-Code solution with Power Automate.

The experiment measures the prediction success rate for both custom and Low-Code/No-Code solutions to predict construction issues—Table 2.

Table 2. Experimental results for accuracy

	#Items	#success rate	#time	#time for implementation
Custom Solution	1000	99.9%	15 min	1 day
Low-Code/No-Code	1000	99.9%	49 nub	10 days

5. Conclusions

Modern Predictive Analysis in Construction can rely on Computer Vision. The overall analysis offered by automation solutions can effectively solve complex cases even on the PoC level using both - custom and Low-Code/No-Code solutions. The approach is domain-agnostic and can be used in different business domains.

Cost and performance optimization in the case of complex solutions with many components is an important new area, where powered from OpenAI solutions can optimize the maintenance cost for the structures

6. Future work:

Research will be extended based on more complex cases and improved MML models. The major focus will be solving real-life prediction and maintenance cases related to manufacturing systems.

7. Abbreviations

1. LLM – Large Language Models
2. SLM – Small Language Models
3. DT Digital Twins
4. ADT Azure Digital Twins
5. DT Cognitive Digital Twins
6. ChatGPT Generative pre-trained transformers
7. ChatGPT-4-TV-> ChatGPT 4-Turbo with Vision.
8. RAG -> use Retrieval Augmented Generation

8. References

- [1] T. Fernandez, "How to Choose the Best OpenAI Model for Your AI Application," Semaphore CI, Aug. 10, 2023. [Online]. Available: <https://semaphoreci.com/blog/openai-models>
- [2] Lu, Yuqian, et al. "Digital Twin-driven Smart Manufacturing: Connotation, Reference Model, Applications and Research Issues." 2020, <https://doi.org/10.1016/j.rcim.2019.101837>.
- [3] "Generative Pre-Trained Transformer." Wikipedia, 26 July 2023, en.wikipedia.org/wiki/Generative_pre-trained_transformer#Foundational_models.
- [4] "GPT-4 Technical Report". Arxiv.Org, 2023, <https://arxiv.org/abs/2303.08774>. Accessed 8 Aug 2023.
- [5] M., Mateev. "INDUSTRY 4.0 AND THE DIGITAL TWIN FOR BUILDING INDUSTRY". Industry 4.0, vol 5, no. 1, 2020, pp. 29-32., <https://stumejournals.com/journals/i4/2020/1/29>. Accessed 9 Aug 2023.
- [6] Mateev M. Predictive Analytics Based on Digital Twins, Generative AI, and ChatGPT, Proceedings of the 27th World Multi-Conference on Systemics, Cybernetics and Informatics: WMSCI 2023, pp. 168-174 (2023); <https://doi.org/10.54808/WMSCI2023.01.168>
- [7] HeidiSteen. 2024. "RAG and Generative AI - Azure AI Search." RAG and Generative AI - Azure AI Search | Microsoft Learn. Accessed June 12. <https://learn.microsoft.com/en-us/azure/search/retrieval-augmented-generation-overview..>
- [8] Mateev, Mihail. "Using Azure AI Vision and Open AI to analyze images for predictive analysis," BSC Technology AI Bootcamp. [Online]. Available: <https://bsc.technology/AIBootcamp/>.

Designing the spinal column of a quadruped robot by using additive manufacturing: A comprehensive approach

Ilo Bodi^{1,*}, Erald Piperi², Pascal Doumalin³, Mariel Hoxha⁴
 Polytechnic University of Tirana, Albania^{1,2,4}
 Institut Pprime, UPR CNRS 3346, Université de Poitiers³
 ibodi@fim.edu.al

Abstract: In this study, we aim to realize a methodology for designing the spinal column of a quadruped robot employing additive manufacturing techniques. Additive manufacturing processes have gained widespread usage owing to their ability to swiftly develop prototypes for research and actualize production-grade components. Our research commenced with the acquisition of vertebral dimensions sourced from real animals such as cheetahs. These dimensions were pivotal inputs for conducting kinematic analyses of the animal's locomotion cycle.

Through meticulous kinematic analysis, it was delineated the various positions assumed by the spinal column throughout the locomotion cycle. Furthermore, we assessed the maximal force to which the spinal column could be subjected. This evaluation formed a robust foundation for exploring diverse modeling approaches to identify the optimal solutions. An innovative solution emerged through the utilization of a beam featuring a variable cross-section. This approach facilitated precise customization of the design to align with our specific requirements. To manufacture all requisite components for the spinal column, we employed a single 3D printer utilizing the Fused Deposition Modelling (FDM) technique with PLA material. To validate the computational methods, it was developed a sophisticated test bench. By juxtaposing theoretical predictions against tangible experimental data, it was affirmed the accuracy of the theoretical approach. This validation serves as a springboard for subsequent phases in the design and production processes of quadruped robots.

Keywords: QUADRUPED ROBOT, FUSED DEPOSITION MODELLING, PLA MATERIAL

1. Introduction

Quadruped robots are designed to emulate the locomotion of animals in the natural world. The focus of this work lies in dissecting and modeling the spine of such robots, with the initial prototype being integrated into a cheetah-inspired robot. The primary impetus behind designing a cheetah-inspired robot lies in its unparalleled speed. Achieving high-speed locomotion in robots poses a formidable challenge, yet it holds immense potential across varied applications. From exploration and surveillance to deployment in hazardous environments such as disaster zones, the versatility of quadruped robots knows few bounds. Moreover, their utility extends to domains like military, agriculture, and education. The quest for swift robotic locomotion necessitates drawing inspiration from nature, where the cheetah reigns supreme as the fastest land animal. The spinal column's significance in orchestrating the movement of quadruped robots cannot be overstated [1]. The results showed that a passive spine where dorsal stiffness is higher than ventral stiffness can run faster than a rigid case, because of the compliance of the spine, which might play a role as a nonlinear spring to enhance locomotion by transferring and storing energy. Serving as a structural scaffold, it enables fluid and coordinated motion akin to walking, running, and turning. Designing an efficacious spine demands a multidisciplinary approach, integrating principles of mechanics, electronics, control systems, and bioinspiration [2]. The forelimb musculature comprised $15.1 \pm 1.2\%$ of its total body mass, substantially less than its hindlimb at $19.8 \pm 2.2\%$ of total body mass. Kawasaki et al [3], revealed that the bounding gait of the robots with a flexible spine is more stable and more natural than that of the robots with a rigid spine, and the robot can run faster by synchronizing the movements of their spine and legs. Matsumoto et al. showed clearly how their prototype will move its spine, i.e., when the feline is running, the main mobilities are at the animal's sacrum [4].

Ultimately, the construction of a robot mirroring the biomechanics of real animals promises invaluable insights into the complexities of emulating organic movement through mechanical means [5]. Our objective in this work is to decipher the intricacies of spine movement and apply them to quadruped robotics. Commencing with a scaled-down prototype produced with the 3D printing technique, was designed the experiment to validate the study methodology, a pivotal step necessitating the establishment of a dedicated test bench.

2. Experiment

The goal of this study was to create a panther spine. To meet the different requirements of design and production specifications, such as the accuracy of a panther spine movement, and to consider the size and weight of the system different aspects were examined. All the work done led us to design the spine of a continuous robot thanks to compatible joints. However, to dimension all the parts of the spine, it is necessary to know its rigidity, displacements, forces, and size. Unfortunately, the information in the literature on these factors is very low, so we have to find our own.

1. Geometry analysis

A fully 3D-scanned panther skeleton in STEP format was used to analyze, design, and produce it with 3D printing techniques. In this file format, we were able to see the different sizes of the vertebrae and precisely how the skeleton is made up as shown in Figure 1.



Fig. 1 Fully 3D scanned panther skeleton.

Thanks to the skeleton, we were able to determine the articular stops of the spine, i.e. the stops resulting from the position of the vertebrae with each other and their shape, which means that their movement can be limited. This information may be important to us because we have relatively little data on the mobility of the spine and the stops ensure that we limit our movement to something feasible for the animal. To find out the joint stops of the spine, as you can see in the following figure, we draw lines from the middle of the joint (presumed center of rotation) to the part of the vertebra most likely to come into contact with the second vertebra. In this way, we can find the joint's limit of rotation at the joint stops as shown in Figure 2.

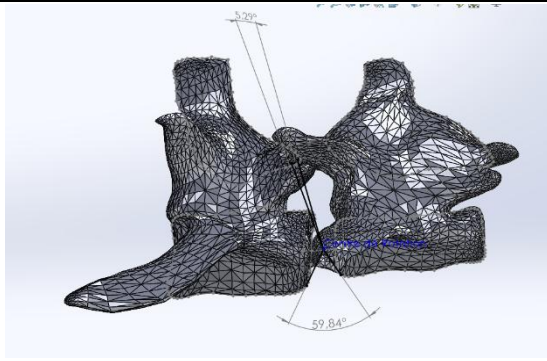


Fig. 2 Joint stop determination

All the vertebrae dimensions are shown in the Table 1.

Table 1. Vertebrae length and height.

Vertebrae	Length (mm)	Height (mm)
T1	54.6	27.2
T2	53.6	25
T3	55.8	29.6
T4	53.8	38.8
T5	50.4	36.2
T6	48.4	36.2
T7	50.2	39.2
T8	59.2	36.2
T9	62	42.6
L1	63.6	38.4
L2	62	43.4
L3	65.2	/
L4	80	35.2
L5	86	41
L6	77	43
L7	82.16	30.3
L8	96.83	32.9
L9	108.4	37.8

2. Kinematic analysis

To dimension the spine accurately, it's essential to gather data regarding its position and the load it bears. This process involves analyzing and processing videos to extract the necessary information. To gather data on the forces exerted on the back and to ascertain the spine's stiffness, we are analyzing a cheetah's running motion. To precisely monitor the positions of the bones throughout the cheetah's run, we utilized a specialized software known as DeepLabCut [5].

DeepLabCut™ is an efficient method for 2D and 3D markerless pose estimation based on transfer learning with deep neural networks that achieves excellent results with minimal training data. We can track various body parts in multiple species. The package is open-source, fast, and robust. It can be used to compute 3D poses. The results of the cheetah running position are shown in Figure 3.

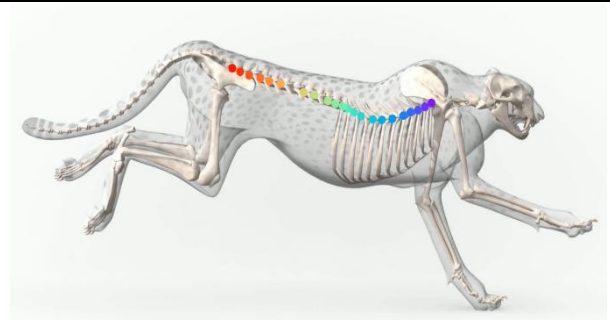


Fig. 3 Modelisation of a running cheetah video.

In our case, we placed, throughout the entire cycle, markers on each vertebra. Then, the software interpolates this point to trace the trajectory for each marker as shown in Figure 4.

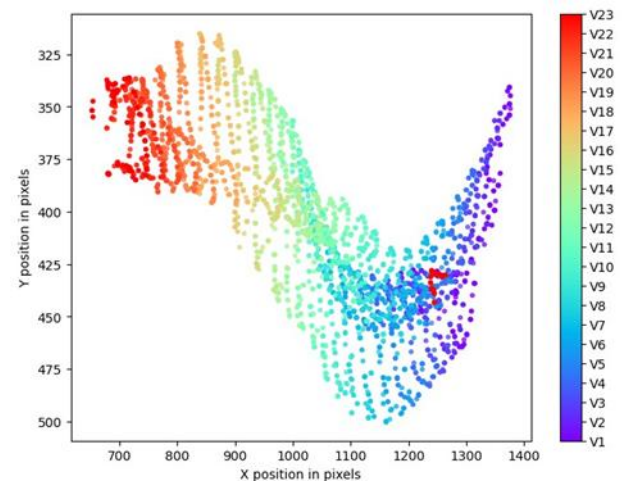


Fig. 4 Position of each vertebrae in pixels.

On this graph, it's evident that each vertebra doesn't move in the same way. For instance, the thoracic vertebrae display less mobility compared to the lumbar ones. Additionally, certain vertebrae show similar amplitudes of movement. Therefore, in our prototype design, we won't model all 22 individual vertebrae separately. Instead, some vertebrae will be combined. For instance, the 11th and 12th vertebrae are closely positioned and exhibit similar movement patterns. This approach was adopted in the spine's modeling process.

3. Materials and methods

The design process began by making a succession of parts corresponding to each vertebra. In the interests of efficiency, we were able to discover and use a SolidWorks software function that consists of using an Excel file to change our rib configurations to be able to make faster studies. We started by creating our column and, as presented in our calculations, by using identical vertebrae lengths as shown in Figure 5. Although this is not realistic, once again, this will simplify our system and enable us to obtain a result.

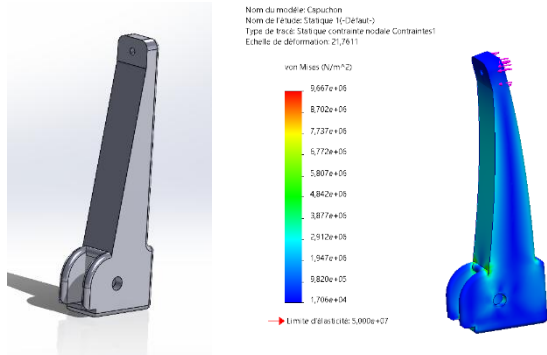


Fig. 5 Fin of the column and deformation test.

For parts production, we utilized a 3D printer to fabricate each component of the spine using PLA material as shown in Figure 6. Following this, we meticulously cleaned and labeled each part to facilitate assembly, ensuring efficiency and accuracy in the process.



Fig. 6 3D printed vertebra parts.

The decision to utilize 3D printing was driven by its rapid and economical nature, given that intricate precision was not paramount for the prototype. We selected PLA material because it was easy to use and with very good characteristics in prototyping. While this method may alter material characteristics, such as anisotropy, the way impressions are done determines the direction of rigidity. For our purposes, this was inconsequential as our focus remained on prototyping. We also fulfill 100 % of each part to limit the porosity. Moreover, the final spine will be in aluminum, so the fabrication process will be way more accurate.

4. Results

A test bench is an essential component in the design and verification process of real parts. It serves as a simulated environment in which designers can test the functionality, performance, and correctness of their designs before implementing them. The main purpose of using a test bench is to catch and correct

errors early in the design process, saving time and resources. The test bench must be resistant to the weight of the spine and vibrations due to movement. The spine must be placed fixed in the device because if we have movement of the spine, we get wrong test results. The weight of the spine and the forces that will act during the test were taken into consideration during the choice of the model. The test bench should be adjustable, the spine should be easily assembled, the dimensions should not be in large proportion to the dimensions of the spine, and it should be suitable for use. Monitoring and recording the results in a simple way of each test case. Identifying and documenting any problems or deviations from expected behavior as shown in Figure 7.

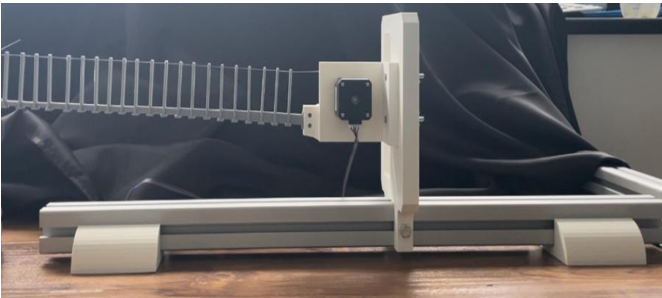


Fig. 7 The assembly of the test bench and the spine.
The components of the test bench are:

- NEMA 17 STEPPER MOTOR: Enables the movements required for testing
- Profile 45x45: Enables the assembly of the testbench board
- Support for Profile: It ensures stability and avoids vibration.
- Support for Motor: It enables the assembly of the motor and the spine during the test.
- Plate of bench: It is resistant to all forces and vibrations during the test.
- Pulley: Perform the necessary rotation motion during the test.

The aim of the tests we're about to carry out is, firstly, to check the reliability of the calculations we've made to dimension our column and, secondly, to verify whether it is indeed possible to activate the deformation of our column using a motor. After acquiring the beam and the bench, our next step was to assess whether the prototype would exhibit the same characteristics as the theoretical spine. To achieve this, we conducted two tests. The first test involved examining the flexion of the beam under a single-weight load (10N). Subsequently, we proceeded to test the response of the spine to various loads. To validate our calculations, we need to apply exactly the force we want to the cable to check that the displacement at the end of the column is correct. We can then compare the experimental value with the theoretical one. To do this, we had several choices. The first was to use the motor and a load cell, which we would place on the cap and retrieve the force value when the desired displacement was reached. The second was to know the load applied to the cable beforehand. Given that it would have been difficult to install a load cell on our system as we might have wished, but also to limit expenditure as much as possible and save the time we had available, we decided to use the second solution. The testing procedure is shown in Figure 8

Step	Instruction	Comment / Picture
1	Prepare the Spine: Ensure the spine is firmly built in on the bench to prevent any movement during the experiment.	

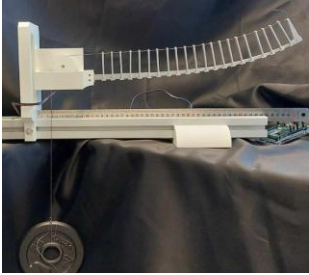
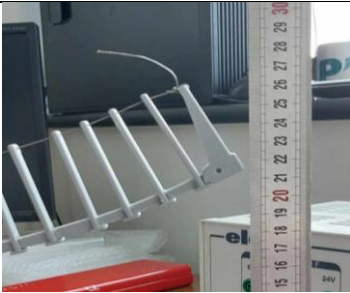
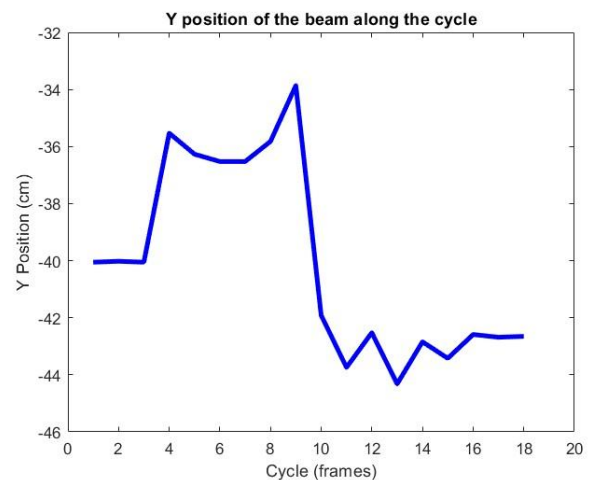
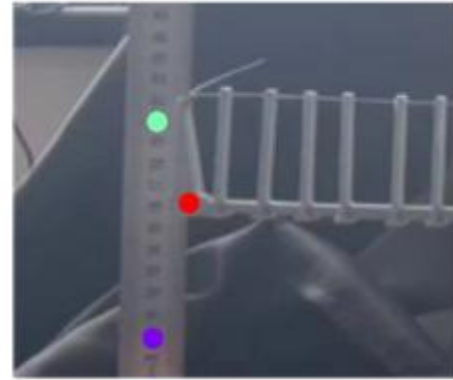
2	Wire Attachment: Affix the wire to the spine.	 <i>Wire attachment</i>
3	Calibration: Place a ruler on the same plane as the spine and wire to capture images for calibration purposes.	
4	Initial Photo: Take an initial picture of the setup before any external force is applied.	 <i>Initial picture with ruler</i>
5	Applying Weight: Place the specified weight on the wire.	 <i>Spine with hanging weight</i>
6	Final Photo: Capture a final image after releasing the weight.	 <i>Final picture with ruler</i> Using the ruler at the rear of the column, we can now calculate the displacement between the phases in which forces are applied and those in which they are not.
7	Deformation Measurement: Measure the deformation of the spine from the initial to the final photo.	

Fig. 8 Step of weight experiment.

We needed to program the motor to replicate the movement of the spine precisely as it occurs in the panther. To achieve this, we utilized the position of the sacrum throughout its cycle and converted it into a binary signal. This signal transitions in increments of 5 steps. Given that we determined that 6 cm corresponds to 45 steps, we could calculate the number of steps required for each centimeter of deformation. With this information, we proceeded to write the corresponding code for the schematic as shown in Figure 9.

**Fig. 9** Step of weight experiment.

As we can observe, we can find the same trace as the real position. However, the motor jumped during the test. During the test, we saw that after the movement, the sacrum position vibrates a lot. To conclude, we found another way to program the motor to suppress jumps.

5. Conclusions

To calculate and dimension the prototype of the spine, we opted for a variable-section beam activated by a cable, mimicking the mechanism of a continuous robot. This approach guided the design of our prototype and test bench. The primary objective of our system was to validate our calculation methodology. Consequently, we conducted two tests for this purpose. The first test aimed to confirm the accuracy of our intended force application by exerting a pure force on the cable. This validation ensured the fidelity of our calculations. Subsequently, the second test aimed to demonstrate the feasibility of activating our system using a motor. By successfully executing these tests, we were able to affirm the validity of our results.

The primary objective of this study was to decipher the intricate movements of a cheetah's spine to replicate them for application in a quadruped robot. Initially, our focus was on comprehending the dynamics of a cheetah's running motion, which provided essential

data regarding spine behavior. Our study primarily delved into the 2D aspects of this behavior, allowing us to validate our hypotheses. However, to develop a more sophisticated spine system, a deeper exploration into the 3D movement is imperative. This necessitates delving into additional articles and studies that shed light on the intricacies of the cheetah movement.

In our spine dimensioning process, we relied on a CAD file as a reference point. However, to further validate this data, acquiring real-world measurements from an actual cheetah would be invaluable. Performing an autopsy on a cheetah would provide precise anatomical data crucial for our study.

Given our focus on 2D analysis, we opted for a beam structure. However, if our study progresses, incorporating compliant joints into the spine design becomes imperative for a more comprehensive understanding.

Regarding the iterative method used to determine the spine's sectional dimensions, we aim to enhance this process through optimization techniques. By refining the method, we can achieve the optimal distribution of sectional properties for enhanced performance.

6. References

- Q. Zhao, B. Ellenberger, H. Sumioka, T. Sandy and R. Pfeifer, "The effect of spine actuation and stiffness on a pneumatically-driven quadruped robot for cheetah-like locomotion," in *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Shenzhen, China, 2013.
- P. E. Hudson, S. A. Corr, R. C. Payne-Davis, S. N. Clancy, E. Lane and A. M. Wilson, "Functional anatomy of the cheetah (*Acinonyx jubatus*) forelimb," *Journal of Anatomy*, vol. 238, pp. 375-385, 2011.
- R. Kawasaki, R. Sato, E. Kazama, A. Ming and M. Shimojo, "Development of a flexible coupled spine mechanism for a small quadruped robot," in *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Qingdao, China, 2016.
- O. Matsumoto, H. Tanaka, T. Kawasetsu and K. Hosoda, "High-speed running quadruped robot with a multi-joint spine adopting a 1DoF closed-loop linkage," *Frontiers in Robotics and AI*, vol. 10, pp. 1-13, 2023.
- "<https://github.com/DeepLabCut/DeepLabCut>," [Online]. Available: <https://github.com/DeepLabCut/DeepLabCut>. [Accessed 20 March 2024].

Photoluminescence properties of self-cleaning coatings on glasses containing Eu_2O_3

Stancho Yordanov

Bulgarian Academy of Sciences, Institute of Metal Science, Equipment and Technologies with Hydro- and Aerodynamics Centre "Acad. Angel Balevski", 67 "Shipchenski prohod" Blvd., 1574 Sofia, Bulgaria,
stancho14@abv.bg

Abstract: Obtained experimental self-cleaning coatings (based on compositions with the participation of Eu_2O_3), applied by the sol-gel method on glass slides. The luminescence of the obtained thin layers of different ratios of the compositions were investigated. A strong emission from the $^5\text{D}_0$ level of Eu^{3+} ions was registered under near UV (392 nm) excitation using the $^7\text{F}_0 \rightarrow ^5\text{L}_6$ transition. The values of the integrated fluorescence intensity ratio $R (^5\text{D}_0 \rightarrow ^7\text{F}_2) / (^5\text{D}_0 \rightarrow ^7\text{F}_1)$ was calculated to give information about the degree of asymmetry in the vicinity around the active ions, suggesting a location of Eu^{3+} in non-centrosymmetric sites. The highest Eu^{3+} luminescence is observed in sample with 3 layers coating. The experimental results represent a prerequisite for the development of a series of additional compositions and a detailed technological regime for obtaining various modifications of resistant, long-lasting self-cleaning coatings, potentially applicable to photovoltaic panels.

Keywords: SOL GEL, SELF-CLEANING COATINGS, LUMINESCENCE, SOLAR CELLS, PHOTOVOLTAICS, Eu_2O_3

1. Introduction

The study of photoluminescence (PL) is widely used in materials science, optoelectronics and photovoltaics. PL spectroscopy characterizes defects, impurities and electronic properties of materials. Researchers study PL spectra to understand carrier dynamics, band structures, and luminescence properties. There are various inorganic compounds used to achieve a luminescent effect. The most popular is the one based on rare earth elements, especially Eu_2O_3 .

Luminescence is described as the emission of electromagnetic radiation in excess of thermal radiation. In most cases, the emitted radiation appears in the visible range, although ultraviolet and infrared emissions have also been observed [1]. The role of this effect can be noticeable in specific applications as well as during emission intensity measurement processes. Luminescence is also used practically in many applications, including light-emitting diodes (LEDs), storage phosphors, persistent luminescent phosphors, scintillators, and in particular in materials for up- and down-conversion mechanisms that are designed for photovoltaics [2]. Eu also plays an important role in the structural and optical properties of the systems/incl. coatings on glasses/. Increasing the Eu concentration improves the transparency, refractive index, and thickness of the films, along with a quadratic increase in PL emission. Some studies have focused on Eu silica-sodium sol-gel glasses. According to molecular dynamics simulations, adding soda to silica can reduce RE clustering and precipitation [3]. New innovative materials based on various luminescent effects have been investigated. The obtained results are discussed from the point of view of application in next-generation photovoltaics. Experiments leading to the production of functional photoconversion layers dedicated to low-cost and large-scale deposition methods are investigated. Various active agents, as well as possible organic carrier matrices, substrate preparation and treatment procedures are presented [4]. Hydrophilic photocatalytically active thin films of ZnO and ZnO doped with Yb_2O_3 and Eu_2O_3 were prepared by a combination of sol-gel technique, spin-coating and high-temperature heat treatment at 500 and 600°C. The methodology presented by the authors to produce thin films featuring the key properties of self-cleaning coatings can be applied to coatings on photovoltaic panels and architectural glass structures [5]. Spectroscopy of lanthanide complexes is presented and the optimal matrix for the luminescent species, which is based on sol-gel technology, is explained and developed. The theory, requirements and conditions for the production of luminescent solar concentrators are presented and the modern methods for their production are outlined. The practical approach to embedding luminescent plates in modern buildings for the utilization of solar energy is presented [6].

2. Experimental procedures

All compounds were supplied as analytical grades and were used without further modification. Thin film coatings were performed on slides from AGC Glass Europe. $\text{ZrOCl}_2 \cdot 8\text{H}_2\text{O}$,

$\text{Y}(\text{NO}_3)_3 \cdot 6\text{H}_2\text{O}$, 99.9% ethanol ($\text{C}_2\text{H}_5\text{OH}$), all purchased from Alfa Aesar Germany. The stabilizer is nitric acid HNO_3 .

Preparation of the glass substrate. Before being cleaned, the slides were cut to size (20 mm x 70 mm x 2 mm). Substrate cleaning is important prior to the deposition process, as this allows the cleaning and removal of submicron particles from the substrate surface. This is decisive for the quality of the adhesive film on the base. The glass substrates were cleaned/washed/in 99.9% ethyl alcohol for 15 minutes and acetone for 15 minutes and then dried for 30 minutes at 100°C. Cleaning the glass substrates improves the adhesion of the coating material to the substrate and prevents contamination of the formed coating. In table 1 presents the compositions and the technology of obtaining.

Table 1

Sample:	Main Composition:	Drying Time T-150°C	T°C-	layer s
GLS-ZrYEu3	$\text{ZrOCl}_2 \cdot 8\text{H}_2\text{O} - 9.889 \text{ g}$ $\text{Y}(\text{NO}_3)_3 \cdot 6\text{H}_2\text{O} - 0.831 \text{ g}$ $\text{Eu}_2\text{O}_3 - 0.407 \text{ mg}$	60 min per layer	420°C- 2 hour Delay 15 min	3
GLS-ZrYEu4	$\text{ZrOCl}_2 \cdot 8\text{H}_2\text{O} - 9.889 \text{ g}$ $\text{Y}(\text{NO}_3)_3 \cdot 6\text{H}_2\text{O} - 0.831 \text{ g}$ $\text{Eu}_2\text{O}_3 - 0.407 \text{ mg}$	60 min per layer	410°C- 2 hour Delay 15 min	4
GLS-ZrYEu6	$\text{ZrOCl}_2 \cdot 8\text{H}_2\text{O} - 2.942 \text{ g}$ $\text{Y}(\text{NO}_3)_3 \cdot 6\text{H}_2\text{O} - 0.831 \text{ g}$ $\text{Eu}_2\text{O}_3 - 0.407 \text{ mg}$	60 min per layer	440°C- 2 hour Delay 15 min	6
GLS-ZrYEu7	$\text{ZrOCl}_2 \cdot 8\text{H}_2\text{O} - 9.889 \text{ g}$ $\text{Y}(\text{NO}_3)_3 \cdot 6\text{H}_2\text{O} - 0.829 \text{ g}$ $\text{Eu}_2\text{O}_3 - 0.407 \text{ mg}$	60 min per layer	400°C- 2 hour Delay 15 min	7

The table shows the compositions and production technology of
GLS-ZrYEu3, 4, 6 and 7

Sol-gel technology is applied. $\text{ZrOCl}_2 \cdot 8\text{H}_2\text{O}$ is dissolved in absolute alcohol - so that a 0.3 - 0.4 molar solution is obtained. $\text{Y}(\text{NO}_3)_3 \cdot 6\text{H}_2\text{O}$ is added to it slowly with vigorous stirring. After 60 min, Eu_2O_3 was added. During hydrolysis, nitric acid is added or stirred until clear, i.e. until completely dissolved. To cover the glass samples with the resulting solution, the Dip Coating technology is applied [7-9].

3. Results and Discussion

Photoluminescence (PL) emission spectra at room temperature for all coatings were measured with a Spectrofluorometer FluoroLog3-22, 2014 (Horiba JobinYvon, Longjumeau, France).

Luminescent properties: Figure 1 shows emission spectra of coatings (ZrO_2 , Y_2O_3 and Eu_2O_3) on glasses

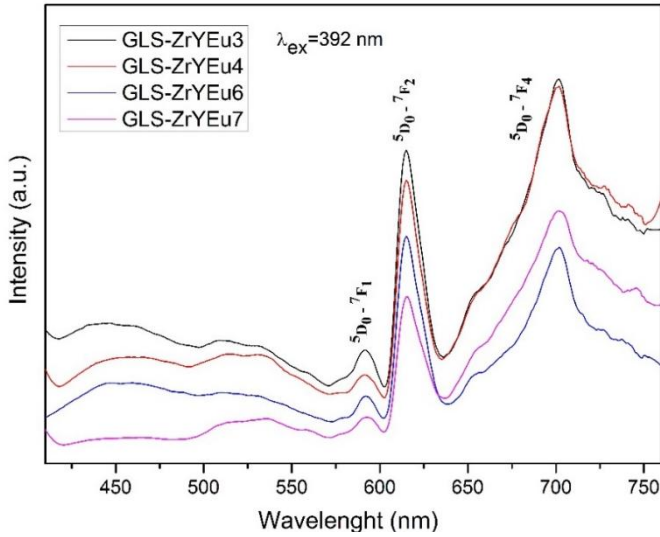


Figure 1. Emission spectra of coatings (ZrO₂, Y₂O₃ and Eu₂O₃) on glasses.

The emission spectra of the obtained coatings (ZrO₂, Y₂O₃ and Eu₂O₃) on glasses upon excitation into the most intensive $^7F_0 \rightarrow ^5L_6$ transition of Eu³⁺ at 392 nm are shown in figure 1. All the spectra are recorded in the same conditions at room temperature and have the same profiles, consisting of two regions: a broad band of Y₂O₃ emission below 570 nm and narrow emission peaks at around 591 nm, 615 nm and 701 nm in the visible region, attributed to $^5D_0 \rightarrow ^7F_1$ transitions of the Eu³⁺ ion from 5D_0 to 7F_1 , 7F_2 and 7F_4 states in order of increasing wavelength, respectively [10]. Since itself ZrO₂ is not luminous, no emission band originating from ZrO₂ under 392 nm excitation is observed. Comparing the spectral peaks, it can be seen that the most intensive is the luminescent band at 701 nm corresponding to $^5D_0 \rightarrow ^7F_4$ transition. This behavior is atypical for the most Eu³⁺ doped materials. It is known that the intensity of the $^5D_0 \rightarrow ^7F_4$ transition is determined by either the chemical composition of the host matrix or by symmetry factors [11].

As can be seen from figure 1, with the decrease of the number of the glass layers, a significant enhancement of the emission intensities was achieved. The most intensive emission is registered in the sample with 3 layers (GLS-ZrYEu3).

Europium is a suitable probe for determining the local symmetry because its electric dipole transition (ED) $^5D_0 \rightarrow ^7F_2$ is hypersensitive to the small changes in the surroundings and usually dominates the emission spectrum when Eu³⁺ ions are located in low site symmetry, while the magnetic dipole transition (MD) $^5D_0 \rightarrow ^7F_1$ transition is insensitive to the environment around the rare earth ion [12, 13]. The fact that the predominant emission is from the ED transition rather than from the MD transition is an indication that Eu³⁺ ions are distributed in a non-inversion symmetry sites. Therefore, the value of integrated fluorescence intensity ratio R of the two transitions ($^5D_0 \rightarrow ^7F_2$)/($^5D_0 \rightarrow ^7F_1$) gives information of the degree of asymmetry in the vicinity of Eu³⁺ ions [14,15]. The lower the value of the asymmetry parameter, the higher the site symmetry of the rare earth ion and the lower Eu–O covalence and emission intensity. Ratio values lower than 1 correspond to Eu³⁺ ions location into high symmetry sites, while values above 1 indicate that the active ion occupies lower symmetry sites.

The calculated values, listed at table 2, are in range 6.57 to 10.80 suggesting more asymmetry around Eu³⁺ ions and stronger Eu–O covalence and thus higher emission intensity.

Table 2

Sample composition	Fluorescence intensity ratio, R
GLS-ZrYEu3	10.80
GLS-ZrYEu4	9.70
GLS-ZrYEu6	8.46
GLS-ZrYEu7	6.57

Fluorescence intensity ratio (R) of the two transitions ($^5D_0 \rightarrow ^7F_2$)/($^5D_0 \rightarrow ^7F_1$) of coatings (ZrO₂, Y₂O₃ and Eu₂O₃) on glasses.

CIE color coordinates and CCT (K) values

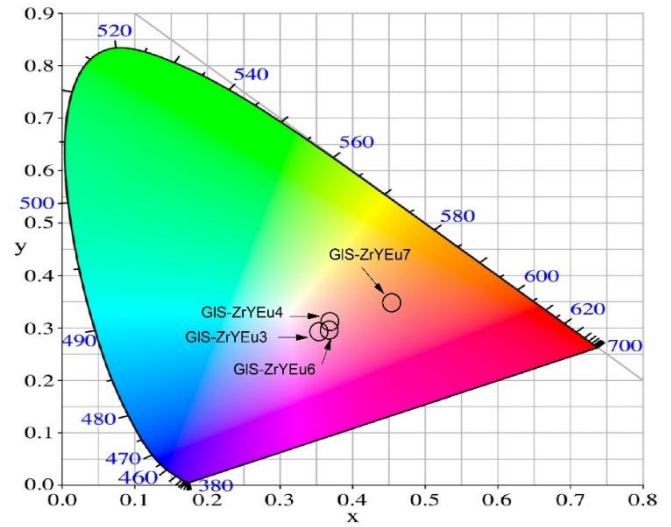


Figure 2. CIE chromaticity diagram of coatings (ZrO₂, Y₂O₃ and Eu₂O₃) on glasses

Table 3

Sample composition	Chromaticity coordinates (x,y)	CCT(K)
GLS-ZrYEu3	(0.351, 0.305)	4518.80
GLS-ZrYEu4	(0.366, 0.324)	4046.61
GLS-ZrYEu6	(0.365, 0.308)	3979.06
GLS-ZrYEu7	(0.451, 0.360)	2196.40

CIE chromaticity coordinates of glass coatings containing ZrO₂, Y₂O₃ and Eu₂O₃ oxides

To visualize the actual color of emission, the standard Commission International de l'Eclairage (CIE) 1931 chromaticity diagram was used [16]. From the luminescence spectra the chromaticity coordinates of samples were calculated using color calculator software (CIE coordinate calculator). The obtained values are listed in table 2 and correspond to pale pink color of emission. The CIE chromaticity diagram is displayed in figure 2. These coordinates are useful not only for determining the exact emission color, but also for calculating the color correlated temperature (CCT) of samples, which in addition, usually serves as another indicator for characterizing the properties of emitted light.

The CCT was calculated by the McCamy empirical formula:

$$CCT = -449n^3 + 3525n^2 - 6823n + 5520.33 \quad [17],$$

where $n = (x - x_e)/(y - y_e)$ is the reciprocal slope, ($x_e = 0.332$, $y_e = 0.186$) are epicenter of convergence and x and y are the chromaticity coordinates.

Generally, luminescent materials, characterized with CCT values below 4000 K are usually considered as a warm source of light, while values above 4000 K, as a cold light source [17]. The calculated CCT values of Eu³⁺ doped glasses (table 3) ranges from 2196.40 K to 4518.80 K. In particular, samples GLS-ZrYEu6 and GLS-ZrYEu7 exhibit a warm light emission with a CCT values of 3979.06 K and 2196.40 K, while GLS-ZrYEu3 and GLS-ZrYEu4 with values 4518.80 K and 4046.61 K respectively, can be considered as a cool light illumination sources.

4. Conclusions

The luminescent properties of the obtained Eu^{3+} doped samples revealed that they could be efficiently excited by 392 nm corresponding to the ${}^7\text{F}_0 \rightarrow {}^5\text{L}_6$ transition and are characterised by a very large value of the luminescence intensity ratio, R (over 8), indicating that Eu^{3+} ions occupy highly distorted sites in the host structure. In addition, the values of CIE chromaticity coordinates and color correlated temperature demonstrate that the obtained samples GLS-ZrYEu3 and GLS-ZrYEu4 can be considered as a cool pink light emitting sources, while GLS-ZrYEu6 and GLS-ZrYEu7 as warm light sources. The obtained results are evidencing that Eu^{3+} ions are successfully incorporated in the prepared coatings. The experimental results represent a prerequisite for the development of a series of additional compositions and a detailed technological regime for obtaining various modifications of resistant, long-lasting self-cleaning coatings, potentially applicable to photovoltaic panels.

5. References

1. K.V.R Murthy, H.S. Virk, Luminescence Phenomena: An Introduction, Defect Diffus. Forum, (2014), 347, 1–34
2. B.S. Richards, Luminescent layers for enhanced silicon solar cell performance: Down-conversion. Sol. Energy Mater. Sol. Cells, (2006), 90, 1189–1207
3. F. Enrichi, G. Mastantuoni, M. Cassetta and all, Structural and optical properties of Eu^{3+} -doped sol-gel silica-soda glasses, Eur. Phys. J. Plus, (2024) 139:346 <https://doi.org/10.1140/epjp/s13360-024-05151-w>
4. Maciej Sibiński, Review of Luminescence-Based Light Spectrum Modifications Methods and Materials for Photovoltaics Applications, *Materials*, (2023), 16(8), 3112
5. T. Tomasz, M. Zaborowska, P. Jarka, A. Woźniak, Hydrophilic ZnO thin films doped with ytterbium and europium oxide, Scientific Reports, (2022), 12(1), 11329
6. R. Reisfeld, Application of Luminescence Spectroscopy in New Materials for Solar Energy Utilization, in Reference Module in Chemistry, Molecular Sciences and Chemical Engineering, (2017)
7. V. Petkov, M. Aleksandrova, V. Blaskov, “Deposition of ZrO_2 thin films obtained by sol-gel method on monolithic composite with layered structure of Nb30Ti/Al”. International Scientific Journal "MACHINES. TECHNOLOGIES. MATERIALS", 16, 2, International Scientific Journals of Scientific Technical Union of Mechanical Engineering "Industry 4.0", (2022), 58-61. ISSN:Print ISSN 1313-0226, Online ISSN 1314-507X
8. M. Aleksandrova, V. Petkov, V. Blaskov, V. Korzhov, V. Kiiko, I. Zheltyakova, Synthesis and Properties of Sol-Gel Coatings Deposited On Hybrid Materials. 6, 76, Доклади на БАН, "Comptes rendus de l'Académie bulgare des Sciences", (2023), ISSN:1310–1331 (Print), 2367–5535 (Online)
9. V. Petkov, M. Aleksandrova, R. Valov, “Partial oxidation of biocompatible titanium alloy Ti 6Al 4V at deposition of glassy carbon coating”. International Journal “NDT Days”, vol. 3, Issue 4, (2020), 225-230, ISSN:2603-4018
7. K. Binnemans, Interpretation of europium(III) spectra. *Coord. Chem. Rev.*, (2015), 295, 1-45
8. L. Xiang, X. Zhou, L. Li, F. Ling, Y. Li, J. Li, & H. Yang, Comparative study on the anomalous intense $5\text{D}_0 \rightarrow 7\text{F}_4$ emission in Eu^{3+} doped $\text{Sr}_2\text{MnNbO}_6$ ($\text{M} = \text{Ga}, \text{In}, \text{Gd}$) phosphors. *Journal of Luminescence*, (2024), 267, 120370
9. K. Binnemans, Interpretation of europium (III) spectra. *Coord. Chem. Rev.* (2015), 295, 1–45.
13. G. Blasse, B.C. Grabmaier, Luminescent Materials, 1st ed.; Springer: Berlin/Heidelber, Germany, (1994), p. 18
14. Devi, C.H.B.; Mahamuda, S.; Swapna, K.; Venkateswarlu, M.; Rao, A.S.; Prakash, G.V. Compositional dependence of red luminescence from Eu^{3+} ions doped single and mixed alkali fluoro tungsten tellurite glasses, *Opt. Mater.* (2017), 73, 260–267.
15. M. Nogami, N. Umehara, T. Hayakawa, Effect of hydroxyl bonds on persistent spectral hole burning in Eu^{3+} doped $\text{BaO-P}_2\text{O}_5$ glasses. *Phys. Rev. B*, (1998), 58, 6166–6171
16. T. Smith, & J. Guild, The CIE colorimetric standards and their use, Transactions of the optical society, (1931), 33(3), 73.
17. McCamy, Calvin S. Correlated color temperature as an explicit function of chromaticity coordinates, *Color Research & Application*, (1992), 17.2, 142-144

Statistical processing of data from the operation of a photovoltaic power plant

Svilen Rachev¹, Raul Turmanidze², Vazha Shilakadze²

¹Technical University – Gabrovo, Bulgaria

²Georgian Technical University (GTU) – Tbilisi, Georgia

srachev@mail.com, inform@gtu.ge, Vazhashilakadze@gmail.com

Abstract: The paper deals with an operating photovoltaic plant in Bulgaria. The wide application of photovoltaic plants in the energy system is known. At the same time, subject matters about their normal functioning are fundamental. The goal of the paper set is to analyze data from the operation of a photovoltaic power plant, which is ultimately related to improving its behavior. Relevant conclusions have been done.

Keywords: SOLAR ENERGY, PHOTOVOLTAIC SYSTEMS, DATA PROCESSING, SYSTEMS OPERATION, STATISTICS

1. Introduction

The development and level of energy resource usage largely determine the standard of living of any industrialized nation. In doing so, the most state-of-the-art technologies, materials, and production methods must be applied.

The use of so-called ‘renewable energy sources’ (RES) in the world is constantly growing. The goals are to achieve sufficient productivity with rational energy consumption and compliance with environmental pollution norms.

For many years, the usage of solar energy has been a matter of great importance. Standard energy resources such as oil, natural gas and coal are gradually being depleted, and new deposits are being developed less and less, as they either have small estimated reserves and will be very expensive to extract, or are too deep, which also makes extraction more expensive. The other important point is the pressing issue of protecting the environment from technological pollution when using traditional methods of energy extraction [1].

An essential problem of humanity is the production of electricity. Obtaining electrical energy from nuclear power plants (NPPs) is by no means the best alternative to thermal power plants (TPPs), as there is a danger of nuclear contamination, and the life of the NPPs themselves is limited to certain years, after which they must be ‘buried’ in expensive facilities. Hydroelectric power plants (HPPs), on the other hand, are of limited application due to the specific conditions they require. Moreover, for all these sources of energy, a lot of funds are needed for their maintenance.

Probably most knowledgeable people know that the end of ‘petroleum humanity’ is approaching. A real shortage of liquid fuels will occur. That is why, in recent decades, scientists from all over the world have focused their efforts on the development and implementation of technologies for the extraction of energy from RES.

One of the developing directions is solar energy usage. The huge amount of solar energy reaching the Earth and its immediate conversion into electricity is not only ecologically clean but also satisfy a significant part of all current and future needs of humanity. The production of electricity from the sun is among the most current directions in the development of modern energy and has many advantages.

In addition, the Renewable Energy Directive RED II (Directive on the promotion of the use of energy from RES) together with the European Green Deal are driving a radical rethink of energy investment across the EU [2, 3].

Bulgaria can be divided into different sunny zones, with the average annual duration of sunshine being about 2150 hours (about 49% of the maximum possible). Although 77% of its territory is occupied by forests, arable land, and nature reserves, the remaining territory could be used for solar power plants under certain conditions.

The geographical conditions of Bulgaria determine a resource of 1300÷1700 kWh/m² average annual radiant flow (amount of radiant energy for all wavelengths falling on a unit area of the country's surface). There are three pretenders for the sunniest city in Bulgaria - Velingrad, Sandanski, Chepelare in the last decade, but it is difficult to say exactly who is the winner.

Regarding the situation in Bulgaria, attention should be paid to several starting points. The energy strategy foresees that in the coming years, the consumption of electrical energy will grow by about 2% per year. In the production of electrical energy for 2023, 39.2% falls on TPPs, 40.2% on NPP ‘Kozloduy’, 7.7% on HPPs, 8.3% on photovoltaic power plants (PVPPs), and 3.9% on wind power plants (WPPs) [4].

On the other hand, the Energy Act of Bulgaria and the State Energy and Water Regulatory Commission guarantee a premium purchase of all electrical energy produced by RES at a preferential price, higher than the purchase price for the rest of the power plants in the country [5].

In connection with everything mentioned above, the goal of the paper is also set – to analyze data from the operation of a photovoltaic power plant, which is ultimately related to improving the behavior of these energy sources.

2. Basic theoretical statements in statistical processing and data analysis

Statistics is a mathematical subject that studies the extraction of information through the analysis and interpretation of empirical data using probability theory. Statistical activity also includes the planning and organization of data collection through surveys and experiments. Mathematical statistics arose in connection with the needs of empirical science and stand out from most branches of mathematics by its applied orientation. Statistical methods are applied in a wide range of fields, such as the natural and social sciences, public administration, and business [6].

One of the major subfields of mathematical statistics is descriptive statistics, which deals with the summarization of systematized data. It is of particular importance in empirical research and the description of experimental results. Through the methods of mathematical statistics, the data can be analyzed, and, based on this, conclusions can be drawn about the regularities contained in them.

Below are some definitions specific to mathematical statistics given [7].

- **Mass phenomenon** – when certain regularities valid for the community of units as a whole are repeated in many individual phenomena.

- **Statistical aggregation** – this is a collection of a large number of units (cases) that characterize the mass phenomenon.

- **Statistical signs** – express the properties (qualities, manifestations, and relations) of the individual units of a given phenomenon, and in general statistical signs are qualitative and quantitative.

▪ **Statistical units** – these are the individual units (cases, representatives) that form the statistical aggregation. Through statistical grouping, the individual statistical units are distributed into groups, distinguished based on the observed signs characterizing the aggregation.

▪ **Statistical data** – this is the collected, organized, and analyzed information necessary for the study of a given phenomenon.

▪ **Aggregation and sampling** – accumulated data can generally be thought of as measurements of certain properties of objects selected from some large *aggregation*, and information about certain characteristics of items is recorded by *sampling*.

▪ **Characteristics of the location** – there are several characteristics of the location of the empirical data set. The most common is the average (*arithmetic mean*) value, in short, it represents the value about which the entire empirical distribution can be ‘balanced’. For raw data (sample values) $x_1, x_2, \dots, x_n$, the *arithmetic mean* value is:

$$\bar{x} = n^{-1} \cdot \sum_{i=1}^n x_i \quad i = 1, 2, \dots, n \quad (1)$$

The *geometric mean* value can also be used as a characteristic of the location. In the case of sample volume n , the *geometric mean* is defined as:

$$(x_1 x_2 \dots x_n)^{\frac{1}{n}} \quad (2)$$

▪ Dispersion characteristics

A commonly used characteristic of the dispersion is the *arithmetic mean deviation* δ , which can be absolute or relative:

$$\bullet \text{ absolute } \delta = \frac{\sum |x_i - \bar{x}|}{n} \quad (3)$$

$$\bullet \text{ relative } V_{\delta} \% = \frac{\delta}{\bar{x}} \cdot 100 \quad (4)$$

One of the most commonly used characteristics of data dispersion is the *root mean square deviation* σ . First, the square of this quantity, called *dispersion* σ^2 , has to be determined:

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \quad (5)$$

The *root mean square deviation* is defined as the positive square root of the *dispersion*:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (6)$$

3. Results obtained

Following the set objective, below are presented processed data obtained from measurements of certain electrical quantities with static electronic electricity meters. Reported: voltage, grid frequency, active power, reactive power, solar radiation, air temperature, module surface temperature. The data is for a PVPP with peak power 25 MWp; location – N42.58674674328831, E25.46936206803762; electrical connection to the power system – 110 kV; 2 pcs available step-up transformers by 16 MVA rated power; a combined module (circuit breaker + disconnector) used to save space by obtaining a small field; single-linear circuit of a switchgear 20/110 kV; solar inverters type PVS800-57 (100÷500 kW); single PV modules type HR-240P-18-Bb with a power of 240 W each.

A general view of the considered PVPP is presented in Fig. 1, and in Fig. 2 – SCADA (control and monitoring system) screen.

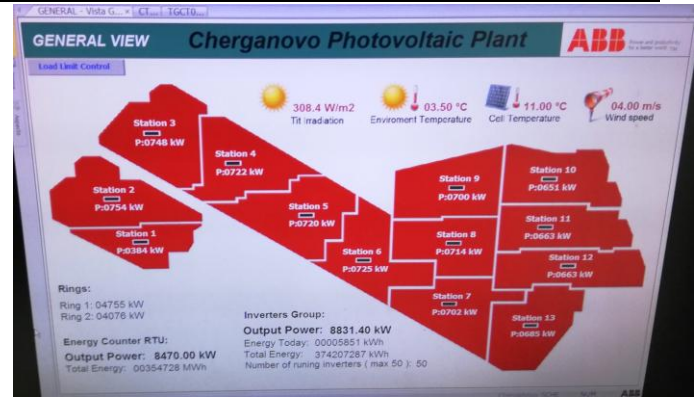


Fig. 1 General view of the PVPP considered.

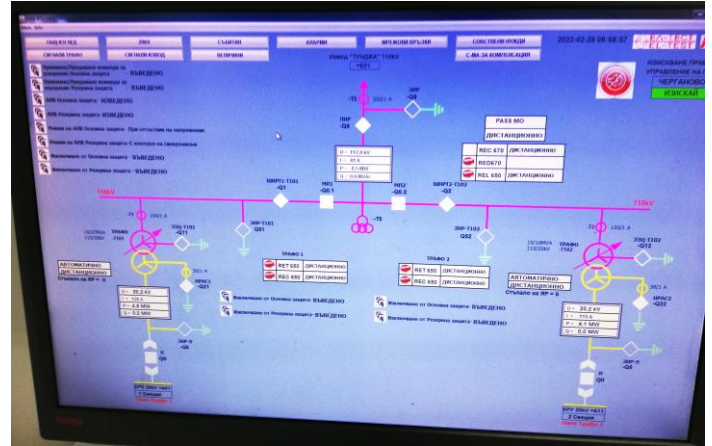


Fig. 2 SCADA screen for the PVPP considered.

Part of the data available for one year is given in the Appendix – only for the daylight part of a particular day, during which generation of electrical energy is possible. The measurement interval (recording or sampling frequency) is 15 minutes, which means a significant amount of data (4 values in 1 hour, 96 values per day).

Table 1 presents the results of the processing reported data during the operation of the PVPP, and the following figures show the graphical change of some quantities. The table and figures prepared, including calculation results, are implemented according to the rules of mathematical statistics [8]. Additionally, analytical expressions for approximating functions (polynomials) are derived, convenient for further calculations.

Table 1: Statistical indicators for the PVPP considered

Month (year 2020)	Total active power, P [MW]	Arithmetic mean value of active power, $\bar{P}$ [MW]	Absolute arithmetic mean deviation of the active power, δ [MW]	Relative arithmetic mean deviation of the active power, V_{δ} [%]	Root mean square deviation of the active power, σ [MW]
V	16124.02	520.13	494.12	94.99	512.69
VI	16915.05	563.84	535.64	94.99	553.02
VII	19020.63	613.57	582.89	94.99	601.46
VIII	19129.22	617.07	586.53	95.05	605.23
IX	16480.39	549.35	521.88	94.99	538.82

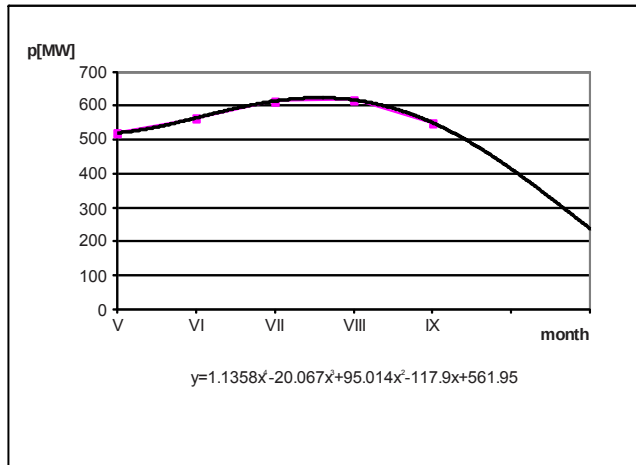


Fig. 3 Change in arithmetic mean value of active power (year 2020).

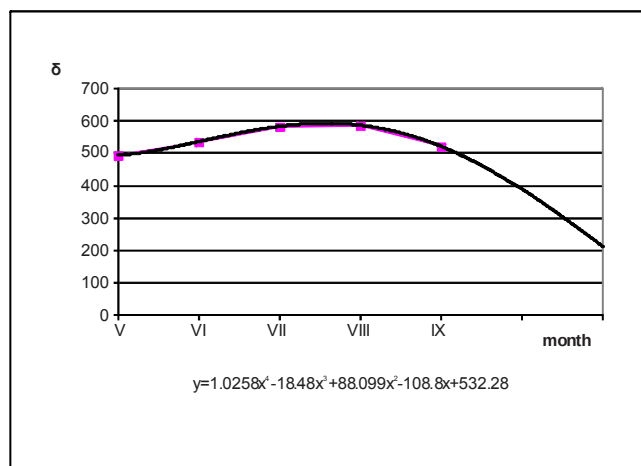


Fig. 4 Change in absolute arithmetic mean deviation of active power δ (year 2020).

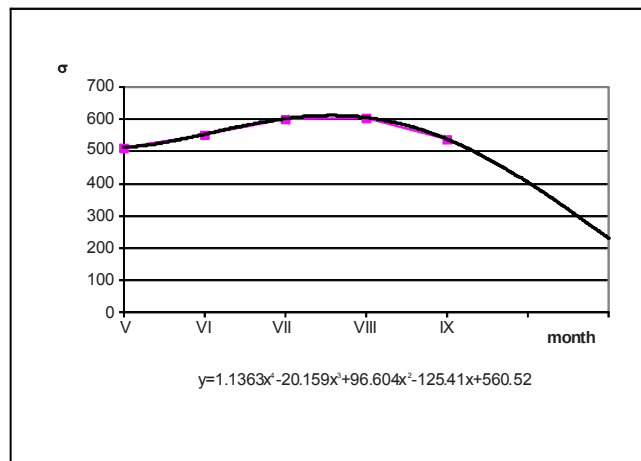


Fig. 5 Change in root mean square deviation of active power σ (year 2020).

4. Conclusions

The treated topic is current, as the results of some developments in this area have not yet been put into practice.

The processed real statistical set of statistical units, which are empirical data on the active and reactive power of a specific PVPP, allows certain regularities to be established based on the analysis.

The dispersion measures *arithmetic mean* and *root mean square deviation* have been determined. It is necessary to clarify that statistical treatments use the data values as they are, and zero is also a value, not the absence of data.

The root mean square and arithmetic mean deviations are the main characteristics of statistical distributions. These are called the parameters of the statistical distribution. The arithmetic mean deviation has the property of underestimating the differences compared to the standard deviation. The root mean square deviation acts as the most accurate and adequate measure of aggregation dispersion and is, therefore, the most widely used. Regarding dispersion, the aim is always to keep it to a minimum.

Despite some inconveniences in the use of RES, the production of 'green energy' and its storage for later use will be at the heart of the sustainable energy of the future.

5. Acknowledgments

The implementation of this paper is thanks to the support from:

- the University Project 2404S 'Applied mathematical research on energy and economic efficiency of electrical equipment in the conditions of energy transition', financed by the Ministry of Education and Science of the Republic of Bulgaria.
- the Project BG05M2OP001-1.002-0023 Competence Center 'Intelligent Mechatronic, Eco- and Energy-Saving Systems and Technologies' financed by the Operational Program 'Science and Education for Intelligent Growth' of the European Regional Development Fund.

6. References

1. <https://bg.wikipedia.org/wiki/>
2. <https://eur-lex.europa.eu/>
3. <https://commission.europa.eu/>
4. <https://www.eso.bg/>
5. <https://www.dker.bg/>
6. Moses, Lincoln E. Think and explain with statistics. ADDISON-WESLEY, 1986. pp. 1-3.
7. Moore, David. Teaching Statistics as a Respectable Subject. // Statistics for the Twenty-first Century. Washington, DC, The Mathematical Association of America, 1992. pp. 14-25.
8. Chance, Beth L. Investigating Statistical Concepts, Applications, and Methods. Duxbury Press, 2005.

Appendix**A piece of measured data for a specific time period**

PV Power Plant 25 MWp	Active power					Reactive power				
	Current Value	Average Value	Integral Value	Minimum Value	Maximum Value	Current Value	Average Value	Integral Value	Minimum Value	Maximum Value
02/08/2020 06:30:00	-0.06	-0.07	-0.02	-0.07	-0.06	-0.02	-0.02	-0.01	-0.04	-0.02
02/08/2020 06:45:00	-0.56	-0.28	-0.07	-0.56	0.00	-0.07	-0.02	-0.01	-0.07	-0.02
02/08/2020 07:00:00	-1.06	-0.76	-0.19	-1.06	-0.56	-0.14	-0.10	-0.03	-0.17	0.17
02/08/2020 07:15:00	-2.03	-1.68	-0.42	-2.08	-1.06	-0.11	-0.11	-0.03	-0.15	-0.07
02/08/2020 07:30:00	-3.62	-2.85	-0.71	-3.62	-2.03	-0.12	-0.25	-0.06	-0.43	0.05
02/08/2020 07:45:00	-4.98	-3.99	-1.00	-4.98	-3.61	-0.15	-0.15	-0.04	-0.38	0.15
02/08/2020 08:00:00	-6.12	-4.81	-1.20	-6.12	-3.42	0.01	-0.03	-0.01	-0.33	0.18
02/08/2020 08:15:00	-6.52	-6.67	-1.67	-6.88	-6.12	0.00	0.02	0.00	-0.35	0.20
02/08/2020 08:30:00	-8.00	-7.11	-1.78	-8.00	-5.70	0.00	0.05	0.01	-0.21	0.21
02/08/2020 08:45:00	-9.57	-8.76	-2.19	-9.57	-8.00	0.23	0.07	0.02	-0.22	0.45
02/08/2020 09:00:00	-11.20	-10.30	-2.58	-11.20	-9.57	0.25	0.17	0.04	-0.45	0.50
02/08/2020 09:15:00	-12.37	-11.80	-2.95	-12.37	-11.20	0.38	0.14	0.03	-0.38	0.75
02/08/2020 09:30:00	-13.34	-12.91	-3.23	-13.34	-12.37	0.40	0.30	0.07	0.00	0.89
02/08/2020 09:45:00	-14.51	-14.00	-3.50	-14.51	-13.34	0.42	0.59	0.15	0.00	1.34
02/08/2020 10:00:00	-15.73	-15.05	-3.76	-15.73	-14.51	0.88	1.03	0.26	0.42	1.58
02/08/2020 10:15:00	-16.86	-16.26	-4.06	-16.86	-15.73	1.66	1.46	0.37	0.54	1.83
02/08/2020 10:30:00	-17.78	-17.36	-4.34	-17.78	-16.86	1.91	1.80	0.45	1.47	1.99
02/08/2020 10:45:00	-18.63	-18.26	-4.57	-18.63	-17.78	2.12	2.10	0.53	1.81	2.28
02/08/2020 11:00:00	-19.13	-19.03	-4.76	-19.25	-18.63	2.51	2.39	0.60	2.12	2.64
02/08/2020 11:15:00	-19.77	-19.48	-4.87	-19.77	-19.13	2.54	2.65	0.66	2.43	2.82
02/08/2020 11:30:00	-20.60	-20.08	-5.02	-20.60	-19.77	2.92	2.82	0.70	2.54	2.98
02/08/2020 11:45:00	-21.60	-20.95	-5.24	-21.60	-20.44	3.42	3.09	0.77	2.86	3.42
02/08/2020 12:00:00	-18.68	-20.94	-5.24	-22.23	-18.52	3.88	3.54	0.88	3.16	3.95
02/08/2020 12:15:00	-17.48	-15.11	-3.78	-20.96	-9.65	2.57	3.63	0.91	2.49	4.38
02/08/2020 12:30:00	-22.26	-21.61	-5.40	-23.37	-17.37	2.36	1.61	0.40	0.00	3.59
02/08/2020 12:45:00	-12.02	-19.03	-4.76	-23.18	-10.86	4.35	4.09	1.02	2.18	5.03
02/08/2020 13:00:00	-22.78	-18.72	-4.68	-23.17	-10.83	0.49	3.13	0.78	0.00	4.88
02/08/2020 13:15:00	-23.17	-22.42	-5.61	-23.24	-20.86	4.68	2.99	0.75	-0.41	4.90
02/08/2020 13:30:00	-23.36	-15.53	-3.88	-23.36	-6.07	4.74	4.49	1.12	3.47	4.95
02/08/2020 13:45:00	-7.78	-15.00	-3.75	-23.36	-7.78	4.89	2.29	0.57	-0.22	5.02
02/08/2020 14:00:00	-16.32	-10.11	-2.53	-16.32	-7.51	0.23	1.83	0.46	-0.49	4.98
02/08/2020 14:15:00	-13.14	-19.95	-4.99	-24.06	-12.77	1.96	0.26	0.06	-0.44	1.96
02/08/2020 14:30:00	-19.44	-12.09	-3.02	-19.44	-7.09	0.51	3.53	0.88	0.50	5.62
02/08/2020 14:45:00	-21.94	-15.73	-3.93	-21.94	-12.59	2.89	0.92	0.23	-0.45	2.89
02/08/2020 15:00:00	-9.74	-15.81	-3.95	-23.34	-8.10	4.12	1.71	0.43	0.00	4.12
02/08/2020 15:15:00	-14.18	-11.29	-2.82	-16.53	-7.31	0.25	2.29	0.57	-0.22	4.99
02/08/2020 15:30:00	-18.87	-15.67	-3.92	-22.02	-10.32	1.41	0.64	0.16	-0.43	2.06
02/08/2020 15:45:00	-9.39	-13.71	-3.43	-18.87	-9.39	2.76	1.86	0.46	0.00	4.23
02/08/2020 16:00:00	-4.74	-5.23	-1.31	-9.39	-4.60	0.25	1.15	0.29	0.00	2.76
02/08/2020 16:15:00	-4.90	-4.80	-1.20	-4.95	-4.66	0.00	0.02	0.00	-0.18	0.25
02/08/2020 16:30:00	-4.02	-4.69	-1.17	-5.30	-3.83	0.00	0.00	0.00	-0.35	0.18
02/08/2020 16:45:00	-6.35	-7.47	-1.87	-13.77	-4.02	-0.15	-0.01	0.00	-0.33	0.18
02/08/2020 17:00:00	-4.30	-5.68	-1.42	-8.55	-4.25	0.20	0.17	0.04	-0.35	1.37

Device and technology for the production of foam granules of "breathable" composite material based on heat-treated rice husks

Lyuben Lakov*, Bojidar Jivov, Marieta Gacheva, Krasimira Toncheva

Bulgarian Academy of Sciences, Institute of Metal Science, Equipment and Technologies with Hydro and Aerodynamics Centre
"Acad. A. Balevski", 67 "Shipchenski prohod" Blvd., 1574 Sofia, Bulgaria, e-mail: silicatchim@mb.bia-bg.com

Abstract: An innovative device and a cost-effective and environmentally friendly technology have been developed to produce porous composite materials by using two types of waste raw materials: residual biomass from agricultural production and municipal waste. Heat-treated rice husk and ground waste glass were used as main components. The technological approach presented here allows for the preparation of "breathable" porous materials with the predominant presence of open pores and increased specific surface area. A promising possibility for the application of the developed materials is their use in bulk or in the preparation of various composite building elements based on inorganic binders. Recipes and an adequate technological regime have been developed taking into account the design characteristics of the production facility, the specificity of the waste raw materials used and the intended target performance of the products.

Keywords: WASTE BIOMASS, RICE HUSKS, MUNICIPAL WASTE, "BREATHABLE" MATERIALS

1. Introduction

The increasing productivity of agricultural production is accompanied by increasing amounts of residual biomass obtained from the processing of different crops. A variety of technological approaches for the efficient treatment, recovery and recycling of heterogeneous renewable agricultural waste (a potential environmental pollutant) of different origins are being applied [1,2]. In this respect, rice husks [3] are a specific waste product characterized by low bulk density, which makes it difficult to store and transport. Untreated rice husks are mainly composed of cellulose, hemicellulose, lignin, other organic substances and hydrated oxide compounds, etc. [4]. After heat treatment at a temperature of 800°C [5] to the state of white rice husk ash, the chemical composition of the resulting product is characterized by the dominant presence of SiO₂ (mostly amorphous), as well as other oxide components: Al₂O₃, K₂O, Fe₂O₃, MgO, CaO, Na₂O, SO₃, etc.

Rice husks and various products obtained from their technological processing are used in various fields [6-14]: the brewing and culinary industries, cosmetics, animal husbandry, energy and chemical industries, metallurgy, the production of construction materials and household articles and many others.

Technological schemes have been developed for obtaining various products from rice husks [15-25]: mesoporous adsorbent (carbon/SiO₂) for the removal of spilled oil and oil products, technological additives for increasing the strength of cements and concretes, biofuels, pigments for ceramic products, fillers for elastomeric mixtures and rubber products, polycrystalline silicon, silicon-containing materials - SiC, Si₃N₄, Si₂N₂O, (Na,K)₂SiO₃ and others.

One possibility for cost-effective utilization of rice husks [6,9,11] is the production of pellets, briquettes and granulates (with high bulk densities). The production of rice husk granules requires relatively limited investment, can be realized using basic machinery, without the need for high-tech equipment and is associated with relatively low operating costs. Rice husk can be used as fuel in various building heating systems, combustion installations [6,11] of production plants, etc.

Pyrolysed (under certain conditions) rice husks have been found to be applicable as a material for flame retardant additives and coatings of foundry moulds in the production of small and medium size steel and cast iron castings [4].

A useful approach to the use of rice husks is the production of a wide range of household utensils, crockery and other reusable products [11] that can be ecologically degraded under natural conditions after a certain service life.

Waste glass is a widespread technogenic waste product in the urban environment, practically undegradable in a natural environment [26]. The continuous generation of significant quantities of waste materials is a prerequisite for the analysis and re-evaluation of existing traditional and standardized approaches for optimal waste management [27,28]. Due to the excellent physicochemical characteristics of glassy waste materials, their

recycling as secondary raw materials to produce various products has been adopted as an appropriate environmental and economic approach [29-35].

The aim of this article is to present an innovative facility and technology for the production of "breathable" porous composite materials based on two types of waste raw materials: heat-treated residual product of agricultural origin (rice husks) and glass household waste (mainly silicate glass).

2. Results and discussion

The development is based on an invention patent [14], on the analysis carried out on the trends in the ecological utilization of waste rice husks and on the technological experience in the field of foam silicate materials production.

The production facility [14] consists of a hopper with a technological additives feeder 1 connected to a technological additives pipeline 3 and a hopper with a raw material feeder 2 connected to a raw material pipeline 4. The tubes of the pipelines are shaped in the form of longitudinal spirals and are loosely connected to the upper end of an inclined rotating drum 6. The hoppers 1 and 2 are movably connected to the upper end of the rotating drum 6 by a housing 5. To the inner side of the rotatable drum 6 is mounted an inner tube 10, to which spiral ribs 11 are attached in zones 'A' and 'B' (fig. 1). On the rotatable drum 6, in zone 'C' spiral ribs 12 are also mounted on its inner side and straight ribs 13 are mounted on its outer side. Below the lower end of the rotating drum 6 is a receiving hopper 16, loosely connected to a system of several separator vibratory screens 15, below which is a set of hoppers 14. The first of them, which collects the finest fractions, such as technological additives, is connected via a product line 17 to the hopper with a technological additives feeder 1. The space between the rotating drum 6 and the inner tube 10 is filled with insulating material 7. To the inner tube 10 is coaxially mounted a combustion tube 8, which is movably connected to a gas burner 22 at one end and a guide body 9 is fixedly attached to its other end. The gas burner 22, the housing 5, the rotating drum 6 and a drive mechanism 18 of the rotating drum 6 are mounted on a tilting frame 21. The frame is pivotally connected to a base 19, and a lifting mechanism 20 is mounted between them, providing an adjustment of the inclination of the frame 21 by a certain angle α . A pyrometer 23 is mounted at the lower end of the housing 5, in the direction of the interior of the rotating drum 6.

The operating mode of the facility consists for a series of technological stages [14]. Loading of hopper with a technological additives feeder 1 and hopper with a raw material feeder 2. Setting in motion the rotating drum 6 and starting the gas burner 22. Starting the feeders of the hoppers 1 and 2, when the required temperature is reached in the rotating drum 6 (as determined by the pyrometer 23). The technological additives and raw material pass through the longitudinal spiral tubes of pipelines 3 and 4 (which are partially heated) and enter the rotating drum 6. Here, in zone "A", marked "Active", the heating of the raw material and the

technological process additives takes place and at the end of zone "A" the raw material is foamed. The heating is combined: radiantly from the heated combustion tube 8, by convection from the hot gases and by thermal conduction through the contact with the heated surfaces of the inner tube 10 and the spiral ribs 11. The gases released through the thermal decomposition of the foaming agent leave the combustion tube 8, changing direction of movement on contact with the guide body 9. Passing through the housing 5, the gases give up some of their energy to the longitudinal spiral tubes of pipelines 3 and 4 and are discharged for further use. At the same time, the foamed raw material and the technological additives enter zone "B", named "Passive", where they are subjected to tempering. Entering zone "C", the resulting products are cooled to the required temperature and conveyed to the receiving hopper 16, and from there to the set in motion separator vibrating screens 15 and the corresponding hoppers 14. The system of separator vibratory screens 15 distributes to the set of hoppers 14 the resulting quantities of foamed granules of composite material and technological additives according to their fractional composition. The spent additives (including the used anti-caking agent) collected in the first hopper (of the hopper set 14) are fed back through the product line 17 into the hopper with a technological additives feeder 1 and reintroduced into the process cycle. The parameters of the technological process are controlled by adjusting the feeders of hoppers 1 and 2, the temperature in the combustion tube 8 (by means of controlling the operation of the gas burner 22) and the inclination angle α and the speed of the rotating drum 6.

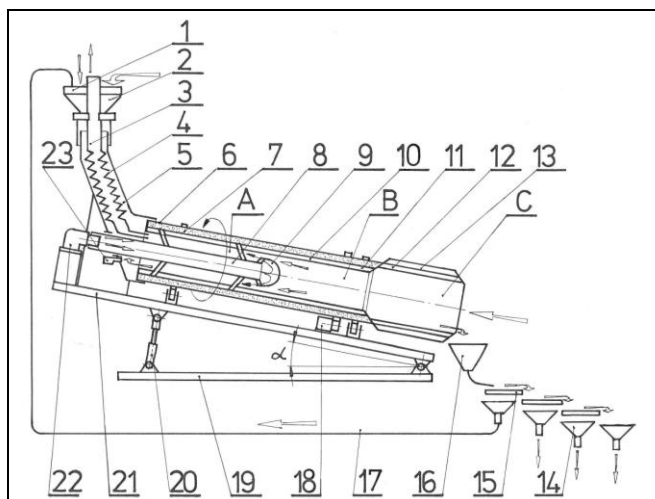


Fig. 1. Scheme of a device for obtaining foamed granules of composite material.

After analyzing the production capabilities of the considered device, a modified experimental prototype of a facility (next version, fig. 2, 3, 4) with improved structural and technological characteristics was developed and prepared.



Fig. 2. Modified prototype of a facility for obtaining foam granules of composite material.



Fig. 3. Modified prototype of a facility for obtaining foam granules of composite material.



Fig. 4. Modified prototype of a facility for obtaining foam granules of composite material.

The necessary quantities of unprocessed rice husks and household glass waste (mainly standard glass packaging like jars, bottles, etc.) were provided to carry out the complete technological cycle.

Under a limited temperature treatment of rice husks up to 600°C (under controlled pyrolysis conditions), the volatile compounds released partially condense on the surface of the resulting solid pyrolysis residue. The product thus formed is a hybrid composite structure of dominant SiO₂ (mostly in amorphous form), other oxides, solid carbon and residual amounts of volatile hydrocarbons.

The primary technological processing of the prepared glass waste materials involves crushing to pieces (crumbs) and grinding to a powder fraction with a suitable granulometric composition.

Recipes and composite raw materials based on a heat-treated rice husk product (from 2 to 20 %) and a powder fraction of ground waste glass were prepared.

The prepared samples are subjected to processing in the prepared modified prototype (fig. 2, 3, 4) by applying a developed adequate temperature regime of several stages: 650°C, 820°C, 820°C and from 570°C to the final cooling temperature.

The designed facility and the used technological approach allow for the obtaining of different foam aggregate fractions of high porosity "breathable" composite materials. Under the thermal decomposition of the residual organic products, the composition of the rice husk produces a characteristic porous structure (with a predominant pore size of up to 600 µm) and a predominant presence of open pores (more than 80 %). At the same time, a significant presence of micro-pores with a size of up to about 50 µm (concentrated in the walls of the main pores) was found, further increasing the high specific surface area of the material and its sorption properties.

The main potential application of the obtained individual granule fractions is their use as energy-saving, thermal and acoustic insulation in bulk or for the preparation of lightweight concretes and an assortment of composite elements for the construction industry (shapes, panels, etc.) involving inorganic binders.

3. Conclusions

A general study of the main issues and existing technological options for the efficient utilization of waste rice husks was carried out. An original facility and a promising technology for the production of porous composite materials based on heat-treated rice husks and waste glass in the form of ground powder fraction are presented.

A modified experimental prototype with improved design and operational characteristics was prepared. The designed facility allows for variation of the parameters of the technological process according to the specific recipe compositions, the fractional composition of the raw product and the planned target performance of the final products. The proposed technological scheme makes it possible to obtain "breathable" porous materials with increased specific surface area, predominant presence of open pores (over 80%) and also presence of micro-pores.

The granule fractions thus obtained are potentially applicable in bulk for thermal insulation or as components for the preparation of lightweight concretes and building composites involving inorganic binders. The presented development is in line with the current trends in the implementation of environmentally friendly technologies and allows for simultaneously utilization two types of waste materials and producing sustainable and durable products.

Acknowledgments

The authors express their gratitude to the Bulgarian National Innovation Fund and the Bulgarian Small and Medium Enterprises Promotion Agency (BSMEPA) for funding the scientific and applied project / contract № 13 IF-02-21/12.12.2022 between contracting authority BSMEPA and contractors - beneficiary "MAG" Ltd. with partner Institute of Metal Science, Equipment and Technologies with Hydro- and Aerodynamics Centre "Acad. A. Balevski".

4. References

1. A. Faaij, Modern biomass conversion technologies, Mitig. Adapt. Strat. Glob. Change, 11, 2006, pp. 335–367.
2. S.R. Decker, J. Sheehan, D.C. Dayton, J.J. Bozell, W.S. Adney, B. Hames, S.R. Thomas, R.L. Bain, S. Czernik, M. Zhang, Biomass Conversion, Kent and Riegel's Handbook of Industrial Chemistry and Biotechnology, Springer, 2007. pp. 1449–1548.
3. A. Yank, M. Ngadi, R. Kok, Physical properties of rice husk and bran briquettes under low pressure densification for rural applications. Biomass Bioenergy, 84(1), 2016, pp.22–30.
4. R. I. Petrov, V. B. Toteva, R. N. Nikolov, Investigation of possibility of application of carbonized rise husks as an antiburnt-on sand mould coat, Proceedings of the XXIII International scientific and technical conference Foundry 2016, 13–15. April 2016, Pleven, Bulgaria, Year XXIV, Issue 7 (193), 2016, pp. 19-21.
5. R. Putra Jaya, M. Al Amin Muhamad Nor, Z. Arifin Ahmad, Z. Mohd Amin, Properties of mortar containing rice husk ash at different temperature and exposed to aggressive environment, Advanced Materials Research Vol. 620, 2013, pp 87-93.
6. A. Jain, R. T. Rao, S. S. Sambi, P. D. Grover, Energy and chemicals from rice husk, Biomass and Bioenergy, 7, 1995, pp. 285–289.
7. L. Sun, K. Gong, Silicon-based materials from rice husks and their applications, Ind. Eng. Chem. Res. 40, 2001, pp. 5861–5877.
8. M. R. Taib, A. Johari, S. Ngo, A. Salema, K. Tan, Utilisation of rice husk waste and its ash, Part 2, Ingineur, 42, 2009, pp. 37–41.
9. C. A. M. Moraes, I. J. Fernandes, D. Calheiro, A. G. Kieling, F. A. Brehm, M. R. Rigon, J. A. Berwanger Filho, I.A. Schneider, E. Osorio, Review of the rice production cycle: by-products and the main applications focusing on rice husk combustion and ash recycling, Waste Manag. Res. 32, 2014, pp. 1034–1048.
10. N. Soltani, A. Bahrami, M.I. Pech-Canul, L.A. González, Review on the physicochemical treatments of rice husk for production of advanced materials, Review, Chemical Engineering Journal, 264, 2015, pp. 899–935.
11. Y. Zou, T. Yang, Rice husk, rice husk ash and their applications, In Rice bran and rice bran oil, Chapter 9 (1), 2019, pp. 207-246.
12. Madhu Vershit CH, Amandeep Kaur, Lomada Vasudha, Exploitation of rice husk and ash: A review, The Pharma Innovation Journal, 12(3), 2023, pp. 5203-5208.
13. L. B. Zhengwuvi, A. M. Usman, R. M. Joshua, S. Abdulhamid, Potentials of rice husk ash utilisation as an alternative source of hybrid materials for engineering application: a critical review, Nigerian Journal of Engineering Science and Technology Research, Vol. 9, No. 2, 2023, pp. 199-211.
14. I. Chorbov, L. Lakov, K. Toncheva, L. Drenchev, N. Guo, H. Shi, "Device for obtaining foamed granules of composite material", Application date 21/10/2020, Patent № 67553 B1, 17.07.2023.
15. B. K. Padhi, C. Patnaik, Development of Si₂N₂O, Si₃N₄ and SiC ceramic materials using rice husk, Ceramics International, 21, 1995, pp. 13–20.
16. J. Watari, A. Nakata, T. Torikai, M. Yada, Fabrication of porous SiO₂/C composite from rice husks, Journal of the European Ceramic Society 26, 2006, pp. 797–801.
17. S. Kumagai, Y. Noguchi, Y. Kurimoto, K. Takeda, Oil adsorbent produced by the carbonization of rice husks, Waste Management, 27, 2007, pp. 554–561.
18. I. Uzunov, S. Uzunova, A. Gigova, D. Klisurski, P. Tsokov, D. Angelova, Synthesis of biogenic carbon/silica material by pyrolysis of rice husks, Comptes Rendus de L'Academie Bulgare des Sciences 62, 2009, pp. 570-578.
19. D. Angelova, S. Uzunova, S. Staykov, I. Uzunov, Preparation of a biogenic carbon/silica based adsorbent for removal of petroleum products spills from aqueous medium, J. Univ. Chem. Tech. Metal. 45, 1, 2010, pp. 25-32.
20. D. Radev, I. Uzunov, Nanosized Silicon Carbide Obtained from Rice Husks, Solid State Phenomena, 159, 2010, pp. 153-156.
21. M. F. M. Zain, M. N. Islam, F. Mahmud, M. Jamil, Production of rice husk ash for use in concrete as a supplementary cementitious material, Construction and Building Materials, vol. 25, 2011, pp. 798-805.

22. A. Boonmee, P. Sabsiroht, K. Jarukumjorn, Preparation and characterization of rice husk ash for using as a filler in natural rubber. *Materials Today Proceedings*, vol. 17, 2019, pp. 2097-2103.
23. Z. Zeng, Y. Zhao, P. Li., H. Yu, Fabrication of rice husk ash/natural rubber composites by the latex process. *Journal of Wuhan University of Technology (Materials Science Edition)*, vol. 35, no. 1, 2020, pp. 42-46.
24. M. Aleksandrova M., V. Petkov, V. P. Korzhov, I. S. Zheltyakova. "Study the influence of immersion in the synthesis of thin layers on a composite substrate", *International scientific journal INDUSTRY 4.0*, VII, 6, Scientific technical union of mechanical engineering "INDUSTRY 4.0" Bulgaria, 2022, pp. 226-228.
25. V. Petkov, M. Aleksandrova, V. Petkov, D. Teodosiev, A. Bouzekova-Penkova, „Investigation of the glassy carbon coating deposited on the titanium alloys, microstructure and mechanical properties”, *Journal of Theoretical and Applied Mechanics*, Vol. 52, No 4, 2022, pp. 381-392.
26. I. M. Pongoh, Y. I. Masjud, The effect of glass waste on climate change, *The Journal of Environment Conflict EnvironC*, Volume 1, Issue 1, 2024, pp. 64–69.
27. L. Giusti, “A review of waste management practices and their impact on human health”, *Waste Management*, 29(8), 2009, pp. 2227-2239.
28. E. Amasuomo, J. Baird, “The Concept of Waste and Waste Management”, *Journal of Management and Sustainability*; Vol. 6, No. 4; 2016, pp. 88-96.
29. T. C. Ling, C. S. Poon, S. C. Kou, Feasibility of using recycled glass in architectural cement mortars, *Cem. Concr. Compos.* 33, 2011, pp. 848–854
30. T. C. Ling, C. S. Poon, Properties of architectural mortar prepared with recycled glass with different particle sizes, *Mater. Des.* 32, 2011, pp. 2675–2684.
31. S. P. Cui, J. G. Zhang, Y. L. Tian, S. B. Sun, “Generation Review on the Production Line Development of Foam Glass at Home and Abroad”, *Advanced Materials Research*, Vol. 915-916, 2014, pp. 524-531.
32. L. Lakov, B. Jivov, M. Aleksandrova, Y. Ivanova, K. Toncheva, „An innovative composite material based on sintered glass foam granules”, *Journal of Chemical Technology and Metallurgy*, 53, 6, 2018, pp. 1081-1086.
33. L. Lakov, B. Jivov, Y. Ivanova, S. Yordanov, M. Marinov, S. Rafailov, “Composite Materials Obtained from Foamed Silicate Products”, *International Journal “NDT Days”*, Volume II, Issue 2, Year 2019, pp. 188-194.
34. T. O. Ogundairo, D. D. Adegoke, I. I. Akinwumi, O. M. Olofinnade, Sustainable use of recycled waste glass as an alternative material for building construction - A review, 1st International Conference on Sustainable Infrastructural Development, 2019, IOP Conf.Series: Materials Science and Engineering 640(2019)012073.
35. D. Patel, R. P. Tiwari, R. Shrivastava, R. K. Yadav, Effective utilization of waste glass powder as the substitution of cement in making paste and mortar. *Construction and Building Materials*, 199, 2019, pp. 406-415.

Experimental self-cleaning glass coatings for photovoltaic systems placed in a real environment

Stancho Yordanov¹, Tatyana Ivanova², Mihaela Aleksandrova¹, Vladimir Petkov¹

¹ Bulgarian Academy of Sciences, Institute of Metal Science, Equipment and Technologies with Hydro- and Aerodynamics Centre "Acad. Angel Balevski", 67 "Shipchenski prohod" Blvd., 1574 Sofia, Bulgaria

² Central Laboratory of Solar Energy and New Energy Sources, Bulgarian Academy of Sciences, Tzarigradsko chaussee 72, 1784 Sofia, Bulgaria
stanco14@abv.bg

Abstract: Experimental vitreous samples with the presence of self-cleaning coatings applied by the sol-gel method were selected, which were tested in a real environment under cyclic atmospheric conditions in different TPPs / TPP-Sofia and TPP-Republika, Pernik/. The resulting thin films were studied by UV-NIS - VIR spectroscopy, AFM, contact angle, etc. The experimental samples are potentially applicable in the operation of photovoltaic panels with increased efficiency and sustainability in the conditions of an industrial environment and under the influence of various meteorological factors.

Keywords: REAL ENVIRONMENT, PHOTOVOLTAIC SYSTEMS, SURFACE COATING, HYDROPHILICITY, TiO_2

1. Introduction

A source of energy loss in photovoltaic (PV) modules is caused by reflection from the glass surface of the PV. Reflection from the cover glass causes a loss of ~4% of the air-glass interaction. [1]. Self-cleaning surfaces have gained prominence in recent years in the field of solar energy. These surfaces are used to avoid the accumulation of dirt on photovoltaic (PV) modules. TiO_2 has been widely used due to its photocatalytic activity and photoinduced superhydrophilicity. However, this oxide has some limitations, as it reduces the permeability of the glass and rapidly recovers the contact angle of water in a dark environment. To circumvent these limitations, $\text{TiO}_2/\text{SiO}_2$ [2] composites have been proposed. For a photovoltaic application, in addition to good transparency in the wavelength range 300–1800 nm and self-cleaning properties, the coating must also have long durability and adequate adhesion to withstand external conditions. The sol-gel method has the advantage of offering small and large area deposition at low cost [3]. It is important that AR /anti-reflection/ coatings are applied to the front glass surface of the PV modules if the output power is to be optimized. The majority of commercial AR coatings are monolayer porous SiO_2 , deposited by sol-gel methods. These coatings are effective in reducing reflection losses of between 2% and 3% absolute compared to uncoated glass [4, 5, 6].

The test glass samples with self-cleaning coating were laid down for a long time on the territories of TPP "Sofia" and TPP "Republika". TPP "Republika" is a combined heat and power plant in Pernik, Bulgaria belonging to "Toplofikatsia" - Pernik. The plant mainly uses brown coal mined in Pernik. Its capacity for electricity production is 105 MW, and for thermal energy – 204 MW. As of 2022, the plant is a permanent air polluter with repeatedly established violations.

Sofia Thermal Power Plant is a combined heat and power plant in Sofia, Bulgaria, owned by Toplofikatsia Sofia. Its main purpose is heating the central and northern part of the city, and the main fuel is natural gas. There is a possibility to produce electricity in 4 turbine units with a total capacity of 125 MW. [8]

2. Experimental procedures

In laboratory conditions, a series of experimental coatings on glass substrates were prepared by changing the ratio of the included components, the speed of application of the solution on the glass substrate and the mode of heat treatment.

All compounds were supplied as analytical grades and were used without addn modification. Thin film coatings were performed on slides by AGC Glass Europe. TiO_2 precursors and solvents are 99% titanium butoxide - $\text{Ti}(\text{OBu})_4$ and 99.9% ethanol ($\text{C}_2\text{H}_5\text{OH}$), all purchased from Alfa Aesar Germany. The hydrolysing agent is distilled water and the stabilizer is nitric acid / HNO_3 /. Preparation of the glass substrate. Before being cleaned, the panes were cut in

size (20 mm x 70 mm). Substrate cleaning is important prior to the deposition process, as this allows the cleaning and removal of submicron particles from the substrate surface. It is decisive for the quality of the adhesive film on the substrate. The glass coasters were cleaned/washed/ in 99.9% ethyl alcohol for 15 min and acetone for 15 min and then dried for 30 min at 100°C. Cleaning glass substrates improves coating adhesion material to the substrate and prevents contamination on the formed coating.

Preparation of TiO_2 solution. The goal of the synthesis is to obtain TiO_2 coatings on glass substrates applying the sol-gel method. Precursor mixture – 5 ml 99 % $\text{Ti}(\text{OBu})_4$ and 5 ml of DEA / $\text{C}_4\text{H}_{11}\text{NO}_2$ /– is added dropwise to the acidic solution composed of ethanol, water and nitric acid in a 1:1:0.5 ratio to prevent gel formation. To ensure adequate hydrolysis, the mixture was continuously stirred at 500 rpm with a magnetic stirrer for 2 h under air pressure to obtain pink translucent solution.

Dip coating machine (Dip Coating technology) with a speed of 42.8 mm/min is used to deposit thin layers of TiO_2 on the substrates. The lower speed of the applied solution on the glass substrate helps the uniform coating, while higher speed helps the solution become thinner. The deposited films are then heated at a temperature 400°C for 2 hours

Applying a coating of titanium oxide a few nm thick to the outside of the glass creates two mechanisms by which the glass is self-cleaning. The first is the photocatalytic effect, where ultraviolet rays catalyze the breakdown of organic compounds on the surface. The second is the hydrophilizing effect, in which water forms a thin layer, washing away the destroyed organic compounds.

3. Results and Discussion

Optical characterization was performed with UV–VIS–NIR Shimadzu 3600 double-beam spectrophotometer (Shimadzu Corporation, Kyoto, Japan) in the spectral region of 240-1800 nm. The transmittance spectra of the samples were taken against air. The reflectance spectra were measured by using the specular reflectance attachment (5° incidence angle) and Al coated mirror as reference.

AFM imaging was performed on the NanoScope V system (Bruker Ltd, Germany) operating in tapping mode in air at room temperature. We used silicon cantilevers (Tap 300Al-G, Budget Sensors, Innovative Solutions Ltd. Bulgaria) with aluminum reflective coatings with a thickness of 30 nm. According to the manufacturer's specifications, the cantilever spring constant and resonant frequency are in the range of 1.5 to 15 N/m and 150 ± 75 kHz, respectively. The tip radius of curvature is less than 10 nm. The scan rate was set at 1 Hz and images were taken in the highest possible resolution mode of the AFM 512 × 512 pixels in JPEG format. NanoScope software was used for cross-sectional and roughness analysis of all images.

Figure 1 shows the transmission spectrum of a clean glass substrate. The figures 2 and 3 shows the transmission and reflection spectrum of an experimental coating on a glass substrate in the region 300 - 800 nm and 400 - 1800 nm /on sample Ti_9/.

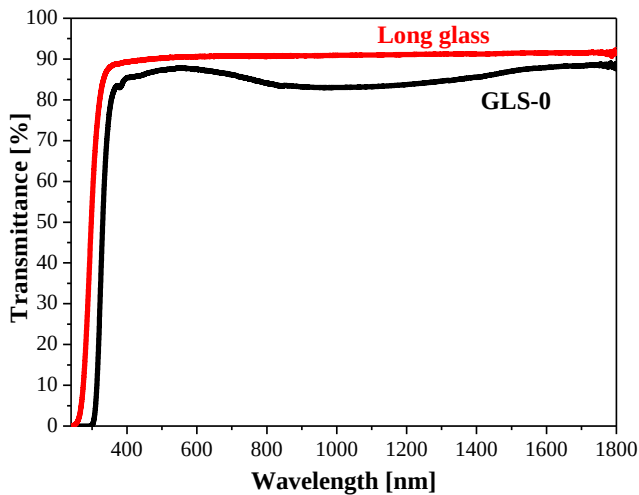


Figure 1. The figure shows the transmission spectrum of a clean glass substrate /GLS-0/

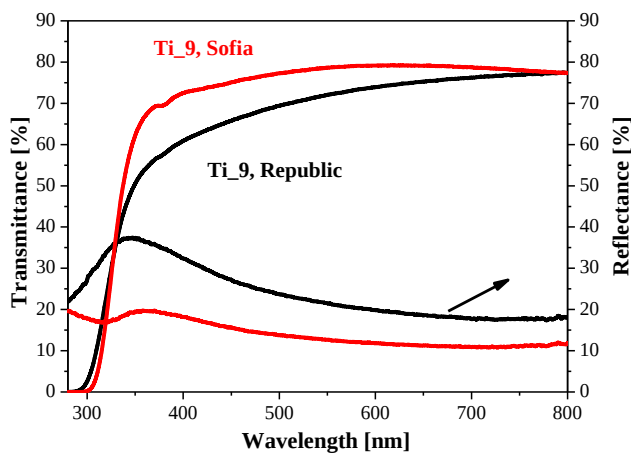


Figure 2. The figure shows the transmission and reflection spectrum of an experimental coating on a glass substrate in the region 300 - 800 nm /Ti_9/

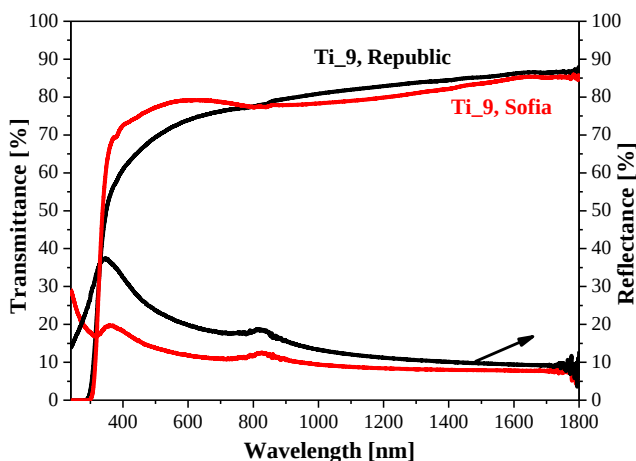


Figure 3. The figure shows the transmission and reflection spectrum of an experimental coating on a glass substrate in the region 400 - 1800 nm /Ti_9/

From the optical characteristics obtained after staying /30 days/ in a real production environment of the territories in TPP-Sofia and TPP-Republika, town of Pernik - under cyclical atmospheric conditions no decrease in permeability is observed /fig.1-3/.

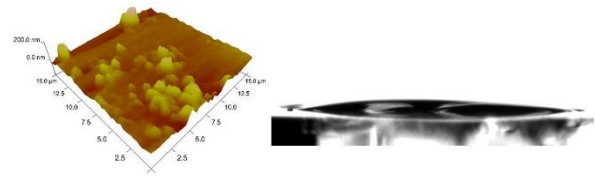


Figure 4. 3D AFM image and contact angle of coated glass /Ti_9/

The AFM image and contact angle of coated glass /Ti_9/ are shown in Figure 4. From the AFM measurements, it was found that the surface topography of the samples was due to the coating on the glass surface. Different structures ranging from 50 nm to 150 nm are observed from the 3D image.

Roughness values for Ra and Rq of coated glass sample:

Ti_9.

Sample	Roughness R_q (nm)	Roughness R_a (nm)
Ti_9	27.8	18.1

Contact angle measurements show that the Ti_9 coating is superhydrophilic with a contact angle of 8.3 degrees /8.3°/. At the measured surface contact angle, a spreading process of the water drop on the glass surface is observed. When sol gel coatings on glass containing TiO_2 surfaces are exposed to ultraviolet (UV) light, they exhibit superhydrophilicity. UV irradiation generates electron-hole pairs on the TiO_2 surface. These electron-hole pairs can react with adsorbed oxygen and water molecules, creating a highly hydrophilic surface. Electrons in the valence band of TiO_2 are excited to the conduction band under UV irradiation conditions, leading to superhydrophilic behavior. This property is particularly useful for self-cleaning surfaces.

The TiO_2 coated glass surface has a large contact angle with water /8.3°/. After UV irradiation, the contact angle of water can be reduced to less than 5 degrees or even to 0 degrees (ie, the water droplets are completely dispersed on the TiO_2 surface). It exhibits a very strong hydrophilicity. The superhydrophilicity of the surface can be maintained from several hours to about 1 week. It then slowly returns to the hydrophobic state before irradiation. Periodic UV irradiation can maintain the superhydrophilic state of the surface.

4. Conclusions

Experimental samples of vitreous body with the presence of self-cleaning coatings were tested in a real production environment under cyclic atmospheric conditions in different TPPs / TPP-Sofia and TPP-Republika, Pernik/. The preparation of the coatings was carried out at different ratios of the included components, speed of application of the solution on the glass substrate and different curing temperature and retainer. They were studied by UV-NIS - VIR spectroscopy, AFM, contact angle, etc. of the resulting thin layers. From the optical characteristics obtained after staying /30 days/ in a real production environment of the territories in TPP-Sofia and TPP-Republika, town of Pernik - under cyclical atmospheric conditions no decrease in permeability is observed.

The experimental samples are potentially applicable in the operation of photovoltaic panels with increased efficiency and sustainability in the conditions of an industrial environment and under the influence of various meteorological factors.

5. References

1. G. Womack, Kenan Isbilir, J. Walls and all, The performance and durability of single-layer sol-gel anti-reflection coatings applied to solar module cover glass, *Surface and Coatings Technology*, Volume 358, (2019), Pages 76-83
2. M. Moraes, A. Mello Ferreira, G. Timò, Superhydrophilic self-cleaning surfaces based on TiO_2 and $\text{TiO}_2/\text{SiO}_2$ composite films for photovoltaic module cover glass, *Applied Adhesion Science*, (2015), 3(5),1-9,DOI:10.1186/s40563-015-0034-4
3. H.E. Çamurlu, O. Kesmez, E. Burunkaya, E. Arpaç, Sol-gel preparation and characterization of anti-reflective and self-cleaning $\text{SiO}_2\text{-TiO}_2$ double-layer nanometric films, *Solar Energy Mater Solar Cells*(2009), 93:1833–1839
5. VDMA, International Technology Roadmap for Photovoltaic. (2021), URL: [https:// www.vdma.org/international-technology-roadmap-photovoltaic](https://www.vdma.org/international-technology-roadmap-photovoltaic).
6. T. Karin, M. Reed, J. Rand, R. Flottemesch, A. Jain, Photovoltaic module antireflection coating degradation survey using color microscopy and spectral reflectance, *Prog. Photovolt. Res. Appl.* (2022), 30 (11), 1270–1288.
7. A. Law, L. Jones, J. Walls, The performance and durability of Anti-reflection coatings for solar module cover glass – a review, *Solar Energy*, (2023), Volume 261, 85-95
8. N. Nabatov and others. The electricity industry of Bulgaria. Sofia, Tangra TanNakRa, (2011). ISBN 978-954-378-081-5. pp. 458 – 460.

Small-Scale Photovoltaic Power Systems: Identifying Vulnerabilities and Recommending Security Protocols

Michal Mikulášek¹, Pavel Paluřík¹, Lukáš Jablončík¹
Brno University of Technology, Czech Republic¹
xmikul58@vutbr.cz

Abstract: The integration of small-scale photovoltaic power systems into residential and commercial buildings is rapidly gaining popularity. Despite the original design specification for these systems to function solely on local networks or without internet connectivity, they are often linked to installers' or building managers' monitoring systems. The substantial surge in demand and deployment of photovoltaic systems necessitates an examination of their susceptibility to a broad spectrum of threats, encompassing both physical and cyber-attacks. This analysis endeavours to identify the components of these systems and evaluate the associated security vulnerabilities. Additionally, it will include a practical demonstration of selected attacks on a representative system to illustrate potential threats. The analysis will culminate in a summary of recommended security protocols aimed at fortifying the security and robustness of PV systems.

KEYWORDS: CYBERSECURITY, PV PLANT, CYBER-ATTACK, SMART HOME

1. Introduction

The transition towards renewable energy sources has led to a significant increase in the adoption of small-scale photovoltaic (PV) power systems, particularly in residential installations. Homeowners are increasingly leveraging solar power to reduce energy costs, decrease carbon footprints, and gain energy independence. These small PV plants, typically installed on rooftops or in small open spaces, convert sunlight into electricity, providing a sustainable and eco-friendly energy solution.

However, as the number of residential PV installations grows, so does the need to address potential vulnerabilities associated with these systems. The integration of PV systems into the home energy grid introduces several security concerns that were previously not a priority. Given that modern PV systems often include sophisticated monitoring and control systems accessible via the internet, they become potential targets for cyberattacks.

Cybersecurity for small-scale PV systems is an emerging field of importance. These systems are equipped with smart inverters, communication networks, and remote monitoring capabilities, making them susceptible to various cyber threats. Unauthorized access, data breaches, and malware attacks can disrupt the functionality of the PV system, leading to power outages, reduced efficiency, and even physical damage to the equipment.

This article explores the critical vulnerabilities inherent in small-scale photovoltaic power systems and provides recommendations for robust security protocols. By understanding the potential threats and implementing effective security measures, homeowners can protect their PV investments and ensure the continuous, safe operation of their solar energy systems. In the following sections, we will delve into the specific cybersecurity risks associated with residential PV installations and outline best practices for safeguarding these systems against cyber threats.

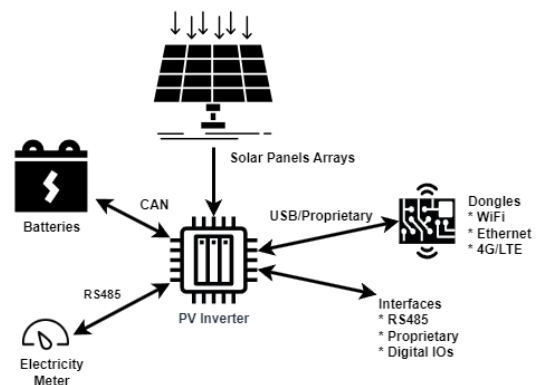
2. Security Concerns

Currently, when installing home solar power plants, remote monitoring devices are installed and connected to smart home systems to enable optimization of energy flow. Unfortunately, this is associated with the risk of attacks from the local site as well as the public one. Depending on the configuration of the private network in these installations, there may be a potential risk of attacks due to a security hole in one of the connected devices in the network, or a poorly set network security policy.

Photovoltaic Systems

Photovoltaic systems themselves consist of solar panels placed on rooftops and other places where they can absorb as much sunlight as possible. Furthermore, power elements, called inverters, which are usually placed in technological rooms. These inverters usually contain a set of interfaces primarily designed to connect the necessary extensions to work properly. These interfaces are the

CAN interface for communication with the batteries, RS485 for communication with the electricity meter for measuring energy flows to the distribution network. In addition to these standard interfaces, extensions in the form of so-called dongles can be connected to these devices, which make them so-called smart and enable supervision, management and possible connection to a smart home systems. The biggest weakness of these extensions is their relatively low complexity considering the price. Furthermore, these devices are typically created as additional accessories from electronics manufacturers that specialize in energy production and processing and instead of information security. Consequently, there is a possibility that circumstances may arise that are unsuitable for deployment in today's internet-connected world.



The most common case of inappropriate use can be directly connecting this dongle to a public network without a firewall to protect the device from attacks.

Another major vulnerability is setting strong enough passwords to access the remote monitoring system and device settings. Fortunately, nowadays device manufacturers often treat this possible security breach by asking users to change default password at the first start/login to the device and devices will not allow further actions without changing to a more robust password.

A big security weakness can be other devices in the local network that can act as a relay for the attackers. These attacks can often be very sophisticated and cause at first glance laggy or otherwise strange behavior. At the same time, thanks to these devices, an attacker can send DoS/DDoS attacks to all devices in the network and thereby destabilize the network or cause some devices to malfunction, e.g.: smart home controllers, control units of photovoltaic power plants, etc [5].

Cyber Attacks

From the point of view of possible attacks on the systems of residential photovoltaic power plants, attacks can be divided into passive and active attacks. Passive attacks are characterized as

attacks that aim to gain information without affecting system resources. The most common form of these attacks is passive monitoring of communication on the local network with the help of a compromised device. Another common attack is port scanning, which makes it easy to see which services are running on attacked devices.

The second category of attacks are so-called active attacks. A characteristic of these attacks is an attempt to change or influence the functionality of system resources. The attacker's goal is to affect authentication, confidentiality and data integrity. Among the typical representatives of these attacks are: Phishing attacks, which most often aim to obtain information about the user (name, password, credit card numbers, PIN, etc.). Its forms can be different from fraudulent SMS or email to fake websites.

Another large category of active attacks are the so-called DoS and DDoS attacks, or Denial of Service. Their main goal is to slow down or completely disable specific services, e.g.: websites, access to the device, etc. Its principle is to send a large number of requests to the attacked device and thereby overload its processing parts. The most typical representatives of this category are ICMP Flood, UDP Flood attack.

Active attacks also include Zero-day attacks, which take advantage of system vulnerabilities that have not yet been patched. Their functionality is highly dependent on device support and the inquisitiveness of the attacker. In these attacks, the result can range from simply collecting data to completely taking control of the device.

The last major category of active attacks are the so-called MITM (Man In the Middle) attacks. The principle of these attacks is to intercept the communication of the participants and then try to modify and read this communication. This attack is popularly used on unsecured, unencrypted communications. A typical example can be connecting to an unsecured public Wi-Fi (Wireless Fidelity) network [1, 2, 3, 4].

3. Measurement Campaign

In light of the potential security threats identified in the previous chapter, a procedure was devised to carry out several basic attacks. These attacks were designed to reveal the security threats and potential risks associated with the system. In the initial stages, a passive attack will be performed to verify active services with the help of a tool for discovering open ports on the communication device. Based on this information, several active attacks will be selected, which will aim to prevent the operation of the device and possibly damage its basic function. These attacks will primarily be carried out from the DoS/DDoS category.

All attacks will be performed on a test workstation equipped with a Windows computer and installed with a Kali Linux virtual system (the attack system), as illustrated in Figure 2. To ensure the accuracy of the results, multiple network cards were employed. One was used for Windows from the user's perspective, one for Kali Linux from the attacker's perspective, and a shared wireless card was used for monitoring and management.

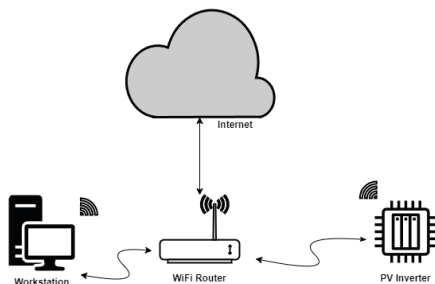


Fig. 2 Testing Setup with Photovoltaic Inverter

To carry out all the attacks, the Linux distribution Kali Linux was chosen, which consists of Debian and many different tools installed in one package. With the help of this system, it is possible to carry out inspections and the attacks themselves according to current needs.

The tested inverter Solax X3-Hybrid G4 photovoltaic system will be gradually connected using Wi-Fi and a LAN dongle to a shared network created using a Mikrotik Wi-Fi router. Due to technological limitations, 4G/LTE dongle testing will be omitted.

During testing, the operation of the inverter will be verified using the system's communication to the Solax cloud servers, by constantly sending ICMP Ping requests, and by sending requests to read registers with the Modbus protocol (if and where possible).

4. Results Discussion

Practical validation tests were done for both categories of attacks: passive and active attacks. Results are discussed in more depth below.

Passive Attacks

Before starting the attacks themselves, so-called lookup attacks were carried out, the aim of which was to find out which ports and services are running on the investigated device in various configurations. The results of these tests were open ports 80 (HTTP) and 502 (Modbus TCP) for the Wi-Fi Dongle, see Figure 3. In the case of the LAN dongle, no ports were open, indicating no potential risks. With this finding in mind, it can be concluded that a LAN dongle is much more secure than a Wi-Fi dongle [4].

```

msf6 > db_nmap -sV 192.168.88.245
[*] Nmap: Starting Nmap 7.94SVN ( https://nmap.org ) at 2024-03-19 10:31 EDT
[*] Nmap: Nmap scan report for 192.168.88.245
[*] Nmap: Host is up (0.028s latency).
[*] Nmap: Not shown: 999 closed tcp ports (conn-refused)
[*] Nmap: PORT      STATE SERVICE VERSION
[*] Nmap: 80/tcp open  nagios-nsc Nagios NSCA
[*] Nmap: Service detection performed. Please report any incorrect results at https://nmap.org/submit/
[*] Nmap: Nmap done: 1 IP address (1 host up) scanned in 14.87 seconds
msf6 > db_nmap -sV -p502 192.168.88.245
[*] Nmap: Starting Nmap 7.94SVN ( https://nmap.org ) at 2024-03-19 10:32 EDT
[*] Nmap: Nmap scan report for 192.168.88.245
[*] Nmap: Host is up (0.042s latency).
[*] Nmap: PORT      STATE SERVICE VERSION
[*] Nmap: 502/tcp open  tcpwrapped
[*] Nmap: Service detection performed. Please report any incorrect results at https://nmap.org/submit/
[*] Nmap: Nmap done: 1 IP address (1 host up) scanned in 0.55 seconds
msf6 >
  
```

Fig. 3 Passive Attack - Nmap for discovering opened ports on device

Furthermore, it can be inferred from the information obtained that the Wi-Fi dongle supports the SCADA (Supervisory Control and Data Acquisition) application protocol Modbus TCP and attacks can therefore be carried out on this protocol, which does not have any form of security for its target deployment.

In the next step, therefore, a scan of the available registers of the Modbus TCP protocol on the Wi-Fi dongle was performed. The scan resulted with 254 available registers being available to read and write, see Figure 4. Thanks to this information, an attacker would be able to attack the target system and potentially disable its operation.

```

msf6 auxiliary(scanner/scada/modbus_findunitid) > use auxiliary/scanner/scada/modbusdetect
msf6 auxiliary(scanner/scada/modbusdetect) > set RHOSTS 192.168.88.245
RHOSTS => 192.168.88.245
msf6 auxiliary(scanner/scada/modbusdetect) > exploit
[*] 192.168.88.245:502 - 192.168.88.245:502 - MODBUS - received correct MODBUS/TCP header (unit-ID: 1)
[*] 192.168.88.245:502 - Scanned 1 of 1 hosts (100% complete)
[*] Auxiliary module execution completed
msf6 auxiliary(scanner/scada/modbusdetect) > use auxiliary/scanner/scada/modbus_findunitid
msf6 auxiliary(scanner/scada/modbus_findunitid) > set RHOSTS 192.168.88.245
RHOSTS => 192.168.88.245
msf6 auxiliary(scanner/scada/modbus_findunitid) > exploit
[*] Running module against 192.168.88.245
[*] 192.168.88.245:502 - Received: correct MODBUS/TCP from stationID 1
[*] 192.168.88.245:502 - Received: correct MODBUS/TCP from stationID 2
[*] 192.168.88.245:502 - Received: correct MODBUS/TCP from stationID 3
[*] 192.168.88.245:502 - Received: correct MODBUS/TCP from stationID 4
[*] 192.168.88.245:502 - Received: correct MODBUS/TCP from stationID 5
[*] 192.168.88.245:502 - Received: correct MODBUS/TCP from stationID 6
[*] 192.168.88.245:502 - Received: correct MODBUS/TCP from stationID 7
[*] 192.168.88.245:502 - Received: correct MODBUS/TCP from stationID 8
[*] 192.168.88.245:502 - Received: correct MODBUS/TCP from stationID 9
[*] 192.168.88.245:502 - Received: correct MODBUS/TCP from stationID 10
[*] 192.168.88.245:502 - Received: correct MODBUS/TCP from stationID 11
[*] 192.168.88.245:502 - Received: correct MODBUS/TCP from stationID 12
[*] 192.168.88.245:502 - Received: correct MODBUS/TCP from stationID 13
[*] 192.168.88.245:502 - Received: correct MODBUS/TCP from stationID 14
  
```

Fig. 4 Passive Attack - Scanning available registers for Modbus TCP

Active Attacks

In the next step, several active attacks were carried out in order to degrade the availability of services on the investigated device. Based on the knowledge of the attacks mentioned in the previous chapter, DoS flood attacks were performed: ICMP Flood, UDP Flood and TCP SYN Flood. Influence of these attacks to whole system was validated by checking the response to ping queries using ICMP messages and the stability of the connection with the Smart home system, in this case the communication stability of the Modbus TCP protocol (in the case of a Wi-Fi dongle), was monitored. Aggregate data from measurements of various combinations are in Table 1.

The results show that the greatest impact of the attacks was on the Wi-Fi dongle during the ICMP Flood attack. In this case, ICMP communication was completely disabled, and in the event of a real attack, the control element would not even respond to basic requests, and the control system would most likely consider device damaged or malfunctioning.

An interesting result is that, in the case of a UDP Flood attack, the Wi-Fi dongle showed very little packet loss compared to the LAN dongle. One difference may be different control electronics, where the Wi-Fi dongle contains software modifications that increase the robustness of the system against these basic attacks.

Table 1 ICMP Ping packet loss for different communication dongles during different DoS attacks

	LAN Dongle	Wi-Fi Dongle	4G/LTE Dongle
ICMP Flood [%]	50	100	-
UDP Flood [%]	88	7	-
TCP SYN Flood [%]	53	81	-

In addition to the mentioned packet losses, some attacks caused a partial and, in some places, even a complete failure of the Modbus TCP service. For ICMP Flood and TCP SYN Flood attacks, there was a complete loss of communication between the user computer and the tested system. Additionally, after these attacks ended, the Wi-Fi dongle was unable to resume normal Modbus TCP operation and required a physical reboot in the form of disconnecting and reconnecting power to the dongle. This discovery is a major security threat, because if an attacker were to aim at disabling the control of the production system, then purely theoretically, it is enough for attacker to attack the Modbus TCP protocol for a limited time and the dongle itself will stop communicating permanently. The worst thing about this discovery is the information that the user may not be able to easily recognize and possibly restore this error without prior knowledge.

5. Conclusion

This article summarizes basic information about photovoltaic systems and their cyber security. As part of the summary of the overview of photovoltaic systems, we have summarized its basic parts with connections to possible points of attack by the attacker. General information about the security for this type of system was summarized in the following text. Subsequently, the basic division of cyber-attacks into passive and active attacks was summarized. A passive attack normally only aims to passively listen and find out information about the internal operation of the system and possibly capture security information. On the contrary, active attacks target already known weaknesses of the system, or discover them with the help of passive attacks.

Practical attacks were carried out at a test site, where a Solax X3-Hybrid G4 photovoltaic system equipped with different configurations of expansion modules, which had different characteristics, was investigated. In the case of LAN extension, it was found that the system had no weaknesses. On the contrary,

several potential weaknesses have been revealed for Wi-Fi extensions.

The first weakness was the available open port 80 (HTTP), which by its properties does not provide any security protection, and therefore attacker could cripple the system thanks to this weakness.

The second weakness was open port 502 (Modbus TCP), which is most often used in industry in SCADA systems. These systems are always isolated from the public network and therefore no security is considered with this protocol. At the same time as there is need to handle many requests in a short period of time, security in a closed industrial system is usually omitted as it only increases the processing delay. Unfortunately, in the case of photovoltaic systems, this protocol is quite often used to connect to Smart home systems, which, on the other hand, are connecting users to the public Internet network. This potentially creates a significant risk in exposing system weaknesses and attacks [6].

6. References

1. KOLOUCH, J. a BAŠTA, P. CyberSecurity. CZ.NIC. 2019. ISBN 978-80-88168-31-7
2. KOLOUCH, J. CyberCrime. CZ.NIC. 2016. ISBN 978-80-88168-15-7
3. BEZPALEC, P. Nové trendy v elektronických komunikacích, Bezpečnost sítí. Online. 2015. [cit. 2024-06-08] Available from: <https://publi.cz/books/223/Cover.html>.
4. BRICKSON, J. Hacking: The Art Of Exploitation. Druhé vydání. San Francisco: No Starch Press. 2008. ISBN 978-1-59327-144-2.
5. Photovoltaic system. Online. In: Energyeducation.ca. [cit. 2024-06-08]. Available from: https://energyeducation.ca/encyclopedia/Photovoltaic_system.
6. ALSABBAGH, W. AMOGBONJAYE, S. URREGO, D. a LANGEN-DOERFER, P. 2022. A Stealthy False Command Injection Attack on Modbus based SCADA Systems. Online. In: Researchgate.net. DOI.10.13140/RG.2.2.28361.83041

Advances in Sensory Systems for Unmanned Underwater Vehicles: A Preliminary Review

Krzysztof Mamcarz
AGH university of Krakow, Poland
mamcarz@student.agh.edu.net.pl

Abstract: Unmanned underwater vehicles (UUVs) are gaining traction in various fields, from oceanographic research to defence applications, due to their versatility and autonomy. This preliminary study explores the integration of Internet of Things (IoT) technology and beamforming techniques to enhance the sensory capabilities of UUVs. The study delves into the evolving landscape of sensory systems, emphasising the incorporation of IoT to enable seamless data exchange, real-time monitoring and adaptive decision making. It also examines the use of beamforming techniques in directional sonar sensing, highlighting their role in improving target detection, localisation and tracking capabilities in underwater environments. By synthesising current research and technological trends, this review provides valuable insights into the synergistic integration of IoT and beamforming techniques to enhance the capabilities of UUV sensor systems for diverse applications.

Keywords: UNMANNED UNDERWATER VEHICLE, BEAMFORMING, INTERNET OF THINGS

1. Introduction

Unmanned underwater vehicles (UUVs) have become essential tools in marine technology, significantly impacting fields like oceanographic research and military operations. The rapid advancements in UUV technology are driven by the need for autonomous systems that can function in tough underwater environments. These vehicles are equipped with advanced sensors that enable them to perform tasks such as mapping the seafloor, monitoring marine ecosystems, and conducting surveillance missions. The success and versatility of UUVs largely depend on the continuous improvements in their sensory systems [1].

A major development in enhancing UUV sensors is the integration of Internet of Things (IoT) technology [2]. IoT creates interconnected networks of sensors and devices, allowing seamless data collection, transmission, and analysis. For UUVs, IoT integration means real-time monitoring, adaptive decision-making, and improved communication with other vehicles and control stations [3]. This connectivity is crucial for efficient and reliable UUV operations, especially in complex underwater environments. IoT technology in UUVs has opened new possibilities for data-driven mission planning and execution, boosting their overall performance and usefulness.

Another critical advancement is the use of beamforming techniques in sonar systems [4]. Beamforming is a signal processing method that directs signals in specific directions, greatly enhancing the detection and localization capabilities of sonar systems. By focusing acoustic energy where it is needed, beamforming reduces noise and interference, leading to more accurate and efficient target detection. This technology is particularly valuable for UUVs, as it allows for precise surface scanning, better obstacle avoidance, and more effective tracking of underwater targets [5]. Innovations in beamforming, such as adaptive beamforming and machine learning integration, have further improved sonar systems, making them indispensable for UUVs.

The combination of IoT and beamforming technologies represents a significant leap in the evolution of UUV sensor systems. IoT provides a strong framework for data exchange and real-time adaptation, while beamforming enhances the precision and effectiveness of sonar-based sensing. Together, these technologies enable UUVs to operate with greater autonomy and accuracy, whether conducting scientific research, performing environmental monitoring, or engaging in defence activities. The integration of these advanced technologies not only improves the operational capabilities of UUVs but also broadens their potential applications, paving the way for more sophisticated and versatile underwater missions.

2. UUVs

Unmanned underwater vehicles (UUVs), also known as autonomous underwater vehicles (AUVs), are robotic devices designed to operate underwater without a human operator. These machines are equipped with various sensors and tools that allow

them to navigate, collect data, and perform specific tasks in underwater environments. UUVs are typically programmed to follow predetermined paths or are remotely controlled from a surface vessel or control station. An example of UUVs in operation can be seen in Figure 1.

UUVs are widely used in oceanographic research to explore and map the ocean floor, monitor marine life, and gather environmental data. They can operate at depths that are difficult or dangerous for human divers, making them invaluable for deep-sea research. These vehicles also play a critical role in monitoring the health of marine ecosystems. They can track changes in water quality, observe coral reefs, and detect pollution, providing essential data for conservation efforts.

In defence and security, UUVs are used by military organisations for surveillance, reconnaissance, and mine detection. Their ability to operate stealthily and access challenging environments makes them ideal for defence applications. The oil and gas industry also benefits from UUV technology, using these vehicles to inspect underwater pipelines, survey potential drilling sites, and assess the seabed for resources. UUVs provide detailed imagery and data that are crucial for safe and efficient resource extraction.

UUVs assist archaeologists in locating and studying submerged historical sites and shipwrecks. Their advanced imaging systems can capture high-resolution photographs and videos, aiding in the preservation and study of underwater cultural heritage. Additionally, these vehicles are used to inspect underwater infrastructure such as bridges, dams, and offshore wind farms. They can identify structural issues and perform maintenance tasks, reducing the need for costly and risky human interventions.



Figure 1. Saab Seaeye Falcon fitted with a special stereoscopic camera array. Nekton Maldives Mission © Nekton 2022.

3. Sensory Systems

UUVs are equipped with a variety of sensory systems that enable them to perform complex tasks in underwater environments. These systems include sonar, cameras, magnetometers, and chemical sensors, each providing critical data for navigation, obstacle avoidance, and environmental monitoring. Sonar systems, including both side-scan and multibeam sonar, are essential for

mapping the seafloor and detecting objects in the water column. Cameras, often equipped with high-definition and low-light capabilities, allow for detailed visual inspections and documentation of underwater features. Magnetometers are used to detect magnetic anomalies, which can indicate the presence of metal objects such as shipwrecks or underwater mines.

The advantages of these sensory systems are numerous. They enable UUVs to operate autonomously in environments that are inhospitable or dangerous for humans, providing valuable data for scientific research, resource exploration, and security operations [6]. For example, UUVs can monitor coral reefs, track marine life, and survey underwater pipelines with a high degree of precision and efficiency.

However, there are limitations to current sensory systems. The underwater environment poses significant challenges such as signal attenuation, noise interference, and limited communication ranges [7]. These factors can affect the accuracy and reliability of sensor data, necessitating advanced signal processing techniques and robust communication protocols. Additionally, integrating and synchronising data from multiple sensors can be complex, requiring sophisticated algorithms and computing power.

Despite these challenges, the applications of UUV sensory systems continue to expand. They are used in scientific research to study ocean currents, temperature, and salinity, providing insights into climate change and marine ecosystems. In the oil and gas industry, sensory systems facilitate the inspection of underwater infrastructure, ensuring the safety and integrity of operations. Military applications include mine detection, surveillance, and reconnaissance, where reliable sensory data is crucial for mission success.

4. Integration with IoT

Integrating Internet of Things (IoT) technology with UUVs marks a significant advancement in underwater operations. IoT enables the creation of interconnected networks of sensors and devices that communicate and share data in real-time [8]. For UUVs, this means enhanced capabilities for data collection, transmission, and analysis, leading to more informed decision-making and efficient mission execution.

One of the primary benefits of IoT integration is real-time monitoring. By connecting UUVs to a network of sensors and control stations, operators can receive immediate feedback on environmental conditions and vehicle status, allowing for adaptive decision-making and timely responses to changing situations. This is particularly valuable in dynamic underwater environments where conditions can shift rapidly.

IoT also facilitates improved communication between UUVs and other vehicles or control stations. This interconnectedness enhances the coordination of multiple UUVs, enabling them to work collaboratively on complex tasks such as large-scale mapping or coordinated surveillance operations. Furthermore, IoT supports the remote control and monitoring of UUVs, reducing the need for human intervention and increasing the safety and efficiency of underwater missions.

Incorporating IoT technology into UUVs has opened new avenues for data-driven mission planning and execution. By leveraging IoT, UUVs can access and analyse vast amounts of data, leading to more accurate predictions and improved performance. This technological synergy is driving innovation in various fields, from environmental monitoring to resource exploration and defence.

5. Beamforming

Beamforming is a signal processing technique that enhances the performance of sonar systems by directing the transmission or reception of signals in specific directions. This technique is

particularly beneficial for UUVs as it improves target detection, localization, and tracking capabilities in underwater environments.

The fundamental principle of beamforming involves combining signals from multiple sensors to create a directional beam of acoustic energy. By focusing this beam in desired directions, beamforming reduces noise and interference, resulting in clearer and more accurate sonar images [9]. This precision is crucial for tasks such as obstacle avoidance, high-resolution mapping, and identifying underwater targets. A diagram showing how beamforming works can be seen in Figure 2.

Recent advancements in beamforming include adaptive beamforming and the integration of machine learning algorithms. Adaptive beamforming dynamically adjusts the beam pattern based on the surrounding environment and target characteristics, enhancing detection and localization accuracy. Machine learning algorithms further refine this process by analysing patterns in the sonar data, enabling more efficient and effective underwater exploration [10].

Combining beamforming with IoT technology represents a significant leap forward in UUV sensory systems. IoT provides a robust framework for data exchange and real-time adaptation, while beamforming enhances the precision of sonar-based sensing. Together, these technologies enable UUVs to perform complex underwater missions with greater autonomy and accuracy.

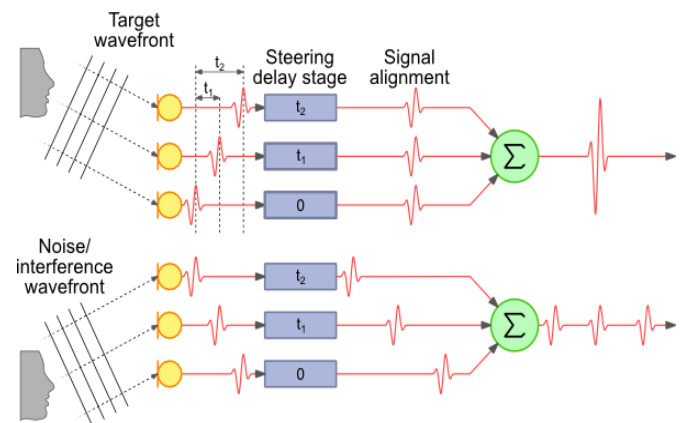


Figure 2. Delay-sum beamforming.

6. Summary

This preliminary highlights the advancements in sensory systems for unmanned underwater vehicles (UUVs), emphasising the integration of Internet of Things (IoT) technology and beamforming techniques. These technologies enhance the sensory capabilities of UUVs, enabling real-time monitoring, adaptive decision-making, and precise target detection. IoT facilitates seamless data exchange and improved communication, while beamforming enhances sonar performance by reducing noise and improving accuracy. The synergistic integration of these technologies is driving innovation in various applications, from oceanographic research to defence. As UUV technology continues to evolve, the combination of IoT and beamforming will play a crucial role in expanding the capabilities and applications of these versatile underwater vehicles.

6. References

- [1] Balestrieri, E., Daponte, P., de Vito, L., & Lamonaca, F. (2021). Sensors and Measurements for Unmanned Systems: An Overview. *Sensors*, 21(4). <https://doi.org/10.3390/s21041518>

- [2] Kabanov, A., & Kramar, V. (2022). Marine Internet of Things Platforms for Interoperability of Marine Robotic Agents: An Overview of Concepts and Architectures. *Journal of Marine Science and Engineering*, 10(9). <https://doi.org/10.3390/jmse10091279>
- [3] Zhang, W., Wei, S., Teng, Y., Zhang, J., Wang, X., & Yan, Z. (2017). Dynamic Obstacle Avoidance for Unmanned Underwater Vehicles Based on an Improved Velocity Obstacle Method. *Sensors*, 17(12). <https://doi.org/10.3390/s17122742>
- [4] Liu, W., & Weiss, S. (2010). *Wideband beamforming: concepts and techniques*. John Wiley & Sons.
- [5] Luo, J., Han, Y., & Fan, L. (2018). Underwater Acoustic Target Tracking: A Review. *Sensors*, 18(1). <https://doi.org/10.3390/s18010112>
- [6] Bovio, E., Cecchi, D., & Baralli, F. (2006). Autonomous underwater vehicles for scientific and naval operations. *Annual Reviews in Control*, 30(2), 117–130.
- [7] Zhufeng, L., Xiaofang, L., Na, W., & Qingyang, Z. (2022). Present status and challenges of underwater acoustic target recognition technology: A review. *Frontiers in Physics*, 10, 1044890.
- [8] Jahanbakht, M., Xiang, W., Hanzo, L., & Azghadi, M. R. (2021). Internet of underwater things and big marine data analytics—a comprehensive survey. *IEEE Communications Surveys & Tutorials*, 23(2), 904–956.
- [9] Gannot, S., & Cohen, I. (2008). Adaptive beamforming and postfiltering. *Springer Handbook of Speech Processing*, 945–978.
- [10] al Kassir, H., Zaharis, Z. D., Lazaridis, P. I., Kantartzis, N. v, Yioultsis, T. v, & Xenos, T. D. (2022). A review of the state of the art and future challenges of deep learning-based beamforming. *IEEE Access*, 10, 80869–80882.

Investigation of composite building elements in a real environment

Krasimira Toncheva

Bulgarian Academy of Sciences, Institute of Metal Science, Equipment and Technologies with Hydro and Aerodynamics Centre “Acad. A. Balevski”, 67 “Shipchenski prohod” Blvd., 1574 Sofia, Bulgaria,
e-mail: krasiton4@abv.bg

Abstract: A study was carried out on the incorporation in a real environment of fabricated innovative building elements. The investigated elements include products such as rectangular plates, “IT” and semi “O” shaped elements, all made of composite material consisting of spherical foam glass granules bonded with a suitable multi-component cement mixture and manufactured by a specially developed technology. The elements have a specific shape, dimensions and particular physical and mechanical characteristics, thermal and acoustic insulation properties and fire-resistant qualities. The aim of the study is to incorporate them into given construction objects, mainly buildings, and to determine the efficiency and efficacy of the process of incorporating the building elements into a real building structure

Keywords: WASTE GLASS, COMPOSITE MATERIALS AND PRODUCTS

1. Introduction

Three groups [1] of innovative best quality building elements made of previously developed and investigated composite materials, consisting of waste glass foam granules and bonding cement mixture with additives were investigated. According to their application, they are divided into the following series:

- A series for insulation panels;
- B series for two types of thermal and acoustic insulation building elements, mainly small and large-sized with and without reinforcement;
- C series for two types (“IT” shaped and semi “O” shaped) thermal insulation building elements.

The main application of the innovative building elements is for protection in a critical situation - to reduce or avoid damage from fire impact. Together with the structure being built, a system of interconnected elements is formed. The object of the research is one such building system, typically including two parts in its structure - non-load-bearing innovative building elements according to their application and the main elements connected with them - load-bearing parts such as beams, columns and other building construction elements.

The definition, calculation and interrelationship between the two parameters of the study - efficacy and efficiency – are discussed in various papers [2 - 9]. The main definitions and specifications are given in the ISO 9000:2015 standard [3]. It defines the baseline data and parameters of the study and determines the feasibility of studying **the given system** as a set of interrelated or interacting elements, showing whether the system so selected is useful and whether it meets the purpose for which it was designed.

The abovementioned innovative building elements in combination with the building structure constitute a system. The study of this system should include the combination of internal and external parameters of the elements of the system which can directly influence the principles of their organization in the development of the system and the achievement of its final objectives.

System components [2].

Each system consists of several key components. The system model and its organization includes the following key elements: input, transformation process, goals, output, feedback and control.

An **organization** is a legal entity engaged in managing one or more systems, which may or may not be linked. The ultimate goal of this process is to obtain a final result.

The **input** of a system can be considered as a specific part of it, in which all its elements which are influenced by the environment are combined to be further transformed in the system through different processes.

The **process** (processing, conversion) is the transformation of everything received at the system input.

The **output** unites all the elements of the system through which it influences the environment. It is another specific part of the system where the results of the transformation process are obtained. It is necessary to control the operation of the system to ensure that its objectives are met.

The **feedback** is typically used to compare the output against the system’s objectives and, depending on the results obtained and if necessary, to exert a controlling influence on the input or on the transformation process.

The **environment** (the boundaries of the system) is a characteristic that is expressed in the definition of its dimensions. Inside the boundaries is the system with its properties and outside the boundaries is the rest of the world. In every organization, it is the management that defines the internal boundaries of each individual activity (i.e., subsystem), and these may be different for each organization.

The **boundaries of the system**, in a broad sense, represent all the elements not included in it, i.e. the world in which the system exists.

On the other hand, ISO 9001:2015 promotes “the adoption of a **process approach** in developing, implementing and improving the **efficacy** of a quality management system in order to increase the satisfaction of the organization’s customer by meeting its requirements.”

At the same time, “understanding and managing interacting processes as a system contributes to the **efficient** and **effective** achievement of the outcomes intended by the organization. This approach allows the organization to manage the interrelationships and interdependencies among the processes of the system so that the performance of the organization as a whole is improved” [2].

According to ISO 9000:2015, a **process** is “a set of interrelated or interacting activities that use **input elements** to deliver an expected result.”

At the same time, an **output element** (result of a process) of the organization that can be produced without any transaction taking place between the organization and the customer is defined as a **product**.



Fig. 1. Diagram of the process definition

Effectiveness is an external measure of performance and indicates how well a process meets the requirements of different stakeholders. In other words, it is “getting the process right” or “doing the right thing.” Process efficacy is responsible for how important the outcome is to the goal that has been set.

The **goal** is the desired state that the organization must achieve within a certain period of time, and its elements include:

- (1) the field it is aimed at;
- (2) a quantitative dimension;
- (3) a scale or measure;
- (4) a time frame.

To define the goal, a particular concept must be used, and it must be specific, measurable, attainable, relevant and time-bound.

Effectiveness is an **internal** performance measure that provides insight into how well a process converts inputs into outputs. The closer the ratio of output to input approaches one (or 100%), the better the efficacy of the process. Defined another way, it is “doing

the things right” or using resources (time, finance, materials, equipment, etc.) correctly.

When analyzing processes in an organization, effectiveness takes precedence over efficiency.

Depending on the degree of development of the two parameters, organizations with

- 1) low efficiency and efficacy **perish quickly**;
- 2) low effectiveness but high efficiency **perish slowly**;
- 3) high efficacy but low efficiency are in the surviving category and
- 4) high efficacy and efficiency **thrive**.

Using the process approach, it is also possible to determine process productivity (Fig. 2).



Fig. 2. Diagram of the “efficiency/efficacy – productivity” relationship

Productivity is defined by comparing the output obtained (efficacy) and the effort put in to achieve the result (efficiency) or in other words, if more can be achieved with less effort, productivity increases. It can be said that productivity is “the extent to which the company generates money” [9].

In conclusion, it can be said that both indicators – “efficiency” and “efficacy” – are important for the organization as they examine the process from different perspectives, respectively its external and internal side.

2. Identification of the elements of the studied system for the incorporation of innovative building elements into the building structure and selection of its parameters to determine efficacy and efficiency criteria

2.1. Object of study

According to the topic of the task formulated as “Study of building elements manufactured and incorporated into a real building,” the investigated elements are products such as rectangular plates, “II” and semi “O” shaped building elements, all made of composite material consisting of spherical foamed glass granules manufactured by a specially developed technology and bonded with a suitable multi-component cement mixture. These elements have a specific shape and size and particular physical and mechanical properties that determine their possible application. In addition to these qualities, they also have thermal and acoustic insulation and fire protection properties. The purpose is to incorporate them into construction objects, mainly buildings, so as to perform certain functions facilitating the implementation and construction of the object itself.

The flat small-sized building elements are placed in the walls around doors and windows, ensuring compliance with Ordinance No I3/1971 of 2009, as amended in 2015, [9], concerning technical building rules and regulations ensuring safety in case of fire - mainly to prevent its spread and limit its destructive consequences. The “II” and semi “O” shaped building elements form channels intended to house the main communication systems of the building such as electrical cables, gas and water pipes, also cables with control functions and other conducting systems. Large-size flat building elements with and without reinforcement serve as panels for the walls, ceilings and floors of new building structures or as panels for thermal and acoustic insulation of existing building structures.

These innovative building elements have primarily a protective function in the event of critical situations **such as fire** so as to reduce or eliminate the risk of destruction of the building structure concerned or of the communication systems located inside the channels until the critical situation is eliminated.

2.2. Components of the system for the incorporation of innovative building elements into the building structure

2.2.1. Input components

The **input** to the system are the building element types together with the building structure as a complete self-contained system that is subject to environmental influences.

The **process** (processing, conversion) represents the transformation of everything obtained from the input of the system, i.e. the construction of the whole building structure with the incorporation of the innovative elements into it.

The **output** brings together all the elements of the system through which it influences the environment, i.e. an undamaged (or slightly damaged) building structure in a critical situation. The transformation process results in increased durability under the same critical impact compared to a conventionally constructed building structure.

It is necessary to **control** the operation of the system in order to ensure the achievement of its objectives by testing the fire resistance of the materials concerned in a specialized laboratory.

The **feedback** is made by comparing the fire damage of the innovative system with that of a conventional one, after carrying out special studies of the fire resistance materials and the manners of laying the communication systems for electricity, water, gas, etc.

The **environment** (outside the boundaries of the system) is a characteristic that is expressed in the definition of its dimensions. It encompasses the entire building structure - i.e. the world in which the system exists.

The **output element** (result of the process) is the newly constructed building structure with its increased durability in case of fire emergency and with the preservation of the qualities of the building structure to a greater extent. This is also the product of the implementation of the building structure.

2.2.2. Real installation scheme.

The impact process is carried if a suitable location is provided for the implementation of the process of incorporation of certain construction elements envisaged for this task. For this purpose, a room in a building site - part of a farmhouse (bedroom) with dimensions 4.2 x 3.5 x 2.5 m; one door 2.0 x 0.9 m; one window 2.6 x 1.3 m, with two external walls 0.25 m thick and two internal walls 0.125 m thick, is selected, as shown in Figure 3.

The total floor area of the sample room includes

- surrounding walls (excluding doors and windows) - 33.5 m²
- thick walls (0.250 m) - 15.85 m² or 4 m³
- thin walls (0.125 m) - 18.65 m² or 2.35 m³
- circumference of doors and windows - 10 m
- ceiling - 14.7 m² or 4.4 m³.

The following composite building elements were used:

- half “O” shaped building elements - 7.7 m or 9 pieces each with a volume of 0.045 m³, total 0.4 m³;
- “II” shaped building elements 5.5 m or 6 pieces each with a volume of 0.0315 m³, total 0.2 m³;
- for doors and windows and covers of building elements - 13.2 m or 14 pieces each with volume 0.015 m³, total 0.2 m³.

Total amount of composite material used: 0.8 m³

When comparing composite products with two types of building material - bricks quads and YTONG blocks, approximately the following results for the cost per cubic meter were obtained:

- average building cost for brick quads 150 BGN/m³
- average building cost for YTONG blocks 370 BGN/m³
- average cost for composite products 500-540 BGN/m³ (estimated).

The materials used include:

Construction of the building using brick quads - approximately 10 m³ or 1,500 BGN;

- Addition of pipes for a 10m-long heating installation and 10m-long wires for electrical installation - 500 BGN including labour;
- Total cost: 2,000 BGN.

For isolation of the communication systems - for 0.8 to 1.0 m³ composite elements – an increase in price by 500 BGN.

Total cost with composites – 2,500 BGN.

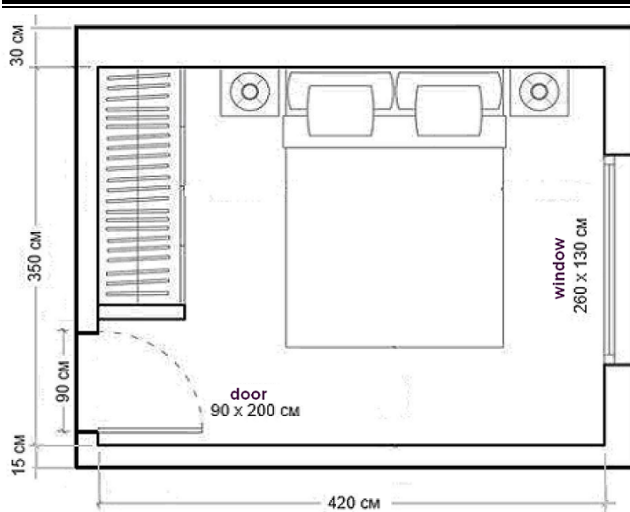


Fig.3. Example diagram showing the incorporation of “II” and semi “O” shaped building elements and small flat ones into a building under construction

Ratio of total cost without and with protection with composite elements (the price increase factor k):

$$k = 2,000/2,500 = 0.8$$

Construction of the site with YTONG blocks - approximately 10 m^3 or 3,700 BGN.

Ratio of total cost without and with protection with composite elements:

$$k = 3,700/4,200 = 0.88$$

The sample calculations are for very small volumes and cannot accurately estimate the price increase factor.

When insulation panels are added to exterior walls and to the ceilings, this factor decreases due to an increase in the denominator. We expertly assume the factor to be in the range from 0.5 to 0.8 and for some more expensive building materials – up to 0.9. For the following calculations, this value of k is assumed, considering only the manufacturer costs of the materials. The selling prices are usually much higher, but the ratio is maintained.

2.2.3. Criteria for efficacy and efficiency of the process of incorporation of new building elements into the studied system - definition and specification of the concepts.

Taking into account the theoretical framework and the management methods, the efficacy and efficiency criteria should be established so as to achieve an accurate balance in the relationship of the two parameters using the limited input resources - the old and new building elements, and the created system should result in an economical engineering solution for the incorporation of these elements into the building structure. This leads to a change in the market demand for precisely such a system that can solve the main problem - creating a long-lasting solution for the fire protection of communication systems in the building structure.

According to ISO 9000:2015:

Efficacy is the degree to which:

- the planned activities have been completed, i.e. the construction elements have been built in and
- the planned results have been achieved, i.e. the durability of the resulting innovative system has been increased.

Efficiency is a correlation between the result achieved, i.e. the increased durability of the resulting innovative system, and the **input resources** used - the embedded innovative building elements.

Following the logic of these definitions, the aim here is to prove that what is defined to be efficient must also be effective, and what is effective must also turn out to be efficient. In other words, the developed system of incorporation of innovative building elements into the building structure appears to be a very efficient solution to improve the durability of the new created system in the event of a critical situation – fire, while also fulfilling the basic function - to

build-in communication systems with additional fire protection capability in the building structure.

Thus, by defining the two parameters, one can also define the concept of **productivity**, which is a relation between efficiency and effectiveness.

At the same time, the above terms also characterize the concept of “**economy**”, which is “the acquisition at lowest cost of the necessary input resources (the new building elements) for the qualitative operation of the system and the realization of an additional quality - fire protection, based on the innovative technical solution.”

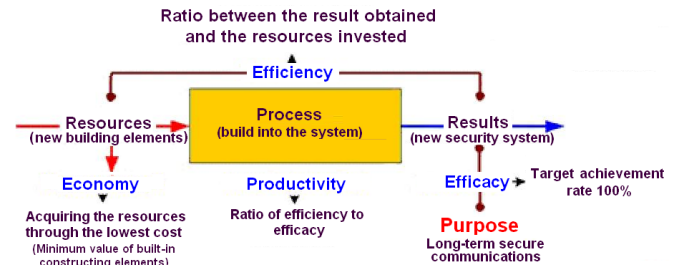


Fig. 4. Diagram of the process in the particular system under consideration and the relationship with efficacy and efficiency

2.3. Measurement of process efficacy and efficiency criteria

2.3.1. Determining the efficacy of the process of incorporating innovative building elements into a building structure

By definition, effectiveness is an external measure of process performance and indicates how well a process meets the requirements of different stakeholders. In other words, it is “getting the process right” or “doing the right things.” The effectiveness of a process is responsible for how important the result is for the goal that has been set. In the case under consideration here, the efficacy of incorporating the innovative building elements into the building structure under construction must be determined, and it must be proved to what extent it is effective in meeting the ultimate stated objective of increasing the durability of the relevant building structure under a critical impact - fire. An expert and logical explanation of this is based on the nature and properties of the foam glass material and the composite material created from foam glass granules and cement-based binders.

The foam glass itself is a highly insulating and non-combustible silicate material with a cellular structure [11] with a very small thermal conductivity factor $\lambda = 0.05 - 0.08 \text{ W/mK}$ and the composite material (also silicate based) used has almost the same properties with $\lambda = 0.08 - 0.095 \text{ W/mK}$ [12, 13]. That is, the high temperature in a critical situation – fire – is not transmitted to the interior of the constructed communication channels within the time required to remedy the critical situation. Thus the interior remains protected from destruction. It can therefore be said that the effectiveness of the system under consideration is 100% because it protects at 100 % the communication systems from destruction and increases their durability.

The defined objective is the desired state that the management of the system must achieve within a certain period of time. Here, the objective is to prevent the spread of fire when using flat building elements or to protect the communication systems when using “II” or semi-“O” shaped building elements.

The elements of the goal are:

- (1) the target area: protection of the building structure in a critical situation - fire;
- (2) quantitative dimension - relative durability - commensurate with the durability of the building structure without the impact of a critical situation - fire;
- (3) scale - 1:1;
- (4) time frame, defined in [6], which must be: measurable, achievable, relevant and defined in time - this is the actual durability, depending on the random event - critical situation (fire).

The effective process thus described - the incorporation of innovative construction elements into a building structure - satisfies the customers, i.e. the users of the structure. The innovative

elements are a quality durable product at the right place and time and with an acceptable price. In this way, a 100% efficacy for the customers is proven. When the management processes of the organization are analyzed, efficacy takes precedence over efficiency.

It can therefore be said that organizations with effective processes are constantly looking for growth opportunities. To some extent, the efficacy of an organization's processes are linked to the ability of its competitors to achieve similar results with the same resources and capabilities, with the seeking organization always pulling others ahead when they try to catch up by creating new effective systems.

2.3.2.Determination of the efficiency of the process of incorporation of innovative building elements into a building structure

Efficiency is an internal performance measure that provides insight into how well a process converts inputs into outputs.

The **inputs** to normal building construction are the ordinary building elements – bricks, concrete or other building blocks (e.g. YTONG blocks) or concrete building panels. For communication systems, there are usually no dedicated grooves, and they are implemented externally on the surface of the building structure. For example, electric cables are run in plastic tubes, while usually for small construction sites such as single-family houses, water, gas, etc pipes are chased into the walls and the grooves are filled with plaster. In large prefab buildings, vertical shafts are constructed, while in the horizontal direction the same methods are used as for single family houses. None of these will be protected in a critical situation such as a fire and will therefore be destroyed. Thus, to eliminate the effects of the critical situation, they must be rebuilt. In Bulgaria, there were no urban gas pipe infrastructure, but recently, over the last 10 - 15 years, gas supply for the homes has been made available in some cities and towns. The output in a critical situation is usually 100% damage to the communication systems.

The proposed innovative building elements made of foamed glass granules, composite material and cement bond are made with the same dimensions as conventional building elements, but they possess protective properties against a critical situation - a fire, where they provide protection for the communication systems of the building structure until the critical situation is eliminated.

Accurate calculations can be made to compare the cost of the two types of building elements.

The cost of conventional and innovative elements is different, and initially, until the innovative elements are established on the market, their cost will be higher. It can be assumed that the cost of the old versus the new elements is in the order of 50-80% - always the new is more expensive until a certain time until it becomes old. But the calculation has to include the additional cost of the communication system, which in a conventional building construction is 80 ÷ 100%, from where it follows that the ratio of output to input is always greater than 1, i.e. the efficiency of the process is better. To put it another way, it is “doing things right” or using resources (time, finance, materials, equipment, etc.) correctly.

So here lies the answer to the question “Can less input resources than planned be used to produce a quality and secure building construction system with innovative building elements?” The answer is Yes, because this solution saves the entire communication system in a critical situation - fire. Therefore, the efficiency is greater than or equal to 1 and involves utilizing every available resource, the idea being to optimize the output-to-input ratio by a single-time utilization of the input resources of the communication system.

The measurement of efficiency is determined by the presence of several interacting environments (physical, economic and managerial), which provides a basis for assessing the following types of efficiency, according to [7]:

- **Physical Efficiency (Ep)** is defined as the ratio between output and input results:

$$Ep = \text{output/input} = \text{product/resources} < 1 \text{ i.e.}$$

$Ep = \text{building elements} + \text{communication system at output divided by building elements} + \text{communication system at input for the new building structure in a critical fire situation relative to the conventional building structure}$

$$Ep = [(0.5 \div 0.8) + (0.8 \div 1.0)] / [1.0 + 1.0] = 0.65 \div 0.9 < 1$$

The goal is to bring Ep close to 1, which means that resources are more fully utilized.

- **Economic efficiency (Ee)** is the ratio between the numerical expression of the achieved utility (preservation of the two parts of the building structure) and the cost of the used production resources, i.e.

$$Ee = \text{utility/cost} > 1 \text{ i.e.}$$

Ee = the utility of the innovative building structure consisting of the retention of its two component parts in relative units - building elements and communication system divided by the total value of the two parts of the building structure relative to the conventional building structure.

$$Ee = [1.0 + 1.0] / [(0.5 + 0.8)/2 * \div (0.8 + 1.0)/2 *] = 3.08 \div 2.22 > 1$$

* It is divided by 2 to get the total relative weight of the system with its 2 parts - building elements and communication system.

- **Total Efficiency (Et)** is the product of Ep and Ee :

$$Et = Ep \cdot Ee > 1$$

$$Et = (0.65 \div 0.9) \cdot (3.08 \div 2.22) = 2.02 \div 1.998$$

$$Et \approx 2 > 1$$

2.4.Efficiency-efficacy matrix of the incorporation process of innovative building elements

It can be concluded for the new building structure that with positive values of both the efficacy and efficiency factors, the considered system with innovative building elements embedded in a building structure falls in the upper right corner of the matrix - Fig. 5. This indicates that the organization and management of the system with the processes that are implemented to protect the same system characterize the enterprise as thriving. Any company that manufactures this product and implements it in real building structures would develop, would make a profit in the timeframe dependent on the development horizon - 10 years or more on average for construction companies.

Using the process approach, the process productivity can also be determined (Fig. 6).

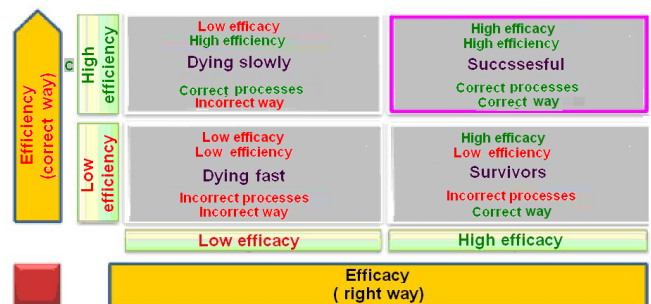


Fig. 5. Diagram of the “efficacy- efficiency” matrix of the process of incorporation of innovative building elements into a building structure in a critical situations - fire

Productivity can be determined by presenting the resulting output (efficacy) with an efficacy factor

$$K_{\text{efficacy}} = 1$$

relative to the efforts made to achieve the result (efficiency)

$$K_{\text{total efficiency}} = 2.02 \div 1.998 \approx 2 > 1$$



Fig. 6. Diagram of the “efficacy/ efficiency - productivity” relationship in the process of incorporation of innovative building elements into a building structure in critical situations - fire

or in other words, if more can be achieved with less effort, productivity increases i.e. productivity is: “the extent to which the company generates money.”

In conclusion, it can be said that both the “efficacy” and “efficiency” indicators are important for the organization as they look at the process from different perspectives, respectively from its external and internal side.

3.CONCLUSIONS

3.1. A system and process approach in the organization is shown, emphasizing the relationship between them through the sequence of the elements of the process "planning-organization-management-control" and the elements of the system "planning-implementation-verification-action" in a system for embedding innovative construction elements in a building structure for protection in a critical situation - fire.

3.2. The characteristics of efficacy and efficiency are derived as well as the approaches to their measurement. The input parameters are defined using their relative values applied to a conventional building construction (without innovative building elements). It is shown that effectiveness is an external parameter of the system and efficiency as an internal measure for the process of incorporation of innovative building elements.

3.3. The logic of the relationship between process efficacy and efficiency is clarified through a matrix that shows the place of an overperforming organization (with a good development horizon - over 10 years lead due to the innovation), while at the same time the productivity problem is also addressed.

The main conclusion of the study is that there is no clear answer to the question which indicator is more important, as different aspects of the processes in the organization are measured.

The organization itself has to determine in the process of its development in which part of the matrix it should evolve. This will depend on how intelligent its management is and what short-term and long-term goals (effectiveness) are set in seeking and solving the ongoing tasks (development of a new process, system of processes, emphasizing efficiency). Ultimately, this will define the organization, in this case one in the construction industry, as surviving or as constantly thriving.

4. References

[1] Project: “Creation and research of a new experimental technological line and technology for the production of fire-resistant, energy-saving, noise and heat-insulating composite building elements based on glass waste” with Contract No. 13 IF 02-21/12.12.2022

[2] Nedyalkov Anton, Process efficacy or process efficiency: Which indicator is more important for organizations?, PROCEEDINGS OF UNIVERSITY OF RUSE “Angel Kanchev” - 2020, volume 59.

[3] ISO 9000:2015, Quality management systems – Fundamentals vocabulary and Quality management systems -- Requirements.

[4] Nedyalkov, A. (2012). Problemi na kachestvoto v operatsionniya menidzhmant na uslugite. Ruse: Primaks.

[5] Vitliemov, P. (2014). Savremenni proizvodstveni sistemi. Ruse: Primaks.

[6] Doran, G. T. (1981). There's a S.M.A.R.T. Way to Write Management's Goals and Objectives. Management Review, 70(11), 35-36.

[7] Mirchev, A. (1996). Proizvodstven menidzhmant. Varna: Princeps.

[8] IATF 16949:2016. Particular requirements for the application of ISO 9001:2015 for automotive production and relevant service part organizations (Section 5.1.1.2: Process efficacy and efficiency).

[9] Goldratt, E., Cox, J. (2016). The Goal: A Process of Ongoing Improvement. Third edition.

[10] Наредба № I3/1971 г. от 2009 г., изменена 2015 г. към ЗУТ.

[11] Toncheva Kr., L. Lakov, TECHNOLOGIES AND STRUCTURES FOR PRODUCTION OF FOAM GLASS - OVERVIEW, SCIENTIFIC CONFERENCE RU & SU'10, October 29-30, 2010, Ruse, Series 1.2, pp. 25-33.

[12] L. Lakov, B. Jivov, M. Aleksandrova, Y. Ivanova, K. Toncheva, “An innovative composite material based on sintered glass foam granules”, Journal of Chemical Technology and Metallurgy, 53, 6, 2018, pp. 1081-1086.

[13] L. Lakov, B. Jivov, Y. Ivanova, S. Yordanov, M. Marinov, S. Rafailov, “Composite Materials Obtained from Foamed Silicate Products”, International Journal “NDT Days”, Volume II, Issue 2, Year 2019, pp. 188-194.

[14] L. Lakov, B. Jivov, Y. Ivanova, S. Yordanov, K. Toncheva, “Alternative possibilities for application of foamed silicate materials”, International Scientific Journal “Machines. Technologies. Materials” Issue 1/2021, pp. 25 - 27.

[15] K. Toncheva, B. Jivov, L. Lakov, A. Staneva, “Technological approach for production of silicate insulating materials and panels for the decorative cladding of buildings”, Journal of Chemical Technology and Metallurgy, 58, 4, 2023, pp. 716-724.

[16] L. Lakov, G. Mutafchieva, G. Peev, M. Aleksandrova, B. Jivov, “Physicochemical and technological researches of marls from the area of the village Lovetz related to the production of the new “Yellow paving stones””, International scientific Journal “Material science, “Non-Equilibrium Phase Transformations”, Year VII, ISSUE 3/2021, pp. 107-109.

[17] Y. Ivanova, T. Partalin, L. Lakov, B. Jivov, “Airborne sound insulation of new composite wall structures”, MATEC Web of Conferences, 145, 05013, 2018, eISSN: 2261-236X, DOI:https://doi.org/10.1051/mateconf/201814505013

[18] L. Lakov, Y. Ivanova, B. Jivov, M. Aleksandrova, N. Barovsky, „Investigation of mechanical properties and coefficients of sound insulation of innovative silicone composite material”, Prace Instytutu Odlewnictwa, Transactions of Foundry Research Institute (Prace IOd, Transactions of FRI), Volume LVII, Number 4, Year 2017, pp. 245-251.

[19] G. Mutafchieva, L. Lakov, M. Aleksandrova, G. Peev, D. Mihailov, B. Jivov, “Technological Processes and Characteristics of Lighting Products Made of Ceramic Materials”, Bulgarian Society for NDT International Journal “NDT Days”, Volume V, Issue 1, Year 2022, pp. 23-28.

[20] M. Aleksandova, L. Lakov, B. Jivov, V. Blaskov, “Synthesis of La doped barium titanate ceramics with application in electronics”, International Journal “NDT Days”, Vol. IV, Issue 5, Year 2021, pp. 311-316.

Utilization of waste chromium electrolyte by addition of Al_2O_3 nanoparticles to obtain thin chromium coatings

Vladimir Petkov, Mihaela Aleksandrova, Stancho Yordanov, Shaban Uzun

Bulgarian Academy of Sciences, Institute of Metal Science, Equipment and Technologies with Hydro- and Aerodynamics Centre "Acad. A. Balevski", Shipchenski Prohod Blvd. 67, 1574 Sofia, Bulgaria, e-mail: vladimir2pe@yahoo.com

Abstract: Chromium coatings were obtained from spent (waste) chromium electrolyte with the addition of nanosized aluminum oxide. The recovery of the waste chromium electrolyte is carried out by the activation of the chromium ions using the addition of Al_2O_3 nanoparticles. Electrochemically thin chromium films were deposited directly on S235 steel. The recovered chromium electrolyte has a concentration of nanoparticles of Al_2O_3 up to 10 g/l, which particles act as intensifiers of the electrochemical process. Analysis was performed with a "Polyvar Met" metallographic microscope and a Bruker D8 Advance powder X-ray diffractometer (XRD). The microstructure of the layer and the matrix were examined and the main phases were determined. SEM-EDS analysis of the chromium layer and the steel matrix was performed. The thickness of the obtained chrome coating with aluminum oxide nanoparticles was determined, which varies between 5 - 10 μm .

Keywords: CHROMIUM COATINGS, NANOPARTICLES, Al_2O_3 , MICROSTRUCTURE, MORPHOLOGY

1. Introduction

Chromium plating is mainly used to improve the properties of metal surfaces such as hardness, corrosion resistance and wear resistance in industry [1, 2, 3]. In our time, galvanic chrome coatings, in addition to increasing hardness, are often used for decorative purposes [4], and the most important conditions for their application are good adhesion to the base, colour range and high resistance to corrosion and wear [5].

Chromium electrolyte recycling is an important process to recover and reuse valuable resources used in anodizing metals such as titanium. Chromium is a key component in the electrolytes used in anodizing titanium, and its efficient recovery can contribute significantly to reducing the waste and costs associated with anodizing. In the literature, the possible ways of beneficiation of chromium from waste electrolytes such as electrochemical reduction and subsequent precipitation are reported [6]. The recovery of chromium electrolytes from anode sludge by a combination of physical and chemical processes, electrochemical deposition or ion exchange, and comparing the performance of different ion exchange resins [7, 8]. These studies contain valuable information on the different methods of chromium electrolyte disposal and can help anode manufacturers to optimize their processes and reduce their environmental impact [9, 10, 11].

The recovery of waste chromium electrolyte is of great importance for the protection of the environment and the recycling of useful resources. In the present study, the possibility of adding Al_2O_3 nanoparticles to the waste chromium electrolyte in order to recover the chromium in the form of a coating is considered. This method will contribute not only to improving the quality of coatings, but also to reducing the harmful effects of chromium in the environment.

2. Experimental part

The aim of the development is to restore depleted chromium electrolyte by adding aluminum oxide nanoparticles. Activation of residual chromium ions in order to deposit them on steel in the form of a thin coating of metallic chromium. The phase composition, morphology, microstructure and elemental composition in the chromium layer and the matrix were investigated. The original test samples of steel are rectangular plates with dimensions - 10 x 20 mm and height h - 5.0 mm (Fig. 1). Figure 1 shows part of the chromium-coated test specimens photographed with an optical microscope. They show the structure of the surface of a chrome coating with a magnification marked as Ax20 and Bx80 with the same degree of illumination. As a rule, the chromium layer follows the surface of the steel, aluminum or other composite test specimen, which is slightly rough and uneven. In principle, chrome coatings have a granular structure, slightly wavy with fine microcracks (Fig. 1) [1, 2, 3].

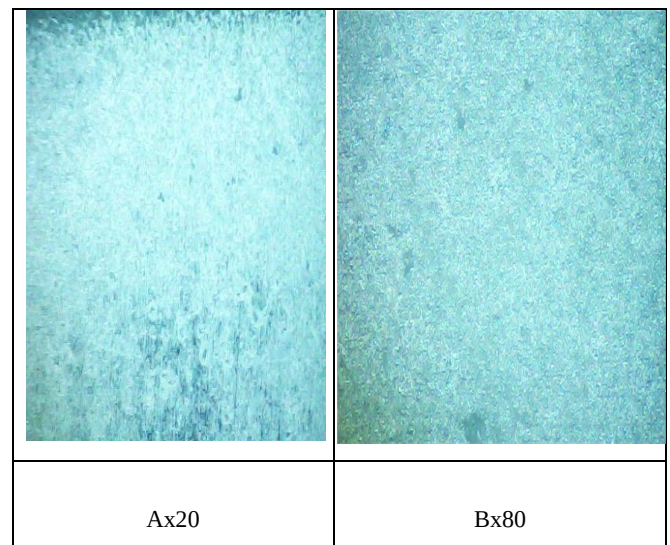


Fig. 1. Chromium coating from reduced chromium electrolyte with Al_2O_3 nanoparticles deposited on steel at magnifications A x20 and B x80

Description of the technological path for obtaining chromium coatings from a reduced chromium electrolyte on a steel matrix: The samples are first ground, then degreased in acetone or another organic solvent, then etched in 20-30% hydrochloric or sulfuric acids [9], before hanging on the cathode and dipping into the chromium plating electrolyte. Precise chemical pre-treatment is of utmost importance to obtain continuous, uniform and dense chrome coatings. In the present experiment, a depleted chromium electrolyte with an initial chemical composition of chromic anhydride sulfuric acid ratio of 100:1 [1, 2] was used. Chemical composition of the electrolyte is: CrO_3 - 220 g/l; H_2SO_4 - 2.2 g/l.

Electrochemical parameters of the process are:

Current density: 50 A/dm^2

Duration: about 50 min

Electrolyte temperature: about 55°C

An electrolyte of aluminum oxide nanoparticles is added to the depleted chromium electrolyte described above. It is in the form of a water-alcohol suspension with a concentration of nanosized Al_2O_3 up to 10%. This is followed by activation of the suspension with Al_2O_3 in an ultrasonic bath and addition of the obtained electrolyte as mentioned to the spent electrolyte for chromium plating. The pretreated steel samples suspended on the cathode are immersed in the new electrolyte thus obtained, where thin layers of chromium are electrodeposited. The end result of the recovered chromium

electrolyte is the direct deposition of thin metallic chromium on steel shown in (Fig. 1).

3. Results and discussion

3.1. X-ray structural analysis

Investigations were done with a Bruker D8 Advance powder X-ray diffractometer. The Bruker DIFFRAC.EVA program was used to determine the qualitative and quantitative phase analysis

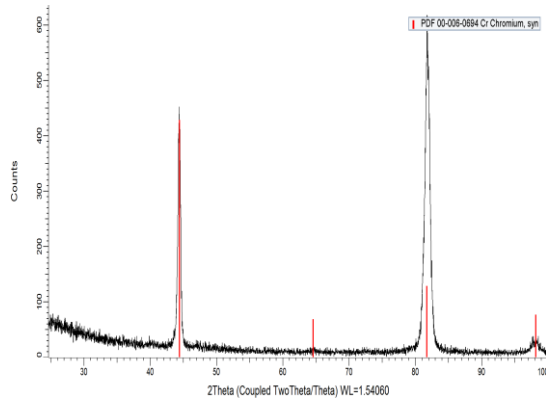


Fig. 2. Diffractogram of chromium coating with Al_2O_3 nanoparticles deposited directly on steel

In the diffractogram shown in Figure 2, a single phase of metallic chromium with a cubic structure is observed, which is evidence of the good density of the chromium coating.

3.2. Microstructure

With a "Polyvar Met" metallographic microscope, the microstructure of the chromium coating with a concentration of 10 g/l of aluminum oxide nanoparticles in the electrolyte deposited on steel was imaged (Fig. 3).

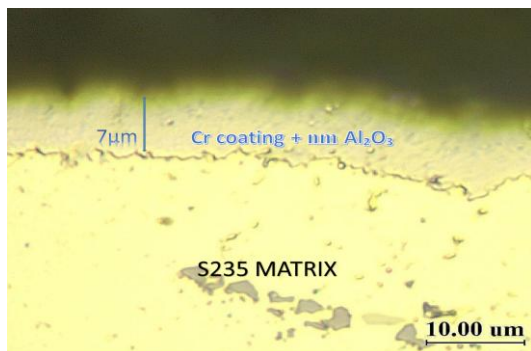


Fig. 3. Microstructure and thickness of chromium coating obtained from reduced chromium electrolyte with addition of aluminum oxide nanoparticles

On the microstructure shown in (Fig. 3) with a magnification of $\times 1000$, a thin chromium layer with a thickness of about $7\text{ }\mu\text{m}$ is observed. It is uniform and continuous following the surface of the steel matrix.

3.3. Surface morphology and elemental composition determination

The investigations were carried out using a scanning electron microscope (SEM) "HIROX SH-5500P" with an energy dispersive spectroscopy (EDS) system "QUANTAX 100 Advanced" with an accelerating voltage: from 1 to 30 kV. The surface of the chrome layer was photographed at a magnification of $\times 10000$. It is slightly wavy, continuous and dense, which morphology is characteristic of chrome coatings. Figure 4 shows the thickness of the chromium coating, which correlates well with the metallographic microscope thickness shown in (Figure 3). It is about 7 microns with a deviation of $\pm 1.5\text{ }\mu\text{m}$, which tells us about a very good uniformity of the chrome layer.

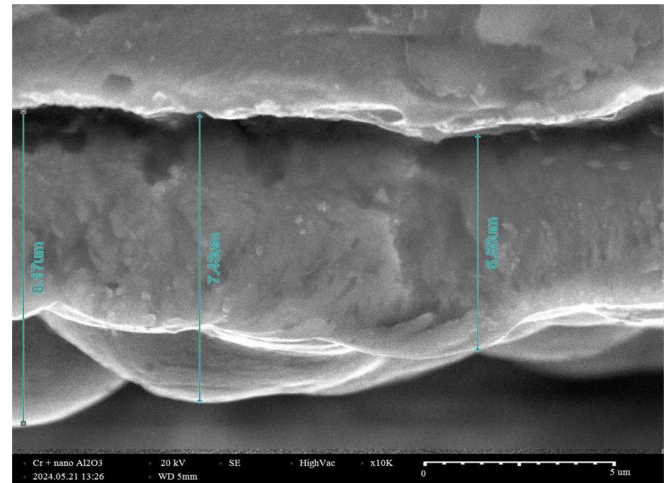


Fig. 4. SEM image of chromium coating with aluminum oxide nanoparticles deposited on a steel matrix

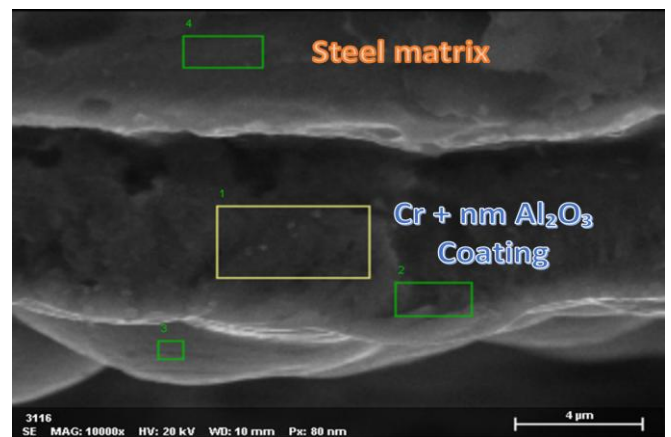


Fig. 5. SEM image of a zonal investigation of a chromium coating with aluminum oxide nanoparticles deposited on a steel matrix

Figure 5 shows a zone-examined and characterized by SEM-EDS analysis of the Al_2O_3 nanoparticle chromium coating and the steel matrix. Using energy dispersive spectroscopy, the elemental composition was determined in three zones: 1, 2, 3 for the chromium layer and zone 4 for the steel matrix.

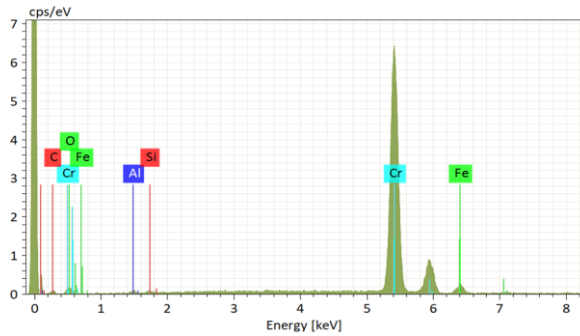


Fig. 6. Spectral analysis of the chemical elements in the chromium layer in zone 1

Element	At. No.	Netto	Mass [%]	Mass Norm. [%]	Atom [%]	abs. error [%] (1 sigma)	rel. error [%] (1 sigma)
Cr	24	28180	79.75	90.41	81.65	2.23	2.80
Fe	26	825	3.77	4.28	3.59	0.20	5.32
C	6	137	1.93	2.18	8.54	0.88	45.51
Al	13	145	1.43	1.62	2.82	0.19	13.05
Si	14	128	0.73	0.83	1.39	0.11	14.73
O	8	142	0.60	0.68	2.01	0.29	47.70
Sum		88.21	100.00	100.00			

Figure 6 shows a spectral analysis of the elemental composition of the chromium layer in zone 1, where it was found that the main element is chromium constituting 90.41 wt%, in very small amounts aluminum, carbon and iron and traces of silicon and oxygen.

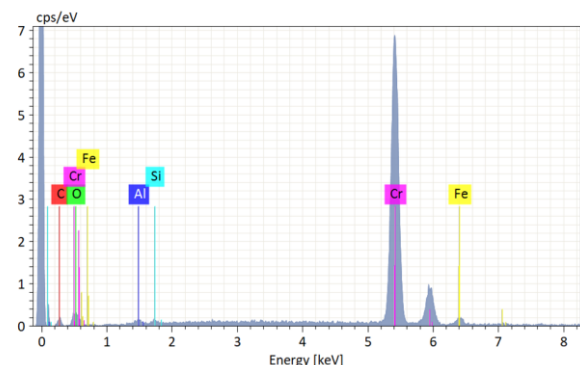


Fig. 7. Spectral analysis of the chemical elements in the chromium layer in zone 2.

Element	At. No.	Netto	Mass [%]	Mass Norm. [%]	Atom [%]	abs. error [%] (1 sigma)	rel. error [%] (1 sigma)
Cr	24	25150	75.35	89.75	76.50	2.12	2.81
C	6	232	3.38	4.02	14.84	1.25	37.01
Fe	26	569	2.76	3.29	2.61	0.17	6.19
Al	13	194	1.11	1.32	2.17	0.14	12.48
O	8	201	0.94	1.11	3.09	0.38	40.87
Si	14	122	0.42	0.50	0.80	0.07	17.43
Sum		83.95	100.00	100.00			

In the area marked with 2, data from researches in the chromium layer are shown in tabular form, from which it is clearly seen that the main element is Cr - 89.75 % and minimal amounts of C, Al, and Fe.

In zone 1 and 2, marked with a yellow and filled rectangle, the data from the studies in the chromium layer are shown, from where it is clearly seen that the main elements are approximately the same in

both zones, with the element chromium having the largest mass percentage. In zone 4 shown in (fig. 5) a SEM-EDS study was made in the steel matrix, which is not relevant for our experiment and we have not analyzed it.

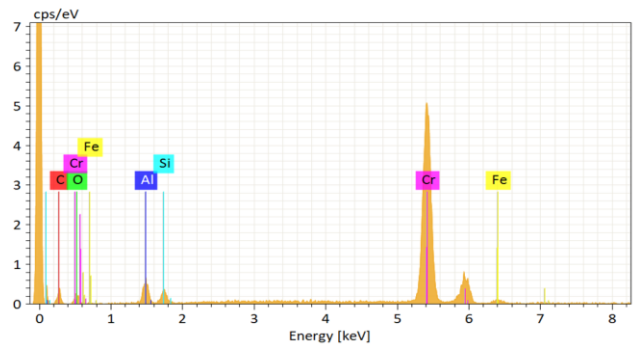


Fig. 8. Spectral analysis of the chemical elements in the chromium layer in zone 3.

Element	At. No.	Netto	Mass [%]	Mass Norm. [%]	Atom [%]	abs. error [%] (1 sigma)	rel. error [%] (1 sigma)
Cr	24	13659	67.56	69.43	44.70	1.95	2.89
Al	13	863	11.09	11.39	14.13	0.77	6.94
C	6	367	10.39	10.67	29.74	3.23	31.13
Si	14	507	4.18	4.29	5.12	0.32	7.72
O	8	237	2.46	2.53	5.30	0.91	36.99
Fe	26	223	1.64	1.69	1.01	0.15	9.16
Sum		97.31	100.00	100.00			

SEM-EDS studies of the chromium layer made in zone 3, appearing closest to the surface of the coating, are of particular interest regarding the obtained spectra of the elements carbon and aluminum. The noted carbon is 10.67 wt%, although with a large margin of error, it is very likely that it is from the used nanodiamonds in the waste chromium electrolyte during chroming and/or from the carbon content of the steel matrix. The other element that increases in this zone is Al with a content of 11.39 wt% probably from the used alumina nanoparticles and impurities from the waste chromium electrolyte. A kind of two-layer coating is obtained on the surface of the investigated layer of chromium and aluminum, for which further research is needed.

4. Conclusion

- Waste chromium electrolytes are utilized by adding nanosized aluminum oxide up to 10% g/l.
- It was found that Al_2O_3 nanoparticles activate the chromium ions and intensify the whole electrochemical process.
- Chromium coatings were obtained from regenerated electrolyte with a thickness of about 7 μm deposited directly on the stoma - S235.
- The presence of chromium with a cubic structure in the chromium coating was established through X-ray structural analysis.
- SEM-EDS analysis revealed an increased carbon content in the layer, most likely from the steel matrix and traces of carbon in the chromium coating, which are most likely from diamond nanoparticles involved in the waste chromium electrolyte.
- In zone 3 of the SEM-EDS study, an increased Al content was observed, most likely from the alumina nanoparticles used

- In our opinion, something like a second layer or composite of chromaluminum oxide is formed on the surface of the chrome plating

Acknowledgements

This work was supported by the European Regional Development Fund within the OP “Science and Education for Smart Growth 2014 - 2020”, Project CoE “National center of mechatronics and clean technologies“, № BG05M2OP001-1.001-0008-C08.

References:

- [1]. Gidikova N., M. Sulowski, V. Petkov, R. Valov, G. Cempura, "Composite coatings of chromium and nanodiamond particles on steel", Arch. Metal. Mater, vol. 62, (2017), 4, 2421-2424
- [2]. Petkov V., R. Valov, S. Simeonova, M. Kandeve, "Characteristics and properties of chromium coatings with diamond nanoparticles deposited directly on aluminum alloys", Archives of Foundry Engineering, (2020), vol. 20, 4, 115-120.
- [3]. Petkov V., M. Aleksandrova, R. Valov, V. Korzhov, V. Kiiko, T. Stroganova, "Studying the surface and microstructure of a chromium coating with nanodiamond particles deposited on monolithic composite with layered structure", International Scientific Journals Machines . Technologies. Materials. vol. 17, 2, (2023), p. 67-69.
- [4]. Metalloberflaeche 62, (2008), 1-2, p. 12-14.
- [5]. Dahlhouse, M. Galvanotechnik 98, (2007), 9, p. 2113-2116.
- [6]. C. Chang, TH Lin and CW Hwang (Journal of Applied Electrochemistry, vol. 29, 3, (1999), p. 271-278.
- [7]. M. A. Al-Obaidi, A. M. Al-Khazraji, S. A. Al-Ani, Journal of Environmental Science and Health Part A, vol. 38, 1, (2003), p. 1-10.
- [8] L. Lakov, B. Jivov, K. Toncheva, „Development of Lead-Free Ceramic Materials for the Electronics. A Review”, Journal of Materials Science and Technology, Vol. 23, No 4, 2015, pp 345-365, ISSN: 0861-9786.
- [9] K. Toncheva, L. Lakov, B. Jivov, “Innovative mill complex for processing glass waste, applicable to the production of foam glass and composite materials”. International Scientific Journal “Machines. Technologies. Materials.”, Year XVII, Issue 8, 2023, p. 298-301, ISSN: Print ISSN 1313-0226, Online ISSN 1314-507X.
- [10]. Y. Zhang, X. Wang and L. Chen (Journal of Hazardous Materials, Vol. 157, (2008), p. 128-135.
- [11]. R. S Kunzmann, D. J. Gauger and J. D. Stankiewicz (Water Research, vol. 35, 12, (2001), p. 3093-3100.

Bio-synthesis of ZnO nanoparticles as potential corrosion inhibitors

Daniela Stoyanova¹, Irina Stambolova¹, Maria Shipochka¹, Ognian Dimitrov², Pavel Markov¹, Kamelia Kamburova³, Tsetska Radeva³

Institute of General and Inorganic Chemistry, Bulgarian Academy of Sciences, Sofia, Bulgaria¹

Institute of Electrochemistry and Energy Systems "Acad. Evgeni Budevsky", Bulgarian Academy of Sciences, Sofia, Bulgaria²

Institute of Physical Chemistry "Acad. R. Kaishev", Bulgarian Academy of Sciences, Sofia, Bulgaria³

Abstract: The interesting physicochemical properties of ZnO determine its wide range of medical, biomedical, industrial and electronics applications. The plants in green synthesis play role as both reducing and capping agents, due to the presence of polyphenols, terpenoids, catechins, etc. The combination of the beneficial properties of ZnO nanoparticles and plant extracts makes them perfect non-toxic materials in the medical aspect, as well as an environmentally friendly candidate for application as corrosion inhibitors in industry. The effectiveness of traditionally obtained ZnO particles as anticorrosive agents for metals and fillers for fabrication of composite anticorrosive coatings has already been reported.

In the present study, ZnO nanoparticles were prepared using a combination of green synthesis, followed by hydrothermal (at 180 °C for 12 h) or microwave treatment. The resulting powder samples were characterized by means of Infrared Spectroscopy (IR), Scanning Electron Microscopy (SEM), Transmission Electron microscopy (TEM), X-ray photoelectron spectrum (XPS). The electrokinetic zeta potentials (ζ) of the particles were determined by electrophoresis mobility measurements using Dynamic light scattering (DLS) technique.

The SEM photographs of the hydrothermally treated sample reveal hexagonal prism-like aggregates. The microwave treated ZnO powders possess irregular shaped plate-like structure.

The XPS analyses proved the presence of nitrogen only in the Z-MT sample. The high resolution spectrum of N1s shows two peaks with binding energies at 398.5 eV and 399.6 eV, associated with imine (-N=) and amine (-NH-) nitrogen atoms, respectively. The VB edge is shifted to lower binding energies in the Z-MT, which indicates a decrease in the band gap of ZnO.

The dispersion of the particles in water at pH 8 leads to positive zeta potential values for the samples prepared by hydrothermal and microwave assisted methods: +7.8 mV and +24.3 mV, respectively.

On the base of the physicochemical measurements it could be supposed that the obtained ZnO nanoparticles have potential as green eco-friendly corrosion inhibitor. Deposition of the particles on the metal surface block the penetration of corrosive medium, thus slowing the rate of the corrosion processes.

1. Introduction

The interesting physicochemical properties of ZnO determine its wide range of medical, biomedical, industrial and electronics applications. The plants in green synthesis play role as both reducing and capping agents, due to the presence of polyphenols, terpenoids, catechins, etc. The combination of the beneficial properties of ZnO nanoparticles and plant extracts makes them perfect non-toxic materials in the medical aspect, as well as an environmentally friendly candidate for application as corrosion inhibitors in industry. The effectiveness of traditionally obtained ZnO particles as anticorrosive agents for metals and fillers for fabrication of composite anticorrosive coatings has been proved by other researches.

2. Problem discussion

The interesting chemical, optical, electrical, electrochemical, gas sensitive etc. properties of ZnO determine its wide range of medical, biomedical, industrial and electronics applications. The effectiveness of ZnO particles as anticorrosive agents for metals and fillers for fabrication of composite anticorrosive coatings is widely reported in the literature [1]. In the last decades the bio synthesis has been proposed as novel ecological alternatives to the chemical methods. The biosynthesis guarantees the production of metal and metal oxide nanoparticles with improved properties compared to traditional methods. In the bio synthesis as biological species can be used plant extracts, which are economically beneficial, playing role as a capping and stabilizing agents. Microwave synthesis offers several advantages in comparison to the conventional thermal treatment: rapid and uniform energy transfer, higher reaction rate (without influencing the activation energy of reactants), energy saving and high product yield and purity. In the scientific literature there are only a few articles, devoted to the preparation of ZnO particles by means of combination of green synthesis and microwave treatment. The application of both *Mentha* leaves as biological template and microwave irradiation could be useful tool to control the final structure, chemical composition and surface properties of ZnO.

[1] T.W. Quadri, L.O. Olasunkanmi, O. E. Fayemi, E. E. Ebenso, Utilization of ZnO-based materials as anticorrosive agents: a review, *Inorganic Anticorrosive Materials*, Elsevier Inc., 2022, 161-182.

3. Objective and research methodologies

3.1 Objective This study aims to investigate the effect of the both the *Mentha* leaves extract as biological template in Zn solution the and treatment with microwave irradiation on the final structure, chemical composition and surface properties of ZnO particles

3.2. Preparation of zinc oxide nanoparticles

- The *Mentha Arvensis* (MA) leaves (from region, located in the north of Bulgaria) were washed and left to dry up. After that 12 g of MA leaves were ground into powder using a mortar and the powder was mixed with 230 ml double-distilled water. The mixture was heated at 60 °C for 10 minutes under continuous magnetic stirring. In order to obtain the final extract, the solution was being centrifuged at 4000 rpm for 20 minutes, after that the material was filtered by KA2 filter paper (100 pcs, PAPIRNA Perstejn s.r.o., CZECH REPUBLIC).

- The first synthesis step is dissolving of 2.2 g zinc acetate in 100 ml of deionized water. Urea aqueous solution (0.6 M) was added dropwise to 0.2 M zinc acetate solution. The pH of solution was adjusted to 7 using ammonium hydroxide. The resulting precursor solution was poured into beaker and then the evaluated quantity of the as-prepared extract MA (10%) was added and stirred at a rate of 200 rpm. The resulting mixture was divided into two parts. One part was transferred into an autoclave and it was maintained at 180 °C. The so obtained samples were denoted as Z-CT. The second part of the solution was subjected on microwave irradiation in domestic MW oven at 235W for 45 min. The resulting sample was denoted as Z-MT.

3.3 Physicochemical characterization

Scanning electron microscopy (SEM) micrographs were recorded on apparatus JEOL JEM-200CX at accelerating voltage: 80 keV.

The TEM observations were carried out using transmission electron microscope JEOL JEM 2100 with GATAN Orius 832 SC1000 CCD Camera at accelerating voltage of 200 kV. **X-ray photoelectron spectrum (XPS)** studies were performed in a VG ESCALAB II electron spectrometer using AlK α radiation with an energy of 1486.6 eV. The changes in composition and chemical surrounding in the depth of the films were determined on the basis of the areas and binding energies of C1s, O1s, Zn2p and N1s photoelectron peaks (after linear subtraction of the background) and Scofield's photoionization cross-sections.

The electrokinetic zeta potentials (ζ) of the particles were measured as follows: Suspensions of both ZnO samples of 0.2 g/L concentration (pH 8) were prepared individually by sonication of

ZnO powder for 1 min in water. In a separate experiment, cationic polysaccharide chitosan (oligosaccharide lactate of molecular weight 5000 DA) was adsorbed on the particles surface aiming to increase their surface charge density (0.1 g/L chitosan was mixed with 0.1 g/L ZnO suspension at pH ~ 7-8 and sonicated for 5 min). DLS particle sizer was used to measure the hydrodynamic diameter of the ZnO nanoparticles. A HeNe laser was used as a light source, and the intensity of scattered light was measured by a detector at 173°. Size measurements utilized 12 mm of square disposable polystyrene cuvettes. The electrokinetic zeta potentials (ζ) of the particles were determined by electrophoresis mobility measurements using a disposable folded capillary cell with the same equipment.

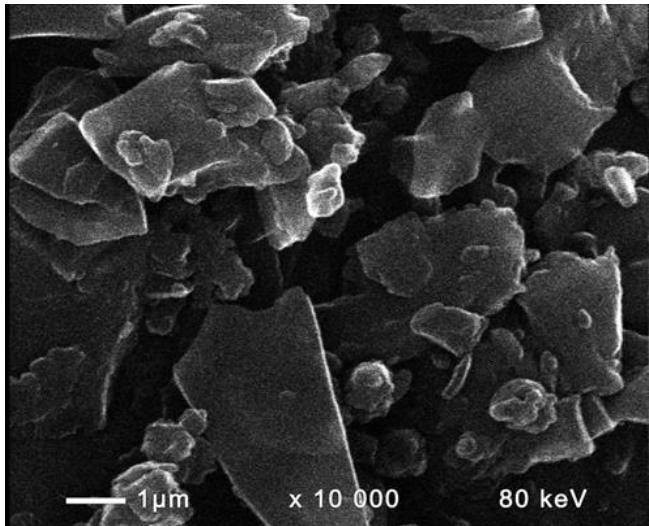


Fig. 1-a SEM photograph of Z-CT powders

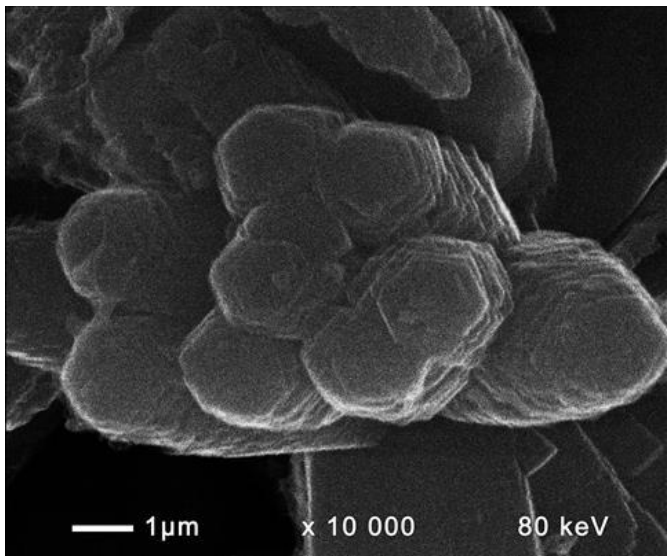


Fig. 1-b SEM photograph of Z-MT powders

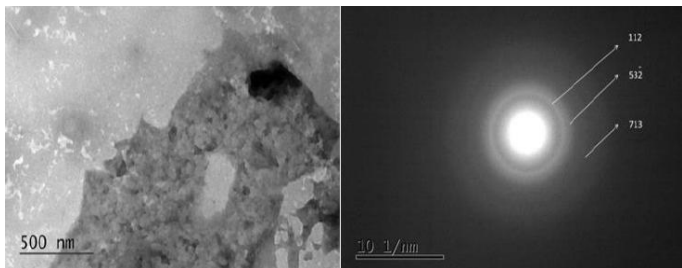


Fig. 2 TEM of Z-CT sample
SAED of Zinc Hydroxide Carbonate
 $\text{Zn}(\text{OH})_6(\text{CO}_3)_2$, (PDF 72-1100)

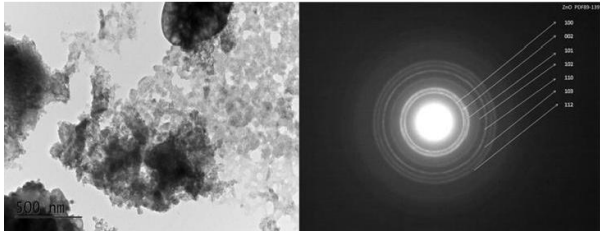


Fig. 3 TEM of Z-MT sample
SAED of hexagonal phase
(PDF 89-1397)

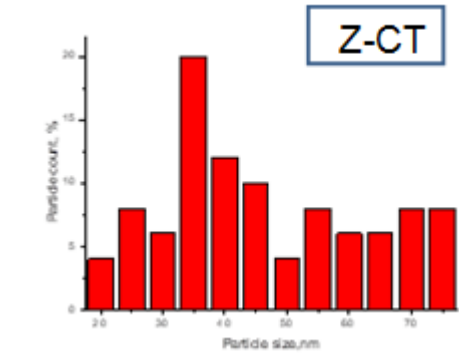


Fig. 4

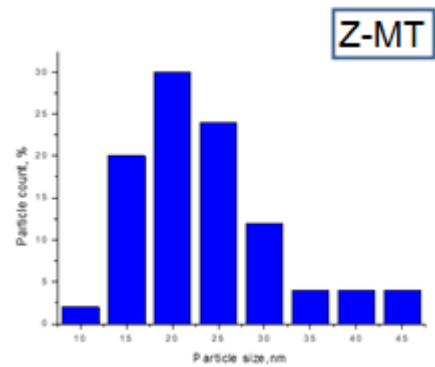


Fig. 5
Particles size distribution
The average particle size values for Z-CT and Z-MT samples are 45 nm and 20 nm, resp.

XPS studies of ZnO particles

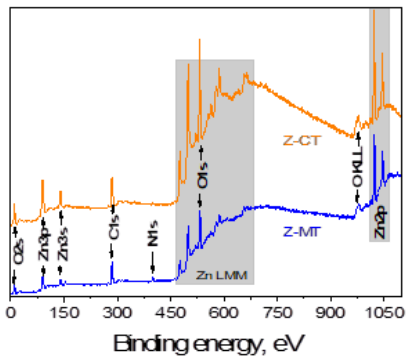


Fig. 6 XPS survey of the ZnO powders

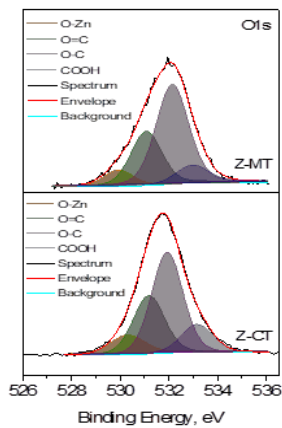


Fig. 7 Deconvolution of O1s spectra of the ZnOpowders

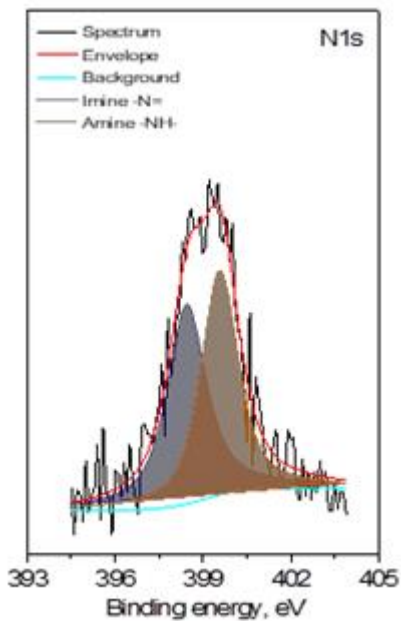


Fig. 8 Deconvolution of N1s spectrum of Z-MT sample

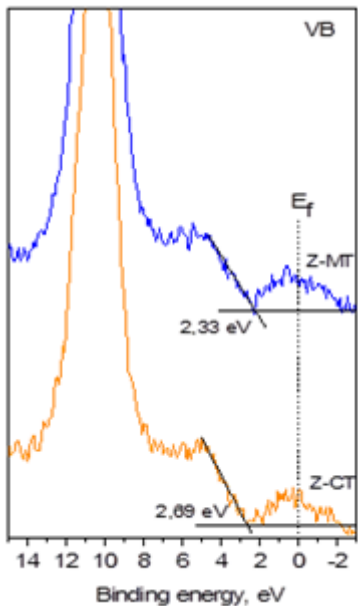


Fig. 9 Valence band (VB) region of ZnO particles

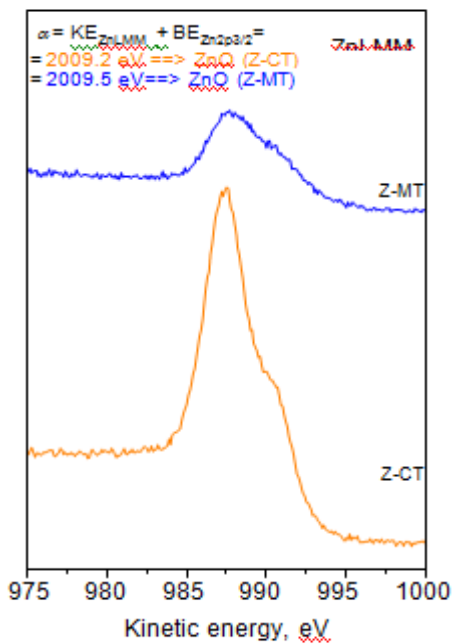
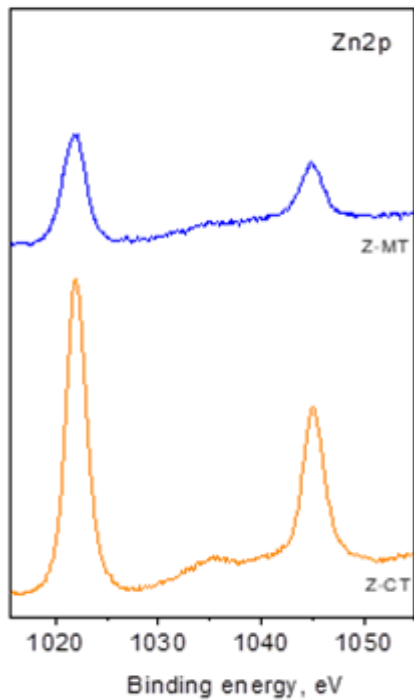


Fig. 10 Zn2p photoelectron spectra and Zn L₃M₄₅M₄₅ Auger area of the ZnO particles

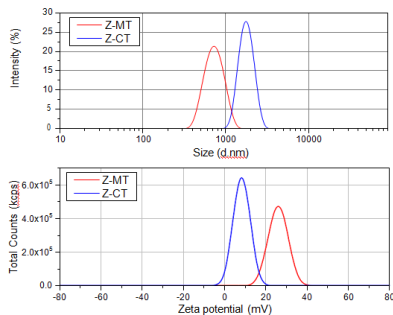


Fig. 11 Particles size distribution and zeta potential of the ZnO

XPS analysis presents that C, O, Zn and N peaks are registered on the surface (Fig 6).

The O1s peaks are deconvoluted by Lorentzian–Gaussian curve fitting into several components corresponded to oxygen in the ZnO lattice (~530.0 eV), oxygen bonds to carbon C=O (~531.0 eV), C-O (~532.0 eV) and carboxyl groups (~533.2 eV). (Fig. 7). The surface composition of the two samples was similar except for the nitrogen peak recorded in the microwave treated one (Z-MT). The high resolution XPS spectrum of N1s shows two peaks with binding energies at 398.5 eV and 399.6 eV, associated with imine (- N=) and amine (-NH-) nitrogen atoms, respectively (Fig. 8).

The valence band (VB) of the ZnO powders was also analyzed (Fig. 9). The results present that the VB edge is shifted to lower binding energies in the microwave treated sample Z-MT, which indicates a decrease in the band gap of ZnO

The binding energies value of the registered Zn2p_{3/2} peaks is 1021.9 eV.

The obtained values of the Auger parameter of the powders is $\alpha = 2009.5$ eV for Z-MT and $\alpha = 2009.2$ eV for Z-CT. (Fig.10) According to literature data, values of this parameter in the range (2009 - 2010) eV corresponds to the Zn²⁺ oxidation state.

Zeta potential measurements

At pH 8, the ZnO particles, prepared using the both microwave and conventional methods showed positive zeta potential values of +24.3 mV and +7.8 mV, respectively. These potentials are below +30 mV, which is generally accepted as a minimum zeta potential value for a dispersion to be electrostatically stabilized. The hydrodynamic diameters of the ZnO particles measured at pH 8 in water suspensions are ~ 750 nm and ~ 1975 nm for the microwave assisted and conventionally prepared samples, respectively, which indicates stronger aggregation in the suspension containing less charged ZnO particles (Fig. 11). From these results, it comes that *Mentha Arvensis* extract is an active reducing agent to convert metal ions into ZnO nanoparticles, but it is not able completely to prevent them from agglomeration at presented conditions. An experiment was performed by adsorption of a low molecular cationic polyelectrolyte chitosan on the particles surface to increase the zeta potential values of the ZnO particles and thus to lower the degree of particles aggregation. The hydrodynamic diameters of the particles were found to decrease from ~750 nm to ~270 nm (zeta potential increased from +24.3 mV to +33.7 mV) for the Z-MT sample, and from ~1975 nm to ~ 370 nm for the Z-CT nanoparticles (zeta potential increased from +7.8 mV to +26.8 mV), respectively.

The results indicated that the particles aggregation could decrease by increasing charge density on the particle surface. It is interesting to note that Sudrajat et al. [H. Sudrajat, S.Babel, A novel visible light active N-doped ZnO for photocatalytic degradation of dyes, J Wat. Proc. Eng.16(2017)309-318] registered zeta potential value + 20 mV at pH 8 for N-doped ZnO nanoparticles (prepared through a combustion reaction from zinc acetate dihydrate and urea), which is close to the zeta potential value for the Z-MT particles, which contain 8% nitrogen

Acknowledgements

This study was accomplished with the financial support of **Project KP-06-H69/8 “ Novel polymer-hybrid materials containing (bio)synthesized metal oxide particles with improved photocatalytic and antimicrobial potential”** of the National Science Fund, Ministry of Education and Science, Bulgaria and contract “**Environmentally friendly anticorrosive hybrid coatings**” is within the Non-currency Equivalent Exchange Bilateral Cooperation between the Bulgarian Academy of Sciences and the Serbian Academy of Sciences and Fine Arts.

Corporate culture and teamwork in the digital era

Nadežda Jankelova

University of Economics in Bratislava, Bratislava, Slovak Republic

nadezda.jankelova@euba.sk

Abstract: The topic of the Fourth Industrial Revolution, the evolution of jobs, and consequently changes in human capital issues is of increasing interest to researchers policy makers, and managers from corporate practice. Corporate culture is a tool for implementing the changes that come with the digital era. Teamwork is part of an adaptive corporate culture that creates the conditions for innovation and technological progress. The present study aims to investigate whether the relationship between Industry 4.0 tools and human resource management in this era (Smart HR 4.0) is mediated by the teamwork climate in the enterprise. The research was conducted by questionnaire survey in 115 industrial enterprises in Slovakia. The respondents were low, middle and top level managers. The PLS-SEM (Partial Least Squares Structural Equation Modelling) method was used to test the proposed hypotheses. The hypotheses were statistically tested at $\alpha = 0.05$ level. The results of the performed analyses were processed using the statistical software SmartPLS 3. The hypothesis of the mediating effect of teamwork in the relationship of Industry 4.0 and SmartHR4.0 was confirmed. Partial mediation was found, with the indirect effect contributing by 32 percent to the total effect. The findings indicate that teamwork is an important mean of support in various innovative management tools, including HR and SmartHR4.0 implementation.

KEYWORDS: CORPORATE CULTURE, TEAMWORK, INDUSTRY 4.0, HUMAN RESOURCES

1. Introduction

The Fourth Industrial Revolution needs no further introduction. Its content, tools, methods are becoming a normal part of work processes. It has brought many challenges in the field of human resources management, which companies and especially HR managers have to gradually adapt to HR processes.

Simultaneously, the topic of the Fourth Industrial Revolution, the evolution of jobs, and consequently changes in human capital issues is of increasing interest to researchers policy makers, and managers from corporate practice [1]. Adoption of the latest automation and digital technologies in manufacturing - such as cyber-physical systems, Internet of Things, cloud computing, Big Data, augmented reality, additive manufacturing, simulation, robotics, cyber-security, systems integration [2] - predict to significantly affect work processes and the work environment [3]. The new era will lead into changes within jobs in terms of content, and new job occupations will emerge. Based on the above fact, higher skill requirements will be demanded from employees. Despite increasing attention being paid to this topic, research addressing the impacts of the fourth industrial revolution on jobs and skills is lacking. It will also be necessary to adapt and build a corporate culture based on teamwork, which contributes significantly to the introduction of technological innovations [4].

The effects of Industry 4.0 are expected to reflect in all areas of scientific progress. Although, it is very difficult to predict certain facts, there are assumptions that human resource management theories will need to rely on building stronger environment, social responsibility and ethical dimensions, as communities and workers demand, that companies increasingly respond to these global challenges more strategically. In the evolving "fight" for talent, companies, that do not respond to a changing environment will have difficulty competing, because branding of employers becomes an essential part of selecting a young generation of workers. Human resource professionals will need to focus more intensively on proactive human resource planning, global and local environmental problem-solving, which will transfer many traditional functions to managers, external service providers using artificial intelligence and robotic technologies. The growing interest in "big data" and a more sophisticated human resource information management system will become an essential part of modern business, as well as greater responsibility for individual strategies, processes and results [5].

Towards both employees and customers, we consider the digitization of analogue data, digital interaction platforms, networking, big data analytics and GDPR to be absolutely essential. Digitization is an absolute priority. Care must be taken to ensure the quality of digitization that makes life easier for the company. Here again, teamwork is coming to the fore, enabling better communication, information sharing, development of different

platforms, simulations in recruitment, gamification. With the help of big data analytics, it is possible to create large databases of employees (even potential ones), track attendance and of different ways of being present at the workplace.

In a study we recently conducted, we investigated which competencies are required from the workforce in Industry 4.0. These are creativity, entrepreneurial thinking, problem solving, conflict resolution, decision making, analytical skills, research skills, efficiency orientation and teamwork. The research highlighted the importance of knowledge development as an indicator of the potential of the company. Employees with entrepreneurial and innovative thinking skills are particularly needed, because they tend to think creatively, take responsibility for their work and are productive. This finding also emphasizes the importance of developing creativity and creative thinking. There is no consensus on well-defined competences responding to the Industry 4.0 environment. However, a critical competency, that future professionals must possess is the ability to use their knowledge in different areas of collaboration to add value [6].

The aim of the present study is to investigate whether the relationship between Industry 4.0 tools and human resource management in this era (Smart HR 4.0) is mediated by the teamwork climate in the enterprise.

2. Theoretical background

Fourth Industrial Revolution

The concept of the fourth industrial revolution, otherwise known as Industry 4.0, was first introduced at the Hannover Messe in 2011, captivating not only manufacturers but also the scientific community. The scientific literature defines the term Industry 4.0 in different ways. Some authors interpret Industry 4.0 as the digital integration of manufacturing processes, in which production processes are automated and products, equipment and services are interconnected. Ongoing industrial revolution is driven by the Internet, through which not only humans but also machines will communicate with each other in a cyber-physical system. Other authors argue that the evolution of manufacturing processes is driven by market demand for more efficient technologies and processes, cost reduction and quality standards, or technological advances [6]. While the third industrial revolution (Industry 3.0) focused on the automation of specific business processes, Industry 4.0 focuses on the digital transformation of companies. This implies a comprehensive digitization of all physical assets and the creation of new digital ecosystems, including the value chain. The generation, analysis and seamless communication of data represent the core benefits, that Industry 4.0 brings, with the aim of bringing together a wide range of new technologies to create value. In the

context of Industry 4.0, new technological systems will be intertwined with organizational processes to transform industries, which will lead into the interconnection of people, machines, and smart objects in real time [7].

A summary of the concepts underpinning the future development of Industry 4.0 will be presented in the Measurement section.

Smart HR 4.0

The theoretical conceptual model of Smart HR 4.0 (SHR 4.0) is based on a study by Sivathan and Pillai [8]. The model has not yet been statistically investigated; we consider it to be original and unique within the scientific literature. By analyzing the model, we will establish the dependence of each variable, thus showing the existence of relationships between them and revealing the importance of the impact of Industry 4.0 on human resource management

The study of the above authors in the proposal of the conceptual model SHR 4.0 highlights the importance of the impact of the organizational structure of the enterprise and the leadership style of the manager on the management of human resources in the conditions of Industry 4.0. In terms of organizational structure, the above authors state that organizational structures should be agile and flat, thus creating a favorable environment for the implementation of SHR 4.0. A flat hierarchy will reduce the number of communication layers and speed up decision making. Decentralization of power will allow teams to work autonomously and adapt immediately to the requirements of a specific project. Leadership styles need to be more open, fostering a culture of learning and innovation; they should also focus on expanding knowledge and rewarding unconventional "out-of-the-box" thinking. The role of management will be to initiate changes in the organizational structure of the business so that conflicts do not arise between multi-generational groups of workers. SHR 4.0 will require upgrading technology in line with the long-term goals of the business to ensure that the most talented from Generation Y and Z are attracted. Based on an in-depth search of the academic literature, there is an apparent consensus on the impact of the Industry 4.0 concept on human resource management, as well as studies by Sivathan and Pillai [8], who describe in detail the issue of the impact of Industry 4.0 on human resource management.

Teamwork climate

Teamwork climate (TWC) is perceived as an important determinant of business performance in general (Zaheer et al., 2018). It represents the perception of cooperation among staff (Weng, Kim and Wu, 2017). Studies have found that TWC is influenced by the culture of communication and mutual trust (Ginsburg and Bain, 2017). However, it also contributes to employee performance (Bogan and Dedeoglu, 2017), but especially to the rate of innovation (Fay et al., 2014) and even innovation in human resource management (Jiang, Wang and Zhao, 2012). SmartHR 4.0 is just such an innovative form of management and therefore we assume a correlation between TWC and Smart HR 4.0.

Based on the above research model, we formulate hypotheses which will be analyzed in detail in the next part of the research.

H1: We hypothesize that Industry 4.0 (I4.0) technologies are positively related to Smart HR 4.0 (SHR4.0)

H2: We hypothesize that the relationship between Industry 4.0 (I4.0) technologies and Smart HR 4.0 (SHR4.0) is mediated by teamwork climate in the firm (TWC).

3. Methodology

We used a questionnaire survey designed to indicate the level of importance and attitude of the respondents, to the study of each of the variables under review. Before running the final version of the questionnaire survey, we conducted a pilot survey. The survey was conducted with managers of 5 companies operating in the Slovak

Republic, while each contacted manager commented on the structure and content of the questions in the questionnaire during an online meeting via MS Teams. During the individual meetings, we received valuable advice and comments, which were subsequently incorporated into the final form of the questionnaire.

An electronic questionnaire was constructed via Google Forms, which we then started to send to industrial enterprises via LinkedIn services, and also via email communication. The electronic questionnaire was sent to 3061 industrial enterprises operating in the Slovak Republic based on the Finstat database, with the final number of completed questionnaires amounting to 115 responses, which also represented the final sample of the research. The respondents were managers of low, middle and top level of industrial enterprises operating in the Slovak Republic.

In order to verify the proposed hypotheses and to better understand the relationships between the selected constructs, we used the PLS-SEM method (Partial Least Squares Structural Equation Modelling). That method allows multiple hypotheses to be tested simultaneously under both direct and indirect effects in a complex system. We chose to use it for several reasons: the first is the relatively small sample size, other reasons are the complexity of the research model, the focus of the study on the prediction of the dependent variables, and the use of latent variable scores for predictive purposes [9]. We used the following methods to test the reliability and validity of the model:

- Methods of convergent validity and reliability: Cronbach's Alpha (Cronbach Alpha), Sperman's rho (ρ_A), Composite reliability (CR), Average Variance Extracted (AVE) - all of these indicators meet the requirements for its confirmation.

- Discriminant validity methods: Heterotrait-Monotrait Ratio (HTMT) and Cross Loadings - all of these indicators meet the requirements for validation.

- Collinearity statistics methods: Variance Inflation Factors (VIF) - all VIF values are below 5, indicating that there are no collinearity problems between predictive indicators.

- Coefficient of Determination (R-Squared) and Adjusted Coefficient of Determination (Adjusted R-Squared)- The coefficient of determination (R²) indicated that the independent latent variables explain 63.1% of the variance of SHR4.0, which means that approximately 63.1% of the variation in SHR4.0 is due to the latent variables in the present model. Adjusted R² provides more efficient values for the correlation between variables.

To check the fit of the structural model, we used the Model Fit: Standardized Root Mean Square Residual (SRMR) calculation - The value of SRMR is 0.095. SRMR values should be less than or equal to 0.100, indicating acceptable model fit [10].

Hypotheses were statistically tested at $\alpha = 0.05$ level. The results of the performed analyses were processed using the statistical software SmartPLS 3.

Measurements

The content of the individual variables is presented in the following table for clarity.

Table 1 Content of individual variables (1 = completely disagree, 5 = totally agree)

Industry 4.0 technologies (is the rate of implementation in the enterprise high?)	Smart HR 4.0 (is the rate of implementation in the company high?)
Additive manufacturing	Posting job offers on smart/mobile apps
Augmented reality	Automated CV search using AI and Big Data
Autonomous robots	Real-time remote video interviewing on a fast data network
Big data and analytics	Acclimating new hires through augmented reality
Cloud computing	Identifying employee skill gaps through artificial intelligence
Cyber protection	Identifying low-performing workers based on Big Data
Horizontal and vertical integration	Anytime, anywhere virtual training
Internet of Things	Continuous feedback
Simulations	Retaining workers through new value propositions and internal opportunities
	Smart IoT-based applications and devices for real-time health tracking to eliminate sick leave
Teamwork climate perceived quality of cooperation between staff	
All team members can ask questions if there is something they do not understand.	
Staff members do receive the support they require from other staff in the performance of their duties.	
The contribution of employees is positively perceived in our company.	
Disagreements in the team are resolved adequately, it is not important who is right, but what is best for the task.	
Team members work together as a well-coordinated team regardless of their functional positions.	
In our company, it is not difficult to express myself critically if I perceive problems in performing performance.	

4. Results

To understand the subjective perception of the amount of each variable, we present both descriptive statistics and correlation analysis.

Source: own elaboration

Note: I4.0 = Industry 4.0, Smart HR 4.0 = Smart Human Resources 4.0, TWC = Teamwork Climate

Based on the above statistics, the reader can visualize respondents' perceptions on different aspects of the summary variables using different measures of descriptive statistics.

The inter-correlations of the summary variables are shown below.

Table 2 Descriptive statistics of the reviewed variables

Technology Industry 4.0 (I4.0)									
	1	2	3	4	5	6	7	8	9
Examined sample	115	115	115	115	115	115	115	115	115
Arithmetic mean	2,69	2,38	2,88	3,21	3,23	4,18	3,22	3,20	3,30
Median	3,00	2,00	2,00	3,00	3,00	4,00	3,00	3,00	3,00
Modus	1	1	1	1	4	5	4	4	1
Standard deviation	1,43	1,39	1,76	1,67	1,65	1,40	1,48	1,54	1,71
Variance	2,04	1,94	3,11	2,80	2,71	1,96	2,19	2,37	2,92
Smart HR 4.0									
	1	2	3	4	5	6	7	8	9
Examined sample	115	115	115	115	115	115	115	115	115
Arithmetic mean	4,20	4,68	2,95	3,38	3,84	4,03	3,77	3,84	4,06
Median	5,00	5,00	3,00	4,00	4,00	4,00	4,00	4,00	4,00
Modus	5	5	1	4	5	4	4	5	5
Standard deviation	1,46	1,07	1,63	1,57	1,33	1,29	1,35	1,44	1,38
Variance	2,14	1,15	2,66	2,45	1,77	1,67	1,83	2,08	1,90
Teamwork climate (TWC)									
	1	2	3	4	5	6			
Examined sample	115	115	115	115	115	115			
Arithmetic mean	3,97	3,99	3,87	3,88	4,03	4,21			
Median	4,00	4,00	4,00	4,00	4,00	4,00			
Modus	5	5	5	4	4	4			
Standard deviation	1,42	1,42	1,50	1,41	1,36	1,23			
Variance	2,02	2,01	2,24	1,99	1,85	1,50			

Table 3 Inter-correlations of summary variables

Kendall's tau_b		I4.0	TWC	SmartHR4.0
I4.0	Correlation Coefficient	--		
	Sig. (2-tailed)			
TWC	Correlation Coefficient	0,612**		
	Sig. (2-tailed)	<0,001		
RLZ	Correlation Coefficient	0,612**	0,507**	--
	Sig. (2-tailed)	<0,001	<0,001	.

** The correlation is significant at the level of 0,01

Source: own elaboration

Note: I4.0 = Industry 4.0, Smart HR 4.0 = Smart Human Resources 4.0, TWC = Teamwork Climate

Based on the results of the analysis using Kendall's tau_b correlation coefficient, we found that all relationships between the summary variables are statistically significant and there are positive correlations between them.

In order to analyze the mediation effects, we use the bootstrapping method to quantify the influence of mediating variables [11]. We constructed 2 hypotheses that we will analyze: H1- the relationship between I4.0 and SHR4.0 and H2- the relationship between I4.0 and SHR4.0 mediated by TWC. The mediation model via TWC between the relationship of I4.0 and SHR4.0 is described in the following table.

Table 4 Paths of latent variables found

Paths of latent variables	Effect type	Original Sample (β)	Sample Mean (β)	Standard deviation	T-value	P-value
I4.0 -> TWC	Direct effects	0.574	0.574	0.075	7.696	0.000
I4.0 -> SHR4.0		0.490	0.495	0.077	6.406	0.000
TWC -> SHR4.0		0.384	0.384	0.103	3.741	0.000
I4.0 -> TWC -> SHR4.0	Total indirect effects	0.232	0.232	0.051	4.570	0.000
I4.0 -> SHR4.0	Total effect	0.723	0.727	0.050	14.318	0.000

Source: own elaboration

Note: I4.0 = Industry 4.0, Smart HR 4.0 = Smart Human Resources 4.0, TWC = Teamwork Climate

The table above shows the coefficients of the structural model. The results of the table indicate the facts that I4.0 has a significant effect on SHR4.0 ($\beta=0.490$, $t=6.180$, $p<0.05$), thus confirming 1 hypothesis, namely that Industry 4.0 (I4.0) technologies are positively related to Smart HR 4.0 (SHR4.0). The path from TWC to SHR4.0 is also significant ($\beta=0.384$, $t=3.658$, $p<0.05$). The path from I4.0 to TWC is also significant ($\beta=0.574$, $t=7.696$, $p<0.05$). All the direct effects variables show T-values greater than 1.96 and P-values less than 0.05 (significance level = 5%). Thus, the prerequisite for the implementation of mediation is given.

Further results suggest that the latent variable TWC represents a partial mediation of the relationship between I4.0 and SHR4.0. Hypothesis 2, which relies on TWC as a mediator, was confirmed. The direct effect is $\beta=0.490$ and the indirect effect is $\beta= 0.232$,

indicating that the mediating effect of TWC in the I4.0 -> SHR4.0 relationship is 32%, and the direct relationship is 68%.

The mediation effect is significant ($\beta= 0.232$, $t= 4.215$, $p<0.05$).

Discussion

Hypothesis 1 was confirmed in the research. Our results are in line with the results of Sharma [12], which show that SHR4.0 driven by modern technology as well as next-generation employees is changing HR processes in the Industry 4.0 era. A recent study by Sivathan and Pillai [8] adds that human resource performance can be achieved by the impact of Industry 4.0 technologies on SHR4.0. The study of Piwowarwa-Sulej [13] adds that the impact of technology on SHR4.0 also translates into the enhancement of collaboration between human capital and machines. Our study also confirmed the results of previous researches that there is a significant dependency between modern Industry 4.0 and SHR4.0 technologies and it causes changes within HR processes.

Hypothesis 2 was confirmed in the thesis research.

This is not a surprising finding as TWC is highlighted by many authors as an important enabler within various innovative management tools. They state that a wider use of teamwork leads not only to a higher number but also to a higher level of innovations. The role of teamwork in organizational innovation is also highlighted by other authors [17, 18]. Factors that tend to be cited as supportive in this context are knowledge sharing, mutual learning and autonomy, leading to responsibility and motivation.

Our results are supported by Malavik and Mohan's study [14], which shows the dependence of effective SHR4.0 implementation on corporate culture and its teamwork aspect. These factors will enable human resource departments to play a more strategic role in the overall growth of businesses. The research findings of Shamim et al. [15] point out that in addition to important elements such as organizational structure and leadership style, HR processes also create conditions for compatibility with Industry 4.0 by continuously promoting innovation and learning. Rehman et al. [16] complement the conceptual model with a corporate culture that significantly influences innovation and facilitates the implementation of Industry 4.0.

Businesses need to implement a Smart HR4.0 strategy to cope with this turbulent environment, as human capital is the main driver behind changes in the way business processes work. Although this is a technological revolution, the working tool of managers in the Industry 4.0 era will be mainly human.

Modern technologies will automate a significant part of human resource management processes with the intention of creating more efficient and leaner HR teams. The strategy should also ensure an influx of the younger generation with diverse skills into businesses. Incorporating both the organizational structure of businesses and the leadership style of managers would support the necessary changes in the effective implementation of SHR 4.0, enabling HR departments to play a more strategic role in business competitiveness, process efficiency and overall growth.

The results of our study add to the so far fragmented and inconsistent knowledge on the promotion of teamwork in the implementation of technological advances in human resource management and strongly point to some benefits for managers and practitioners. The findings suggest that it is important for firms to support their own employees in the creation and implementation of new solutions, as this has been shown to have a direct positive impact on firm performance. The focus needs to be not only on technology, but especially on soft skills, namely team environment and opportunities for quality teamwork and the subsequent implementation of team solutions, as this is an area where significant gaps have been demonstrated. Thus, the effect of technology in HR processes can be intensified by orienting

managers to work in teams, especially in the area of knowledge and mindset. Such support benefits both parties, businesses through better results and competitiveness, employees through higher job satisfaction.

Despite the usefulness of the findings, this study has some theoretical and practical limitations that could indicate the direction of further research. Although this exploratory model proved to be very useful and convincing, future research could consider including additional variables in the framework to help better understand the relationships examined. Another limitation of our study is its restriction to firms operating in the Slovak market. Given the nature of the relationships examined and their treatment in many studies, we can assume that the findings have universal validity. Another limitation of this study was the use of a cross-sectional survey design and the questionnaire form of data collection.

Conclusion

If businesses want to succeed in Industry 4.0, they should consider new approaches in people management not only through new technological innovations, but also through the implementation of soft management tools and teamwork. The ability to work in teams will become crucial. Expertise is of course important, but recruitment and selection should emphasize openness to change, adaptability, teamwork, creativity. In order to meet the various challenges that lie ahead, businesses need to pay considerable attention to strategic workforce planning.

Acknowledgement

This article is a partial output of research project VEGA no. 1/0010/23 entitled Adaptability of corporate culture - a factor supporting resilience and sustainability of enterprises in Slovakia in the post-covid period.

References

- [1] L.Gehrke, et al.: Towards Industry 4.0: mapping digital technologies for supply chain management-marketing integration. *Bus. Process Manag. J.*, 25 (2019)
- [2] T. Melason: What Industry 4.0 means for manufactures, <https://aethon.com/mobile-robots-and-industry4-0/>, last accessed 2024/8/4 (2018).
- [3] G. Chryssolouris, et al.: Manufacturing systems: skills and competencies for the future. In: *Procedia CIRP*, pp. 17-24. ScienceDirect (2013).
- [4] D. Puhovichová, N. Jankelová: Changes in Human Resource Management in Context of Impact of the Fourth Industrial Revolution. *Economics, Finance and Business Management. Proceedings of Scientific Papers Issued to Mark the 80th Anniversary of the University of Economics in Bratislava. Bratislava : EKONÓM*, 2020, p. 440-449. ISBN 978-80-225-4749-9.
- [5] A. Nankervis, M. Baird, J. Coffey, J. Shields. *Human resource management. Strategy and practise*. 2019
- [6] D. Puhovichová, N. Jankelová: Effects of Industry 4.0 on Human Capital and Future of Jobs / EDAMBA 2021: International Scientific Conference for Doctoral Students and Post-Doctoral Scholars. Conference Proceedings; Bratislava : EKONÓM, 2022, p. 400-412. ISBN 978-80-225-4930-1.
- [7] PWC. Industry 4.0: Building the digital enterprise, <https://www.pwc.com/gx/en/industries/industries-4.0/landing-page/industry-4.0-building-your-digital-enterprise-april-2016.pdf>, last accessed 2024/8/4 (2016).
- [8] B. Sivathanu, B. Pillai: Smart HR 4.0—how industry 4.0 is disrupting HR. *Human Resource Management International Digest* 2018, 26, (4): 7-11.
- [9] J.-M. Becker et al. Estimating Moderating Effects in PLS-SEM and PLSc-SEM: Interaction Term Generation*Data Treatment. *Journal of Applied Structural Equation Modeling*. 2018, 2 (2): 1-21.
- [10] J.F. Hair: When to use and how to report the results of PLS-SEM. *European Business Review*. 2019, 31 (1): 2-24
- [11] J.H. Bolin: Review of Introduction to mediation, moderation, and conditional process analysis: A regression- based approach. *Journal of Educational Measurement*, 2014, 51, (3): 335-337.
- [12] P.N. Sharma et al. PLS-Based Model Selection: The Role of Alternative Explanations in Information Systems Research. *Journal of the Association for Information Systems*. 2019, 20 (4): 346-397.
- [13] K. Piwowar-Sulej. Human resource management in the context of Industry 4.0. *Organization and Management scientific quarterly*. 2020, 1 (49): 103-113.
- [14] S.K. Malavik, P.A. Mohana. A global revolution of work with Smart Hr 4.0. *Psychology and education*. 2021, 58 (2): 2829-2835.
- [15] S. Shamin et al. Management approaches for Industry 4.0: A human resource management perspective. 2016 IEEE Congress on Evolutionary Computation (CEC) [online], 2016. Available on: https://www.researchgate.net/publication/311251654_Management_approaches_for_Industry_40_A_human_resource_management_perspective, last accessed 2024/8/4 (2016).
- [16] H.M. Rehman et al. Organization innovation for Industry 4.0 in management perspective. In *International Conference on Business, Accounting, Finance and Economics (BAFE, 2019)* [online]. Available on: https://www.researchgate.net/publication/339828286_Organization_innovation_for_Industry_40_in_Management_Perspective, last accessed 2024/8/4 (2019).
- [17] D. Fay et al. Teamwork and Organizational Innovation: The Moderating Role of the HRM Context. *Teamwork and Organizational Innovation*, 2014, Vol. 00, N. 00, 1-17. doi: 10.1111/caim.12100.
- [18] J. Jiang et al. Does HRM Facilitate Employee Creativity and Organizational Innovation? A Study of Chinese Firms. *International Journal of Human Resource Management*, 2012. 23, 4025-47.

Impact of brain thinking preference on engineering project performance

¹Saja Al-Btoosh, ²Salloom A. Al-Juboori

²Professor in Mechanical Engineering (B.Sc, PhD, M. Sc., Leeds University, U.K) / Engineering Faculty /Mutah University /Jordan

Email: ¹Sajanief94@gmail.com , ²Email: sajabury@mutah.edu.jo

Abstract: The main objective of this study is to determine the extent of the effect of brain thinking preference (brain quadrants) on engineering project performance. A descriptive approach methodology was used to achieve the objective of this work by developing a special questionnaire contains many tracks. The data collected by the questionnaires with a sample size of (500) from different engineering projects in public and private sectors were analyzed using a statistical package for the social science (SPSS –V21). The main results showed that the most influencing criteria on the project's performance was "the freedom to choose a job", which had an impact on the project's performance by 84.3%, and the respondents' answers in favor of this parameter was 78.2%. This indicates that the parameter of the independent coefficient (preference for mental thinking) raises the level of project performance and thus directly affects the project performance measures which are: The effectiveness, efficiency, employee loyalty and project quality. While the lowest effect is the parameter "the creativity and skill "which has 61.3%, with respondents answers in favors of this parameter is 36.6%. Also the results showed that when employees have the freedom to choice their own preferred job will have a great impact on project quality and thus increasing the percentage of profits. So, giving employees the freedom to choose a job and measuring the extent of creativity and skill will give the opportunity to increase a job performance. Also, results have shown that spreading awareness of the preference for cerebral thinking and moving towards the preferred cerebral quadrant of a person's brain will unleash the best level of creativity, skill and conscious future outlook in all scientific and practical fields.

KEYWORDS: BRAIN THINKING PARAMETERS, QUADRANTS, PROJECT PERFORMANCE PARAMETERS, EMPLOYEES

Research background: The purpose of this work was to better understand what constitutes a competent thinker and decision-maker, the competences required to be such that to study the variables that may affect a person's brain decision. Furthermore, making wise decisions is essential for thriving in the business world, particularly when it comes to project performance. Wise decision-making calls for managers to be aware of their surroundings and how they may affect the operations of their projects [1,2]. The brain is a complex organ that controls thought, memory, emotion, touch, motor skills, vision, breathing, temperature, hunger, and every process that regulates our body. Together, the brain and spinal cord that extends from it make up the central nervous system or CNS. To be effective in today's complex and changing world, managers need new insights and skills that up- end conventional thinking about human potential, trust, energy, initiative, and commitment. Neuroscience holds the key to accomplishing this goal. The old belief was that all sensory input went directly to the brain, where a behavioral response was shaped. The new findings show that the gut-brain and the heart-brain are involved first. Brain Thinking: Way of solving problems that combines a rational and convergent approach with creative and divergent thought processes. It includes four fundamental dimensions: creativity, future vision, intelligent opportunism, and thinking in the time [3,4]. Understood the concept of brain dominance and knew that there were three different types of thinking according to the dominant half:

- The thinking style associated with the right half.
- The thinking pattern associated with the left half.
- Integrated thinking style.

Figure (1) is a schematic representation of Herrmann's metaphoric whole brain model and functions associated with each part.

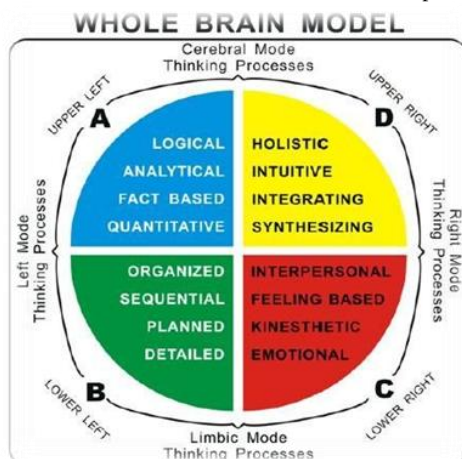


Figure (1) Herrmann's whole brain model [10]

According to [6] these days' lack of brain thinking among managers is obvious. Moreover [7,8, 9] they added that organizations are being involved in competitive and chaotic situations. Because of that, the researcher focused on this problem throughout the current research to explore whether the decisions are being made based on strategic and brain thinking preferences of the project performance of employees, so this work will explain the impact and understanding of the Brain thinking preference and produce a finding of practical relevance through answering the following main question: What is The Impact of Brain thinking preference on project performance?"

Quadrant A [10]: (The Left Cerebral Brain Thinking), is associated with logical thinking, analysis of facts and processing numbers, technical, problem solving. A person with a quadrant dominance is rational and realistic, thinks critically, and likes to deal with numbers and technical matters opinions and feelings are generally not considered as important as facts. People with a strong preference to think in these ways are also attracted to professions that require such thinking. Examples of such careers include technical, legal, and financial areas, including accounting and tax law, engineering, mathematics, and some middle management positions.

Quadrant B (The Left Limbic Brain Thinking) the (B) quadrant has some similarity to the (A) quadrant in that both are left-brain thinking preferences, deals with planning, organizing facts and activities, and careful detailed review, administrative conservative and controlled. A person who relies heavily on quadrant (B) thinking is well organized, reliable, and neat. These people like to establish plans and procedures and get things done on time.

Quadrant C (The Right Limbic Brain Thinking), is associated with interpersonal relationships and affects intuitive and emotional thought processes. (C) quadrant individuals are sensitive to others and enjoy interacting with and teaching others they are typically emotional and expressive, outgoing, and supportive of others. They are concerned with people's feelings more than with facts and procedures. People with single-dominant profiles in the (A) or (B) quadrants probably see individuals with strong (C) quadrant preferences as being very "touchy-feely". Individuals with single dominant (C) profiles would be very "feeling" and people oriented.

Quadrant D (The Right Cerebral Brain Thinking), is associated with conceptualizing, synthesizing, and integrating facts and patterns, with seeing the big picture rather than the details. Words that describe this quadrant are artistic, holistic, imaginative, synthesizer, and conceptualizer A person with a quadrant (D) preference is visionary and imaginative, likes to speculate, break the rules, take risks and may be impetuous, these people are curious and enjoy experimentation, playfulness, and change.

How Does Your Profile Compare to Others? :-

Herrmann's research led to the development of a scientifically validated instrument that can quantify the degree of thinking preferences for specific modes within the whole brain model. The Herrmann dominance instrument is a questionnaire consisting of (120) items that quantifies mental thinking preferences. The results of the questionnaire are a visual plot of the thinking preferences of the individual accompanied by a data summary sheet, the model gives a breakdown of what each individual selected and registers it according to the different quadrants [11]. Finally, these data are complemented by an explanation report that describes and highlights important aspects of the profile in order to understand the profile, it is important to explain (quadrants and modes of thinking, the preference code, the adjective pair data, the profile score) shown in figure (2).

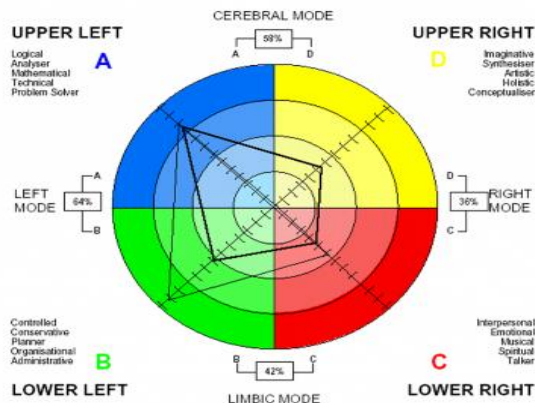


Figure (2) The Ultimate Guide to HBDI, Herrmann Brain Dominance Instrument [11]

Quadrants and Modes of Thinking: Thinking preferences are measured not only by the four quadrants (A, B, C or D) [12], but also by four modes (upper left, lower left, upper right and lower right). The upper modes right and left, combining quadrants (A and D) quadrants, are more cognitive and intellectual, preferring thinking in abstract, conceptual modes. The lower modes combining the (B and C) quadrants, are grounded and emotional in nature. These modes often prefer visceral, "gut" and concrete approaches [9]. The left modes combining the (A and B) quadrants, prefer concise efficient processes with realistic, disciplined and orderly approaches. The right modes combining the (C and D) quadrants, include key mental processes such as intuitive and perceptive thinking, as well as idealistic, expressive and open approaches, all cases of these styles were detailed according to

A and B Dominant Profiles: An (A and B) dominant profile is the profile of the left brain thinker [9]. There is an internal coherence between the (A and B) quadrants. That is, they tend to reinforce each another. For example, the analytical and logical processes of the (A) quadrant support the structural and procedural characteristics of the (B) quadrant. This is a very common profile, especially among males. People with this profile are logical, analytical, technically oriented, and good problem solvers. They are also good planners, organizers, and implementers. They do not shy away from administrative details.

A and D dominant Profiles: The (A and D) dominant profile is the profile of the cerebral thinker. The great minds and great inventors of our time exhibit these two thinking styles. People with this profile have all of the thinking preferences of the (A) quadrant plus a preference for the conceptual, creative, and integrative styles. They like to invent things and new ways of doing things and are quite happy to be individual contributors.

C and D Dominant Profiles: The (C and D) dominant profile is the profile of the right-brain thinker. People with this profile are creative and inventive because of their (D) quadrant preferences but are also people. They are intuitive, holistic, and interpersonal. Entrepreneurs, teachers, salespersons, and trainers often exhibit this thinking style pattern.

B and D Dominant Profiles: So far we have talked about double dominant thinking styles where the two thinking styles were compatible with each another. They were both right brain, or left brain, or cerebral, or limbic. Interesting cases arise when the two dominant thinking styles are fundamentally different from one another, either (B and D) or (A and C). This individual has a double dominant profile, but interestingly, it is across the diagonal between quadrants (B and D). The person was an interior designer. Think of this person in a project manager's role. We would guess that she or he would be good at seeing the "big picture" of the project and at developing project strategy, but she or he would also be interested in doing detailed implementation planning as well. In other words, she could see both the forest and the trees.

B and C Dominant Profiles: This is the profile of the limbic thinker. Because of his or her dominant [18] quadrant score this person has a strong preference for conserve thinking and the need for structure and detailed procedures for all that he or she does. The dominant (C) quadrant score indicates a strong preference for interpersonal relations and concern for how the individual will react to situations. These people are results oriented, and they respond well to deadlines and enjoy working with others.

A, B, and C Dominant Profiles: Simply put, this is a left brain individual with the added concern for the individual. He or she makes a good team player and enjoys analyzing data, building things, and seeing things work. Technical managers will often display this thinking style profile.

A, B, C, and D Dominant Profiles: Only about 3 percent of the population exhibits a preference for all four quadrants. People with this profile are well balanced and can understand the thinking modes of all types of people. They are good facilitators and can deal effectively with every situation. They are adaptable and make good leaders because of their whole brain approach and even handed ways. the successful CEO will often display a four quadrant thinking preference.

Research Objectives: The main aim of this study is to find the extent of the "Impact of Brain thinking preference on engineering project performance." While the sub-objectives are to study the followings:

1. Investigating the impact of the brain thinking parameters (creativity and skill, future vision, intelligent opportunism, mental tendencies, freedom to choice the job) on loyalty of employee.
2. Identifying the individual of brain thinking parameters on engineering project performance.
3. Identifying the level of brain thinking preference awareness of the employees.
4. Determine the overall brain thinking parameters impact on engineering project performance.

Research Variables: As can be seen from the research model in figure (3), Brain thinking is the independent variable, and its dimensions are (creativity and skill, future vision, intelligent opportunism, mental tendencies, and freedom to choice the job). Also, these dimensions were adopted by well-known researchers in the field of management [13,14,15]. Which are hypothesized to affect the dependent variables of the engineering project performance (i.e., efficiency of employee, effectiveness employee, project quality, loyalty of employees). The model in figure 3 represents the predicted association and relationships between brain thinking and organization performance.

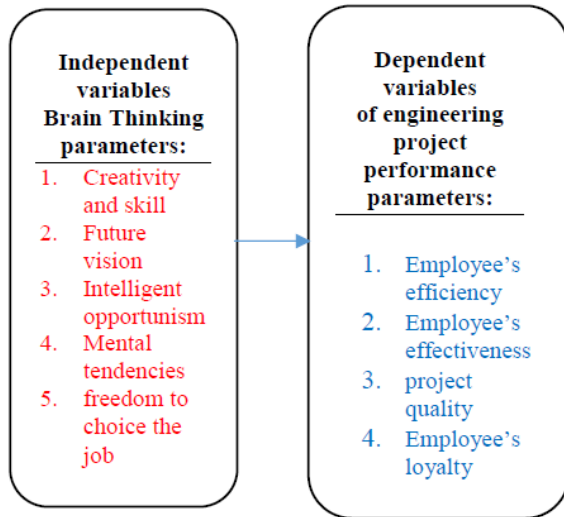


Figure (3): The Independent and dependent variables

Research Population: The study population consisted of all employees in the public sector. The total number of them was (4500) employees, surveying this large of population would consume considerable budget, time, and effort. For this reason, a study sample is calculated based on this number using a formula, which is expressed [16].

$$n = \frac{\hat{n}}{1 + \frac{z^2 \times p^*(1-p^*)}{\epsilon^2 N}} \dots \dots \dots (1)$$

Where:

$\hat{n}$ – Number of the sampel magnitude related to unlimited population

p^* – The proporation of the population

ϵ -The errors frontier

z -The score

N -The realistic population number

$\hat{n}$ can be computed vie the expression: (calculator. net,2023)

$$\hat{n} = \frac{z^2 \times p^*(1-p^*)}{\epsilon^2} \dots \dots \dots (2)$$

Substitung ϵ as 0.05, p^* as 0.5, and z as 1.96 can lead to $\hat{n}$:

$$\hat{n} = \frac{1.96^2 \times 0.5(1-0.5)}{0.05^2} = 384.2$$

Thus, the number of study sample equals will be calculated by substitung N as 4500 persons which is calculated using the equation (1):

$$n = \frac{384.2}{1 + \frac{1.96^2 \times 0.5(1-0.5)}{0.05^2 \times 4500}} = 354$$

As shown in the results of the calculation of the formula, of the sample for this study is 354, but a study sample of 500 was taken in this work to be more convenient for reducing the error.

Sampling and Sampling Frame: After determining the size of the population on which the study will be conducted, estimated at (4500) employees. The sample was selected randomly with a total number of (500) respondents based on the (500) questionnaires, which it was distributed electronically via an electronic link using social networking sites (Facebook, and WhatsApp), and was also distributed on paper in the public and private sectors, on 2023.

Questionnaire Design: The questionnaire consists of three parts; first, the title of the study that includes the important question as filtering question which is (Did your mental orientation influence the choice of your preferred type of work?), If your answer is YES, please proceed to the next question. Part two consists of six questions about the respondent's demographic profile; included is an

instrument that uses nominal and ordinal scales, such as Gender, Age, Qualification, Experience in work, Business sector, and regional distribution. Part three of the questionnaire was designed to measure variables, (Creativity and skill, future vision, intelligent opportunism, Mental tendencies, freedom to choice the job, efficiency of the employee, effectiveness of the employee, project quality, and loyalty of employees) [10].

Validity of the Study Tool: The study tool (questionnaire) was presented to a group of arbitrators with academic competence and experience in a number of public and private Jordanian universities. Their opinion on questionnaire and ascertaining its suitability, in terms of the number of paragraphs, clarity, accuracy, coherence, and coherence, and any other notes. They deem it appropriate, to delete, change, or add and consider this review process and the subsequent correction and modification of most of them as a test of the apparent validity of the tool.

Reliability of the Study Tool: To ensure the stability of the tool, Cronbach's Alpha was calculated using the formula-3 and the results were shown in figure (4). It is clear all the values are greater 0.6, which indicate the model of the study is valid [17].

$$\alpha = \frac{N \cdot \bar{c}}{\bar{v} + (N-1) \cdot \bar{c}} \dots \dots \dots (3)$$

Where:

N - number of items

$\bar{c}$ -bar the mean internal variance between items

$\bar{v}$ -bar the mean of the total variance of the items

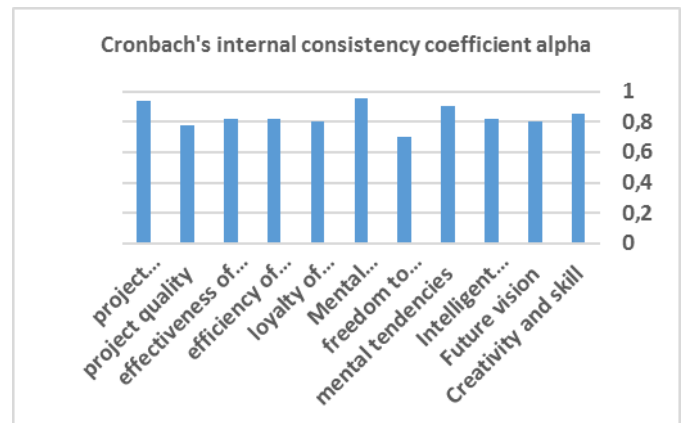


Figure (4) Cronbach's internal consistency coefficient alpha

Statistical Analysis: A Five-point Likert scale was adopted to correct the study tools, by giving criteria of numerical weight (strongly satisfied = 5, satisfied = 4, satisfactory = 3, unsatisfactory = 2, highly unsatisfactory = 1) and it is numerically represented (5, 4, 3, 2, 1) on ranking, and the following scale has been adopted for the purposes of analyzing the results [18]:

- From 1.00 - 2.33 is low
- From 2.34 - 3.67 is medium
- From 3.68 - 5.00 is large

The scale was calculated using the following equation: [19].

$$\frac{\text{Scale upper limit (5)} - \text{Scale lower limit (1)}}{\text{Number of categories required (3)}} \dots \dots \dots (1)$$

$$= (5-1) / 3 = (1.33)$$

Then adding the answer (1.33) to the end of each category.

Normal Distribution Test: The normal distribution test for the data was carried out based on the value of the skewness coefficient, to test whether the data used in the analysis followed a normal distribution or not, so the results were as shown in figure (5).

These dimensions/variables are hypothesized to affect the dependent variable project performance. It is noted from figure (5), that the value of the Skewness coefficient for all paragraphs of the questionnaire was less than one, and this means that the data follows a normal distribution.

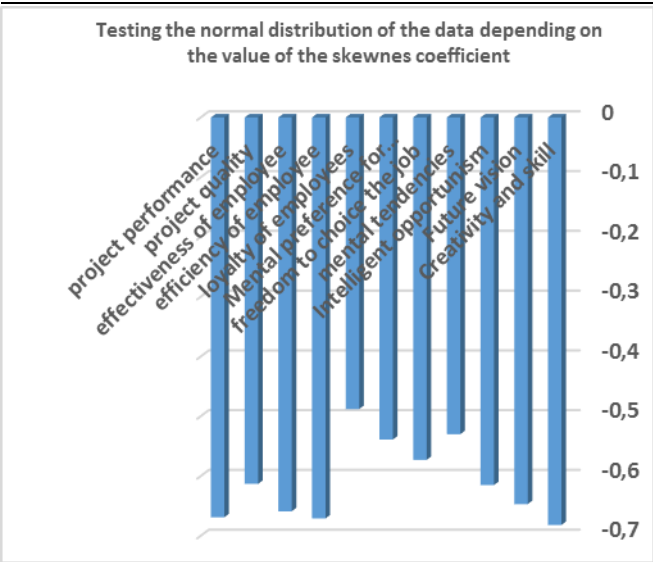


Figure (5) Normal distribution of the data

Testing the Suitability of the Study: The linear correlation test was used to ensure that there is no high correlation between the independent variables, based on the variance inflation coefficient (VIF) test, and the tolerance test for each of the independent variables, as the independent variables of the model must be independent among themselves. Also to ensure that purpose, will use this test, which is one of the ways to get rid of the problem of multi-collinearity, knowing that the variance inflation coefficient must not exceed the value (10). While the value of the allowable variance test must be greater than (0.05), and by calculating the previous coefficients for each independent variable by using the (VIF) equation as shown in formula (4), the obtained results are listed in the following figure (6). [20, 21].

$$VIF = \frac{1}{1 - R_i^2} = \frac{1}{\text{Tolerance}} \dots\dots\dots(4)$$

Where: R_i^2 : represents the unadjusted coefficient of determination for regressing the independent variable (i) on the remaining variables

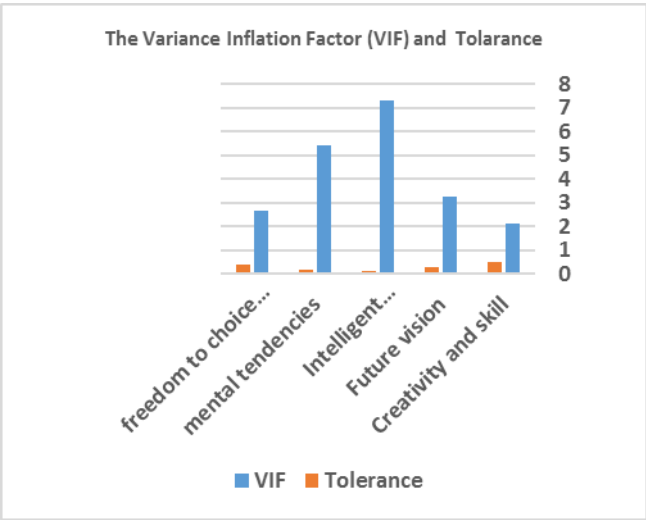


Figure (6): The Variance Inflation Factor (VIF) and tolerance

Analysis of Results: The data were analyzed using a statistical package for the social science version 21 (SPSS-V21) as follows: Respondents answers to the brain thinking parameters: “What is the level of brain thinking (i.e., creativity and skills, future vision, intelligent opportunism, mental tendencies, and freedom to choice the job)” at projects as perceived by employees? The arithmetic means and standard deviations of the study sample's responses to

the mental preference for choosing the type of work were extracted. As shown in figure (7), the dimensions/variables are hypothesized to affect the dependent variable project performance.

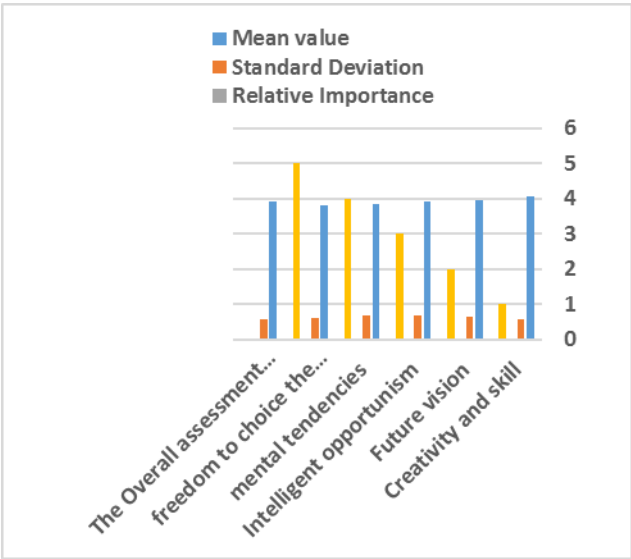


Figure (7) The Arithmetic Means and Standard deviations related of the mental Preference for choosing the type of work

Figure (7) shows that the relative importance of the mental preference for choosing the type of work was high, as the general arithmetic mean in overall assessment question was (3.93) and with a standard deviation of (0.563), where creativity and skill came in the first place with an arithmetic mean of (4.07) and a standard deviation of (0.590) and of high relative importance. Whereas freedom of occupational choice came in the last rank, with an arithmetic mean of (3.81) and a standard deviation of (0.627), with high relative importance. The participants ranked creativity and skills as being the most important influencing factors, which was followed by future vision, Intelligent opportunism, mental tendencies, and then the freedom to choose the job. This result is attributed to successful innovation depends on employees’ creativity, so firms must innovate to survive and succeed. Creativity is essential for organizations so it affects employee efficiency and job satisfaction, as the successful implementation of creativity-based training programs contributes to the development of creative skills. The arithmetic means, standard deviations, and relative importance of each dimension of mental preference for choosing the type of work were extracted separately and the results are shown in figures (8 to 11).

First parameter: “Creativity and skill”

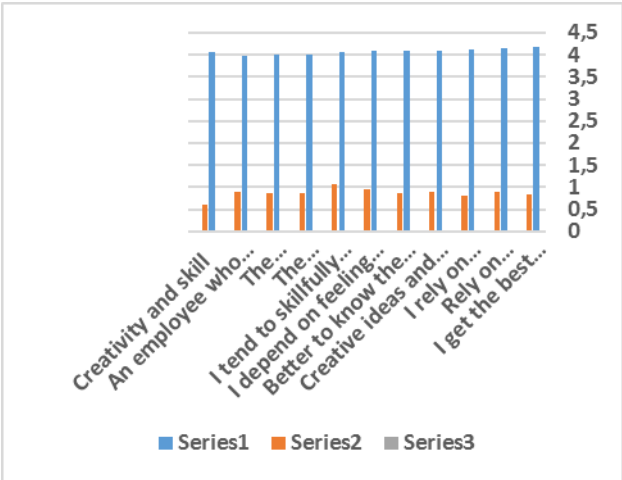


Figure (8): The arithmetic means and standard deviations related to creativity and skill

Figure (8) shows that the relative importance of creativity and skill was high, as the general arithmetic mean was (4.07) with a standard deviation of (0.590), this indicates there is a great agreement about the importance of this parameter in brain thinking of the employees and the paragraph that states “I get the best ideas and creative plans when I am not busy with something” came ranked first among the paragraphs, with an arithmetic mean of (4.17) and a standard deviation of (0.832), with high relative importance. Whereas, the paragraph stating “the employee who thinks and works creatively helps achieve the strategic goals and increase their competitiveness” ranked last among the paragraphs, with an arithmetic mean (3.96) and a standard deviation (0.901), and a high relative importance. The majority of respondents agreed that the best creative ideas and plans can be generated when employees are not busy. These results are attributable to the fact that silence is useful for the creative process as it enables one to generate ideas without restrictions, guidelines, or interruptions.

Second parameter: “The future vision”

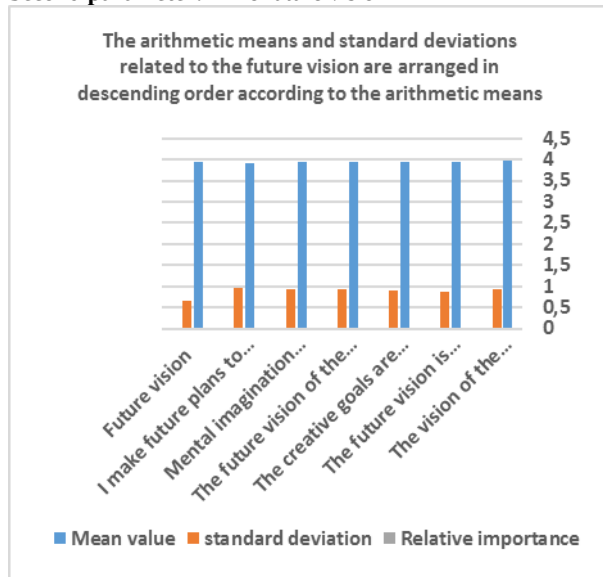


Figure (9) The arithmetic means and standard deviations related to the future vision

Figure (9) shows that the relative importance of the future vision was high, as the general arithmetic mean was (3.94) with a standard deviation of (0.647), and the paragraph stating “the vision of the institution/company reflects the future direction of the program it offers” It ranked first among the paragraphs, with an arithmetic mean of (3.96), a standard deviation of (0.924), and a high relative importance. This indicates there is a great agreement about the importance of this parameter in brain thinking of the employees. Whereas, the paragraph stating “I make future plans to overcome risks and mistakes at work” ranked last among the paragraphs, with an arithmetic mean (3.92) and a standard deviation (0.953), with a high relative importance. This result is attributable to the fact that prosperous businesses have the propensity to have visionary goals like hiring competent, knowledgeable, confident, experienced, loyal, and skilled employees who will devote their full efforts, attitude, and determination to giving the company a competitive edge in the market.

Third parameter: “Loyalty of employees”

It is clear from figure (10) that the relative importance of loyalty of employees was high, as the general arithmetic mean was (3.79) with a standard deviation of (0.640), and the paragraph stating “I maintain the property of the institution/company honestly and faithfully” came in the first place. Between paragraphs with an arithmetic mean of (3.81), a standard deviation of (0.908), and a high relative importance. Whereas, the paragraph stating “I feel that I am part of the administrative system of the institution/company” ranked last among the paragraphs, with an arithmetic mean (3.77) and a standard deviation (0.957), with a high relative importance. The majority of participants choose the paragraph that states “I

maintain the property of the institution/company honestly and faithfully”. These results are in line with a study conducted by [22], the researchers said that integrity has a significant and positive impact on employee performance, and a high-integrity employee will perform better performance.

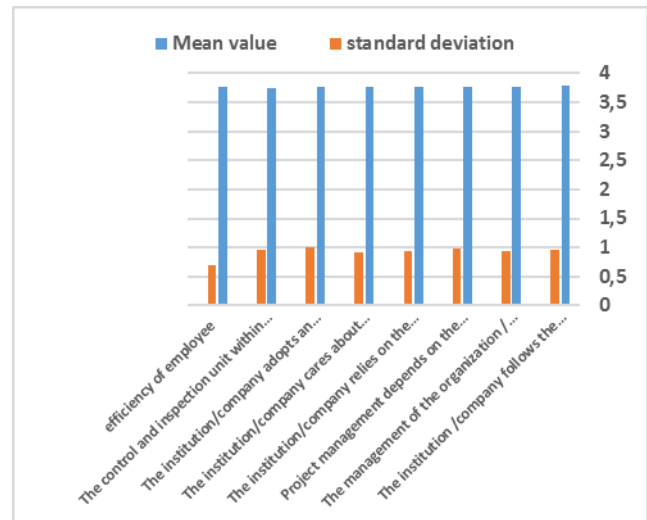


Figure (10). The arithmetic means and standard deviations related to employee loyalty

Forth parameter: “Efficiency of Employees”

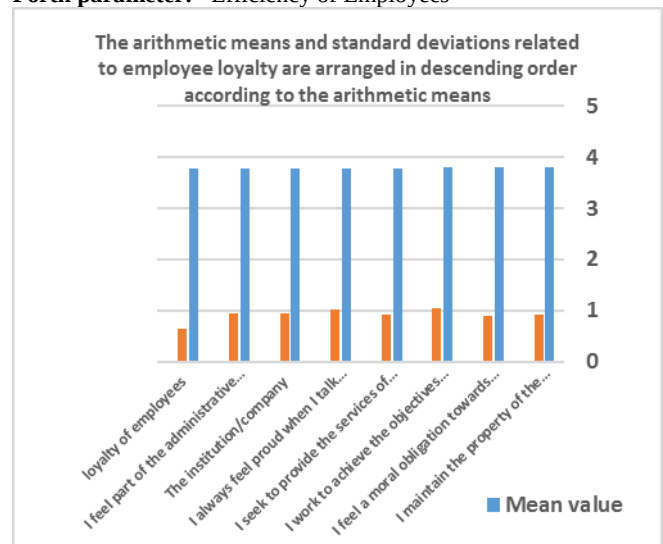


Figure (11): The arithmetic means and standard deviations related to employee efficiency

While figure (11) shows that the relative importance of the employee's efficiency was high, as the general arithmetic mean was (3.77) with a standard deviation of (0.694), and the paragraph stating “the institution/company follows the approach of maximizing value and reducing cost during the production process in Projects” ranked first among the paragraphs, with an arithmetic mean of (3.78) and a standard deviation of (0.948), with a high relative importance. Whereas, the paragraph stating “The control and inspection unit within the institution/company enjoys high efficiency” ranked last among the paragraphs, with an arithmetic mean (3.75) and a standard deviation (0.952) with a high relative importance. These results are attributable to the fact that an organization's performance can be enhanced by granting its employees autonomy in the workplace, freedom of choice regarding how they complete tasks, and the opportunity to contribute their ideas. Effective and efficient employees are likely to perform less well if they are not adequately motivated. So, employees need to work in an environment where creativity is unrestricted.

Hypothesis Tests: ‘First sub-hypothesis’ to verify the validity of the first sub-hypothesis, a simple linear regression was used for the impact of the creativity and skills dimension on the performance of projects in Jordan, and the results as shown in figure (12).



Figure (12): Results of a simple linear regression of the impact of creativity and skills

Figure (12) shows that there is a statistically significant effect of the dimension of creativity and skills in the performance of projects in Jordan, where the correlation coefficient ($R = 0.605$), which indicates the existence of a statistically significant correlation between the independent variable (creativity and skills) and the dependent variable (project performance). The value of the coefficient of determination ($R^2 = 0.366$), which indicates that (creativity and skills) explained 36.6% of the variation in (project performance), while the remainder is due to other variables that were not included in the model, and the value ($F = 287.681$) at a confidence level equal to ($\text{sig} = 0.00$), and this confirms the significance of the regression at a significance level of ($\alpha < 0.05$). It appears from the figure of coefficients that the values of (B) in the dimension of creativity and skills amounted to (0.613), and that the value of (t) was (16.961) with a statistical significance of (0.000), which indicates that the effect of this dimension is significant. These results indicate the rejection of the first sub-hypothesis and the acceptance of the alternative hypothesis, which states: "There is a statistically significant effect at the significance level ($\alpha \leq 0.05$) for creativity and skills dimension in the performance of projects in Jordan."

Second sub-hypothesis: "There is no statistically significant effect at the significance level ($\alpha \leq 0.05$) of the future vision dimension on the performance of projects in Jordan".

To verify the validity of the second sub-hypothesis, a simple linear regression was used to effect the future vision dimension on the performance of projects in Jordan, and the results as shown in figure (13):

The results indicate that there is a statistically significant effect of the future vision dimension on the performance of projects in Jordan, where the correlation coefficient ($R = 0.739$) indicates that there is a statistically significant correlation between the independent variable (future vision) and the dependent variable (project performance). The value of the determination coefficient ($R^2 = 0.546$), which indicates that (future vision) explained 54.6% of the variation in (project performance), while the remainder is due to other variables that were not included in the model, and the value ($F = 599.778$) at a confidence level equal to ($\text{sig} = 0.00$), and this confirms the significance of the regression at the level of significance ($\alpha \leq 0.05$). Third sub-hypothesis: To verify the validity of the third sub-hypothesis, a simple linear regression was used for the effect of the intelligent opportunism dimension on the performance of projects in Jordan, and the results were as follows:

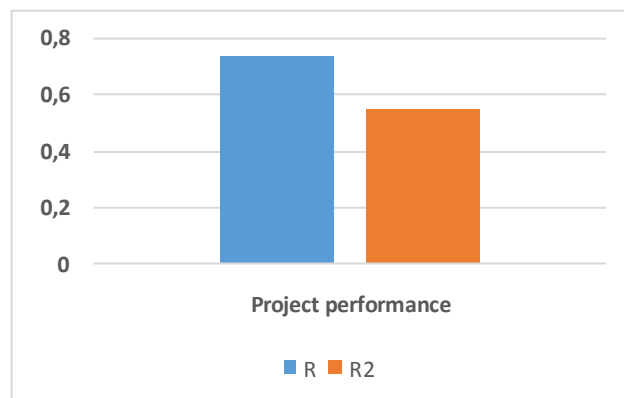


Figure (13): The results of a simple linear regression of the impact of the future vision

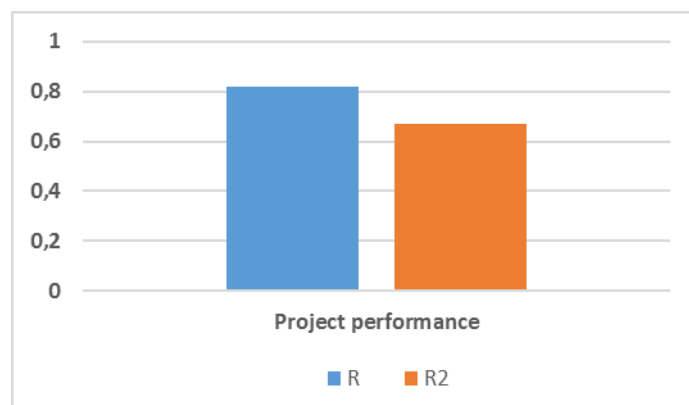


Figure (14): The results of a simple linear regression of the impact of the intelligent opportunism

The results indicate that there is a statistically significant effect of the intelligent opportunism dimension on the performance of projects in Jordan, where the correlation coefficient ($R=0.817$) indicates that there is a statistically significant correlation between the independent variable (intelligent opportunism) and the dependent variable (project performance). The value of the coefficient of determination ($R^2 = 0.668$), which indicates that (intelligent opportunism) explained 66.8% of the variation in (project performance), while the remainder is due to other variables that were not included in the model, and the value ($F = 1002.993$) at a confidence level equal to ($\text{sig} = 0.00$), and this confirms the significance of the regression at a significance level of ($\alpha \leq 0.05$).

Summary of The Main Results: The summary of main results of this work (The dependent variables) can be shown in figure (15).

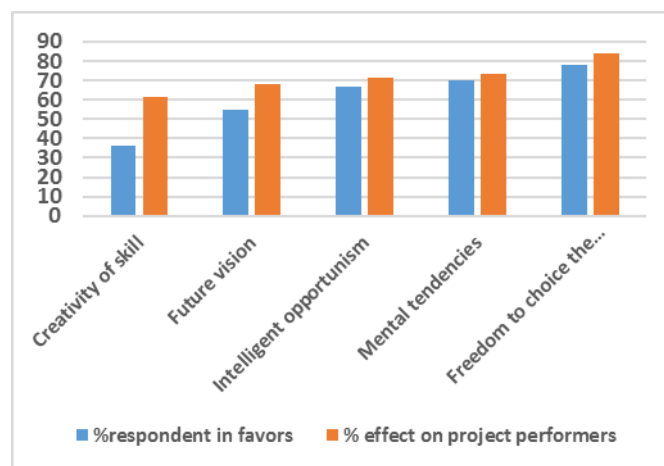


Figure (15): Summary of the main results (dependent variables) of this study

Conclusions:

From the above of the results of this discussion of this work, the following main points can be concluded.

- Creativity and skill parameters of the brain thinking preference has an effect of 61.3% on engineering project performance with 36.6% of respondent's agreement and with (VIF) of 2.097% indicating low correlation between variables.
- Future vision parameter of the brain thinking preference has an effect of 68.3%, with 54.6% of respondent's agreement, indicating good impact.
- Intelligent opportunism parameter of the brain thinking preference has an effect of 71.4%, with 66.8% of respondent's agreement, indicating a high impact.
- Mental tendencies parameter of the brain thinking preference has an effect of 73.1%, with 70.3% of respondent's agreement, indicating a high impact
- The freedom to choose a job affected has an effect on Project performance by 84.3%, and 78.2% of the respondents agreed that this measure of their mental thinking raises their project performance.
- The overall effects of the five criteria's according to the multiple regression test of these criteria (creativity and skill, future vision, smart opportunism, mental inclinations, freedom to choose a job) are 84.4% on the engineering project's performance (efficiency of employees, employees effectiveness, loyalty of employees, project quality).

References:

- [1] Vaihtovirta, N. (2011). Measurement of Business Performance in Strategy Research: A Comparison of Approaches. *The Academy of Management Review*, Vol. 11.
- [2] Morris J., Small, C., & Steinning, K. (2019). The role of representation in teaching and learning critical thinking. *Educational Review*, 54 (1).
- [3] Aizikovitch-Udi, E., & Amit, M. (2011). Developing the skills of critical and creative thinking by probability teaching. *Procedia Social and Behavioral Sciences*, 15, 1087–1091 Doi: 10.1016/j.sbspro.
- [4] Balcom, Dz., & Heinonen, E. (2019). Holistic approach in reorienting teacher education towards the aim of sustainable education: the case study from the regional university in Latvia, *Proceedings of the 5th World Conference on Educational Sciences*, 116.
- [5] Baumeister, R. F. (2008). Free will in scientific psychology. *Perspectives on Psychological Science*, 3, 14-19.
- [6] Brown, V. A. & Harris, J. A. (2016). *Tackling wicked problems: Through the transdisciplinary imagination*. UK: Earthscan
- [9] Herrmann, R. K. (2009). *Image theory and strategic interaction in international relations*. In D. O. Sears, L. Huddy, & R. Jervis (Eds.), *Oxford handbook of political psychology* (pp. 285–314), Oxford.
- [10] Herrmann I (2018) Boundaries and otherness in science-fiction: how we cannot escape the human condition. *Text Matt* 8(8):212–226
- [11] Herrmann, N. (2000). The Theory Behind the HBDI and Whole Brain Technology, The HBDI Accreditation Process. *Herrmann International* [online].
- [12] Keelin, R.J., Cavusgil, S.T., Arnold, Y. (2020). Learning Orientation, Firm Innovation Capability, and Firm Performance. *Industrial Marketing Management*, Vol. 31, No. 6.
- [13] Dufy, G. and Gallagher, T. (2017). Shared education in contested spaces: How collaborative networks improve communities and schools. *Journal of Educational Change*, 18(1).
- [14] Guilford, A. and Torrance S. (2020). Harnessing Knowledge for Innovation: An Integrated Management Framework. *Journal of Knowledge Management*, Vol. 9, No. 4.
- [15] Gamil, E. F., & Rahman, A. (2019). Building a culture that encourages strategic thinking. *Journal of Leadership & Organizational Studies*, 17(2).
- [16] Joubert, s & Christine Swart (2019). Developing a team performance framework utilising Neethling's thinking style preference instrument, ISSN: (Online) 2071-078X, (Print) 1683-7584
- [17] Sternberg, T. F., & Lubart, L. E. (2018). Personal philosophy and personnel achievement: Belief in free will predicts better job performance. *Social Psychological and Personality Science*, 1, 43-50
- [18] Gamil, E. F., & Rahman, A. (2019). Building a culture that encourages strategic thinking. *Journal of Leadership & Organizational Studies*, 17(2), 119–128
- [19] De Bono, E. (2019). *Serious Creativity: Using the Power of Internal Thinking to Create New Ideas*. New York: Harper Collins.
- [20] Renaud, R. D., & Murray, H. G. (2018). A comparison of a subject-specific and a general measure of critical thinking. *Thinking Skills and Creativity*, 3, 85–93.
- [21] Laumann D., & Etz, A. J. (2017). *Too true to be bad: When sets of studies with significant and non-significant findings are probably true*. *Social Psychological and Personality Science*.
- [22] Runco, M. A. (2017). *Conclusions concerning problem finding, problem solving, and creativity*. (Ed. Runco, M. A. in *Problem finding, problem solving, and creativity*). Norwood, New Jersey: Ablex Publishing
- [23] Shriven, E., & Shojaie, E. (2011). Improving high-level thinking skills by development of learning PBL approach on the learning mathematics for Senior High School students. *International Education Studies*, 10(8), 12–20
- [24] Kaufman H. (2013). Is creativity unidimensional or multidimensional? Analyses of the Torrance Tests of Creative Thinking. *Creativity Research Journal*, 18(3)
- [25] Feldhusen, J. F., & Goh, B. E. (1995). Assessing and accessing creativity: An integrative review of theory, research, and development. Doi: 10.1207/s15326934crj0803_3.
- [26] Alshmaeleh S. Altarawneh S. and Al-Juboori S. "Employees Brain Dominance Thinking Style of Arab Potash Company". *Engineering and Technology (IJGET)*, ISSN(P): 2278-9928; ISSN(E): 2278-9936, Vol. 8, Issue 1, Dec- Jan 2019; 19-38

A simple e-voting system ensuring identification, privacy and verifiability

Mariam Haroutunian¹, Karen Mastoyan², Arsen Margaryan³
 Institute for Informatics and Automation Problems of NAS RA^{1,2}, Yerevan, Armenia
 MoodMap AI³, Yerevan, Armenia
 armar@sci.am¹, kmastoyan@yandex.com², margaryanarsen83@gmail.com³

Abstract: *Electronic voting (e-voting) offers many advantages compared to traditional types of voting. Today there are many e-voting schemes and systems, which allow their use for various purposes. Because of its fundamental importance for democratic societies, e-voting is subject to high legal standards, in particular security requirements for the voting method. It is hard to create a system or a protocol fulfilling all security requirements. Designing effective e-voting systems is challenging because these aspects often conflict with each other. For example, one of such challenges is trying to ensure identification and keep votes private while still being able to verify them. There are quite a few cryptographic schemes which fulfill wide requirements for electronic elections, but they use sophisticated cryptographic tools that make them hard to implement and require expertise in various fields. In this paper we suggest a new approach that ensures this requirements without complex cryptographic methods, using only face recognition, properties of entropy, hash functions and steganography.*

KEYWORDS: E-VOTING, PRIVACY, VERIFIABILITY, HASH FUNCTION, ENTROPY

1. Introduction

E-voting refers to the process of casting and counting votes using electronic means, typically through computers or the internet. E-voting systems can vary widely in their design and implementation, ranging from simple online voting platforms to more complex electronic systems. One of the challenging types is the internet voting (i-voting), when voters can cast their ballots online from any location.

Advocates of e-voting often highlight several potential benefits:

- **Accessibility:** E-voting can make the voting process more accessible to individuals with disabilities by providing options such as screen readers, enlarged text, or alternative input methods.
- **Speed and Efficiency:** Electronic systems can streamline the voting process, potentially reducing long lines at polling stations and speeding up the counting of ballots.
- **Convenience:** E-voting allows voters to cast their ballots remotely, eliminating the need to travel to a physical polling location.
- **Cost Savings:** Potential long-term savings on printing, staffing and logistics.
- **Accuracy:** Electronic systems can reduce the risk of errors associated with manual vote counting, such as miscounting or illegible ballots.
- **Increased Participation:** Easier access may encourage higher voter turnout.

However, there are also concerns and challenges associated with e-voting:

- **Security Risks:** E-voting systems are vulnerable to various security threats, including hacking, malware, and tampering. Ensuring the security and integrity of electronic voting systems is a significant challenge.
- **Trust and Transparency:** Transparency in the voting process is essential for maintaining public trust in election outcomes. E-voting systems must be designed to provide transparency at every stage, from ballot casting to vote tabulation.
- **Privacy:** Protecting the privacy of voters is crucial in any voting system. E-voting systems must ensure that voter information remains confidential and cannot be linked back to individual ballots.
- **Digital Divide:** Not everyone has equal access to technology or the internet, which can create disparities in access to e-voting systems. Ensuring equitable access for all voters is a significant concern. Older or less tech-savvy populations may find e-voting challenging.
- **Accountability:** In the event of technical issues or discrepancies, it may be challenging to audit or verify electronic voting systems effectively. Establishing accountability mechanisms is essential for addressing these concerns.

Estonia has been a pioneer in i-voting, allowing citizens to vote online in national elections since 2005. A huge number of publications are devoted to the problems related to e-voting systems. In our previous research [1-3] we studied existing e-voting systems, literature on this topic, investigated the security requirements. We can state that from one side it is necessary to use different cryptographic methods, but at the same time the solution must be user-friendly to ensure maximum efficiency. There are quite a few cryptographic schemes which fulfill wide requirements for electronic elections. Their only disadvantage is inconvenience: they use sophisticated cryptographic tools that make them hard to implement and require expertise in various fields.

Thus, while e-voting holds potential benefits for improving the efficiency and accessibility of the voting process, it also presents significant challenges that must be addressed to ensure the security, integrity, and fairness of elections.

Requirements

The e-voting system must be **reliable** to prevent election fraud (insider attacks) or attacks on the system from outside. When implementing a technical solution, it is necessary to take into account all types of attacks. **Interoperability** of all technical components and services used must be ensured, compatibility and correct functionality of the system must be assured, therefore suitable open electronic data interchange standards should be used. The system must be operational under any circumstances, even if some problems arise (for example, with the network), equipment failures, hacker attacks, etc.). The reliability of the system must be tested, evaluated and improved constantly. The i-voting system must be **uninterruptedly accessible** to voters. Its interface should be easy to use and all voters should be able to access it equally, regardless of age, education or physical condition. Ballots should be written as clearly as possible so that voters can minimize the possibility of votes being miscounted due to any misleading aspect of the ballot design. Another requirement is the **scalability**, ensuring that the system can handle a large number of voters and high traffic, especially during peak voting periods. At last we would like to mention the **simplicity** requirement ensuring that the system is easy to understand and use to prevent user errors.

The main **security characteristics** of e-voting system can be formulated in the following list:

- **Authentication or Eligibility:** Ensuring that only eligible voters can cast a vote.
- **Integrity or Completeness:** Guaranteeing that votes are recorded as cast and counted as recorded.
- **Stability:** No ballot is subject to change after being sent to ballot storage.
- **Privacy:** Ensures that there is no link between a voter and his vote, in other words, protects any information about the voter.

- **Anonymity:** Make ballots indistinguishable from one another, not to find who is the voter.
- **Confidentiality:** Protecting the contents of the vote from being disclosed to unauthorized parties.
- **Verifiability:** Enabling voters to verify that their vote was counted correctly without compromising their anonymity.
- **Unreusability or uniqueness:** voter can vote only once, does not permit any voter to cast more than once.

Designing effective voting systems is challenging because these aspects often conflict with each other. One of such challenges is trying to ensure identification and keep votes private while still being able to verify them. In this paper we suggest a new approach that ensures this requirements without complex cryptographic methods, using only face recognition, properties of entropy, hash functions and steganography.

Our aim is to develop a simple and at the same time secure online voting system in Armenia, suitable for a variety of choices and adaptable to different cases.

2. Existing cryptographic solutions

As mentioned above, an e-voting system is vulnerable to many threats; fraud, compromised computers, attackers compromising system components or other attacks launched over the network. The need for better mechanisms of authentication, anonymity, integrity, confidentiality, secrecy, etc. is crucial. With the development of i-voting systems, the need for cryptographic means both for providing security against these threats and for verification purposes is increasing.

Cryptography is one of the fundamental areas involved in ensuring the security of internet voting. In addition for security other processes and implementations are also important.

One of the main cryptographic methods is encryption. **Symmetric key encryption** is critical to many aspects of secure online voting systems. It is involved in ensuring the security of voting data both at storage and in transit. **Public-key cryptosystems** like RSA can provide confidentiality, authentication and key exchange.

To protect the anonymity of voting and ensure its security, Estonian e-voting uses a **double envelope system**. Once a user accesses the voting application on his computer, he votes and a random number is added to the data. The vote and random number are then encrypted with an election-specific public key. This forms an inner envelope which is used to keep the vote anonymous. Random numbers must be generated by **cryptographically secure pseudo-random number generators**, that can produce a high degree of entropy to prevent patterns and correlations from appearing in the numbers. Such generators are used to produce random values for key generation, salting, nonces, and vote shuffling, increasing the randomness and unpredictability of the voting process, thereby preventing attacks and preserving election anonymity.

Homomorphic encryption, an emerging cryptographic technique that **allows computations to be performed on encrypted data**, has the potential to solve some problems in online voting schemes. One of applications is the usage for tallying encrypted votes without compromising individual voter anonymity. It is still in its early stages of development but holds promise for future online voting schemes.

Cryptographic hash functions, that map variable-sized data volumes into fixed-sized outputs by a one-way mathematical functions, can be used as checksum algorithms. These have a variety of uses, such as verifying the integrity of software. Combining public-key encryption and hash functions, **digital signatures** authenticate the source and integrity of data, which is critical in e-voting systems. This process involves signing data with a private key and verifying it with a corresponding public key.

A **mix network** is a technique based on public key cryptography which includes mixing components between the corresponding parties in the cryptosystem, that process each sent data before it is delivered. A mix net takes encrypted data as input,

shuffles and reorganizes the data and decrypts it. The purpose of a mix network is to hide the correspondences between the items in its input and those in its output. Mix net can be included as a component in the voting process to ensure anonymity, privacy, verifiability. This step can also help to prevent intimidation.

A **zero knowledge proof** is a method to prove knowledge of some fact without revealing this knowledge. Such proofs involve interactions between two communicating parties, the prover and the verifier. The prover demonstrates the possession of some information (authentication information) to the verifier without disclosing it. Today's internet applications of zero knowledge proof are very wide including e-commerce, e-voting, access authorization and entity authentication.

Several remote e-voting solutions based on **blockchain technology** have been proposed. The survey [4] provides a comprehensive overview of the blockchain - based e-voting systems currently being implemented by various countries and companies and proposed for academic research. Indeed, in the scientific literature, more and more research works propose e-voting applications based on blockchain. Nevertheless, only some of the proposed solutions have been implemented in real life and none of them have been tested on a large scale. Therefore, it is very difficult to conclude that blockchain is a completely secure alternative for national election at present. Although the principles on which blockchain is based are secure, e-voting applications are still vulnerable to several attacks. This makes it very challenging to guarantee the integrity of an election, which is problematic given the stakes of such an application. The National Academies of Sciences report [5] states that "blockchain technology does little to solve the fundamental security issues of elections, and indeed, blockchains introduce additional security vulnerabilities". Blockchain technology is designed to keep information secure once it is received. It cannot defend against the multitude of threats to that information before it is entered in the blockchain, and voters cannot verify their votes are entered into the blockchain correctly without compromising ballot secrecy. Recording ballots on a blockchain also risks ballot secrecy if encryption keys are not properly protected or software errors allow decryption of individual ballots.

3. Preliminary information

Entropy

Our approach uses the basic notion in information theory, that is Shannon's entropy, which is the measure of uncertainty [6]. For given random variable X with values from the set χ entropy is calculated as

$$H(X) = - \sum_{x \in \chi} p(x) \log p(x)$$

with log of base 2. Here we omit the properties of entropy. We are interested in using this measure for preserving privacy in face images.

The information theory approach has proved to be effective in solving many Computer Vision and Pattern Recognition problems [7]. One such approach was suggested in [8].

The idea that we suggest here is based on the fact that permuting pixels of the image does not change the entropy of the distorted image (see Fig. 1).

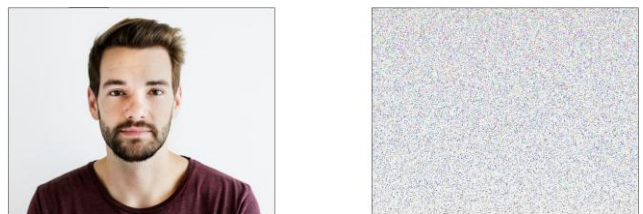


Fig. 1. Source image and the distorted image by permutation of same pixels have the same entropy $H = 5.456204603175102$

Face recognition

In the rapidly advancing world of facial recognition technology, finding the delicate balance between its benefits and protecting people's rights to privacy is of utmost importance and is a challenging problem. Facial recognition relies on capturing and analysing biometric data, which is highly sensitive and unique to individuals. Unlike password or identification numbers, biometric data cannot be changed or reset.

In our system we use face recognition for identification of the voters. The important requirement is to choose a face recognition model which is accurate and can handle a large volume of recognition requests efficiently to facilitate smooth voting process.

Facial recognition technology often involves integrating with various platforms and databases, which can create a complex web of interconnected personal data. If not properly regulated, sharing facial recognition data across platforms can lead to significant privacy risks.

Our solution, described in the next section, addresses all these challenges.

Election structure

The online election process starts with so called *setting phase*, where the staff creates ballot, election parameters are formed and published. These parameters include encryption keys, the number of candidates, list of voters, etc. The election is announced with sending login credentials to users.

The next is the *ballot filling phase*, during which the voter login to user interface and prepare his/her vote. The result should be an e-ballot with voter's personal info embedded in it.

In the *ballot registration phase* the ballot with embedded info and voter's personal number are sent to public ballot storage. The voter's info must be removed from the ballot and as a result a ballot without voter info is collected, this is called *ballot anonymity phase*.

At last in the *counting phase* the committee uses a personal key to decode all the ballots and count the voices.

In the next section we describe in detail processes and our suggestions to ensure the main security characteristics.

4. Our solution

Designing an electronic voting system involves several key components to ensure the main security characteristics. Let us consider all steps of e-voting with our suggestions.

Process 1 is the *Login* step of the staff, which requires secure managing the administrative access to the voting system.

Process 2 is the *Creation of the Ballot* which needs designing a ballot that is easy to understand and use. It must enable handling a variable number of candidates and ensuring the system scales accordingly.

Process 3 implements the *transmission of Login Credentials* to voters. Here the challenges are maintaining an accurate, secure, and up-to-date list of eligible voters and securely distributing login credentials to voters.

In our e-voting system we suppose that the list of voters include their ID cards with face images. The staff send an e-mail to each voter with the voting information and the link to the interface of the e-voting system. The authentication of the voter is realized by face recognition system, that compares a person's face with ID card face image. The system should check the *eligibility* of the voter. The photo image of the user is then immediately distorted by a random permutation (Fig. 1). The entropy H of this image is sent to the voter's e-mail.

Once authenticated, voter can cast his vote electronically (**process 4**). After that the hash value h_1 of the filled ballot with distorted image (**process 5**) is sent to the voter's e-mail.

In the **process 6** the system records each vote with distorted face image to the database. At the same time the hash value h_2 of the existing list is sent to the voter's e-mail.

After the voting period ends, the system should tabulate the votes and generate results guaranteeing *completeness* and *integrity* (**process 7**). The final stage **process 8** must record results

accurately. For avoiding any changes on this step we suggest to transmit the results in a hidden way using steganographic methods.

Discussion on satisfaction of requirements

Now let us discuss the security, privacy and verifiability aspects of suggested system. Which are the roles of distorted images, image entropy and hash values?

The entropy of the voter's face image or the shuffled image will be the **identification number** of the voter. It can be considered as a unique number, because from one side it can not be repeated even for the same person (if it would be allowed), but also because the entropy is a 16 digits number (see Fig. 1) and hence there can be about 10^{16} various values of entropy, that is much greater than the possible number of voters.

As we told, the entropy or the identification number in the system is invariant for any permutation of the image pixels. We claim that the distorted image guarantees the **privacy** of the voter, because it is practically impossible to find the source image from the shuffled one. To prove that, consider an example of an image with only $P = 5 \times 5 = 25$ pixels, then the number of possible permutations equals $|P|! = 25!$, which is a number of order 10^{24} . We know that 1 year equals 31 536 000 seconds that is greater than 3×10^7 . If the supercomputer can check one permutation in 10^{-10} seconds, then it will need about 10^7 years to find the source image. But the real image can have much more pixels. The considered on the Figure 1 image has about $600 \times 500 = 300000$ pixels, hence the number of possible permutations is an astronomically large number. When pixels are shuffled randomly, the process is typically non-deterministic and does not retain any history or record of the original positions. Each pixel's position in the shuffled image is independent of its original position unless additional information (such as a mapping function or shuffle key) is stored. This makes it computationally infeasible to reverse the specific shuffle applied to the pixels without knowing the exact permutation. This is without taking into account also that each pixel has 3 components, that may be also permuted. Hence, we proved that it is impossible to revert a randomly shuffled image to its original state without additional data or a record of the shuffle operations.

At the end of the election each voter will have possibility to **verify** the election by sending a query with identification number (entropy of the image). The system will check the existence of such number, will compute the hash value h_1 of corresponding vote and the hash value h_2 of the previous list and send back to the voter. If all accepted hash values coincide with previously received by e-mail values, it means that

- the voters ballot is not ignored or erased (integrity),
- the vote is recorded correctly (stability),
- the list of votes is not changed (completeness).

Otherwise, if the values of h_1 or h_2 are changed, the voter may record an election fraud.

5. Conclusion

The suggested e-voting system is simple in use and at the same time secure, suitable for a variety of choices and adaptable to different cases. It does not need trusted persons, does not need any exchange of secret keys. It is very simple for users, does not need knowledge of cryptography, and hence can be used by wide range of individuals without preliminary trainings. Privacy and verifiability, as well as the fact that final results are not encrypted and are available to all voters add trust to this system.

6. References

- [1] A. Avetisyan, M. Haroutunian and P. Hakobyan, "Discussion on steganographic methods from the perspective of e-voting implementation", Proc. of 14th International Conf. Computer science and information technologies CSIT 23, pp. 214-217, 2023. https://doi.org/10.51408/csit2023_51
- [2] M. Haroutunian, A. Margaryan and K. Mastoyan, "Secure online voting scheme using steganography", CSIT workshop on Data Analytics and Mathematical Modeling, Tbilisi, 2024.

- [3] M. Haroutunian, P. Hakobyan and A. Avetisyan, "Two-stage optimal hypotheses testing for a model of stegosystem with an active adversary", JUCS - Journal of Universal Computer Science, vol. 29, no. 11, pp. 1254-1273, 2023. <https://doi.org/10.3897/jucs.112913>
- [4] M. V. Vladucu, Z. Dong, J. Medina and R. Rojas-Cessa, "E-voting meets blockchain: A survey," in *IEEE Access*, vol. 11, pp. 23293-23308, 2023. doi: 10.1109/ACCESS.2023.3253682
- [5] A Consensus Study Report of National Academies of Sciences, Engineering, and Medicine, "Securing the vote: protecting American democracy", Washington, DC: The National Academies Press, 2018. <https://doi.org/10.17226/25120>
- [6] T. M. Cover and J. A. Thomas, Elements of information theory. 2nd ed. A Wiley-Interscience Publication, USA 2006.
- [7] F. Escolano, P. Suau and B. Bonev, Information theory in computer vision and pattern recognition, Springer Verlag London Limited 2009.
- [8] M. Haroutunian, D. Asatryan and K. Mastoyan, "Analyzing the quality of distorted images by the Normalized Mutual Information measure", *Mathematical Problems of Computer Science*, vol. 61, 2024.

Acknowledgment

The work was supported by the Science Committee of RA, in the frames of the research 21T-1B151.

Contrastive Analysis of Using Japanese Fillers and the Relationship between Fillers and Politeness: Japanese, Chinese, Korean and Vietnamese

Chie Yamane-Yoshinaga^{1,*},
Sanyo Gakuen Junior College, Okayama, Japan¹
chie@sguc.ac.jp

Abstract: Japanese has various kinds of fillers. Their functions include not only hesitation but also marking the softening of utterances and indicating breaks between utterances. This study analyzes native Japanese, Chinese, Korean and Vietnamese speaker brain's cerebral blood flow related to fillers by using fNIRS (functional near-infrared spectroscopy). Additionally, we seek to explore the relationship between fillers and politeness of utterances by analyzing the sounds of fillers. We conclude fillers are connected to the increase of cerebral blood flow and filler's kinds, pitch, length and strength are influenced to politeness.

Keywords: FILLERS, BRAIN, OPTICAL TOPOGRAPHY, fNIRS, CEREBRAL BLOOD FLOW, BROCA'S AREA, POLITENESS

1. Introduction

Japanese has various kinds of fillers [3]. This purpose of this study is to seek the relationship between the brain activity and fillers using fNIRS in both native Japanese speakers and foreign students (they can speak Japanese) living in Japan whose mother tongues are Chinese, Korean and Vietnamese [5]. Additionally, we seek to explore the relationship between fillers and politeness by analyzing their position in the utterance (beginning, end), their kinds (ano, ee, eeto, maa etc.) and their sounds (pitch, length, strength) [4].

2. Methods

2.1. Brain activity experiment using fNIRS [2]

Each participant (2 students for each group: Japanese, Chinese, Korean, Vietnamese) says 1 minute's monologue about happy story in his or her life and 1 minute's monologue about sad story in his or her life.

2.2. Listening experiment from A to G

Each participant (7 students for each group: Japanese, Chinese, Korean, 8 students Vietnamese) listens to the utterances from A to G below (within fillers, within pause, without fillers) and evaluates the level of politeness from the number 1 to 4 (4: very polite, 3: polite, 2: not very polite, 1: not polite). Kinds of fillers are "ano (o)", "chotto", "saa", "ee", "maa", "hai", "eet", "koo".

The utterances of Japanese and English translation are as follows:

A: (Ano, Anoo, Ano(strong, pitch sound), Anoo(strong, pitch sound)) sumimasen, konohen ni kooban arimasu ka. [Excuse me, but is there any police box around this area?]

B: (Chotto, Saa, Saa chotto) wakarimasen. [I don't know very well.]

C: (Ee), kyoo wa konoba ni omaneki itadaki arigatoo gozaimasu. [Um, thank you very much for inviting me this place.]

D: Kono hon omoshiroidesukara (maa) hima ga attara yonde kudasai. [This book is interesting, therefore if you have a time, please try to read.]

E: Ano hito hontoo ni yasashikute iihito nandesu yo (hai). [That person is really kind and good, right.]

F: 123 + 234 wa? (Eeto) 357[What about 123 + 234? Well, 357.]

G: Watashi wa (pause, ano) imamade (pause, ee) nakanaka jibun no kimochi o(pause, koo) umaku hyoogen dekinakattanodesu ga.

[I have never been able to express my feelings well until now.]

2.3. Used equipment and software

A: About brain activity experiment

Hitachi Medical Corporation Optical Topography ETG-7100

The name of fNIRS is often called instead of optical topography. fNIRS is a noninvasive method that measures relative changes in oxygenated hemoglobin by detecting changes in cerebral blood flow. Using fNIRS, we can visualize brain activity. Additionally, a significant advantage for research participants is that they can sit in a chair during the experiment, rather than lying down.

B: about listening experiment

Multi-Speech 3700 KAY Elemetric Company

This study verifies three hypotheses.

A: Cerebral blood flow in brain increases when fillers appear in the utterance.

B: The evaluation of politeness of fillers differs between Chinese & Vietnamese and Korean & Japanese. This is because Chinese and Vietnamese are isolating languages, whereas Korean and Japanese are agglutinative languages.

C: The evaluation to politeness of fillers depends on kinds, pitch of sound, length of sound and strength of sound.

3. Results

3.1. Result of hypothesis A

The oxygenated hemoglobin by detecting changes in cerebral blood flow of Broca's area is shown as orange wave.

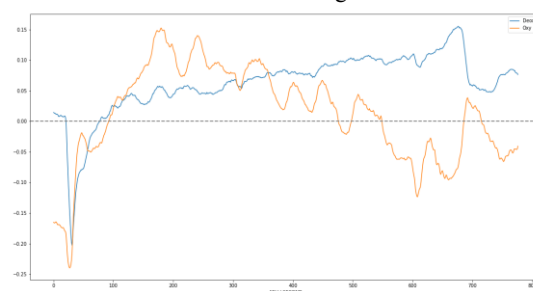


Fig1 Japanese 1

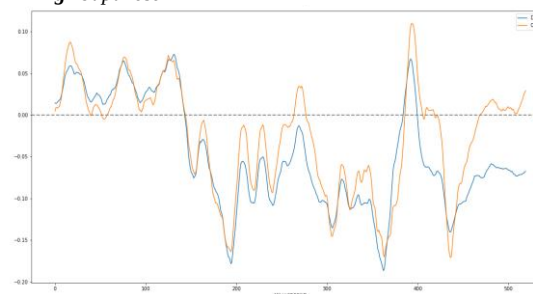
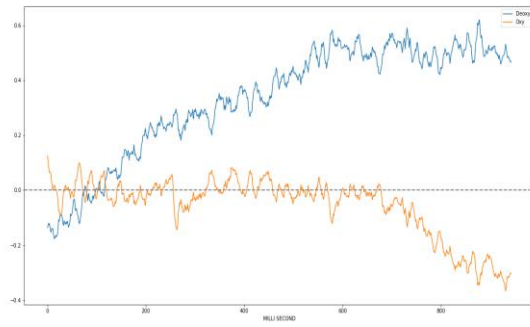
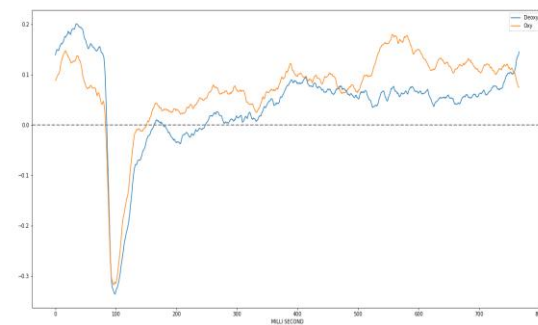
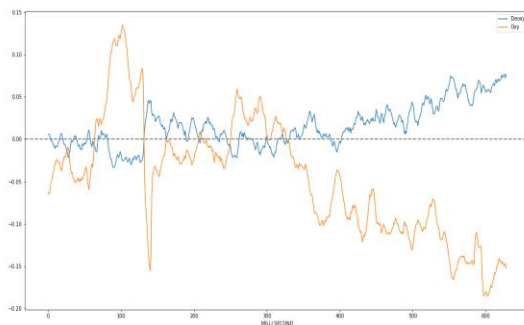
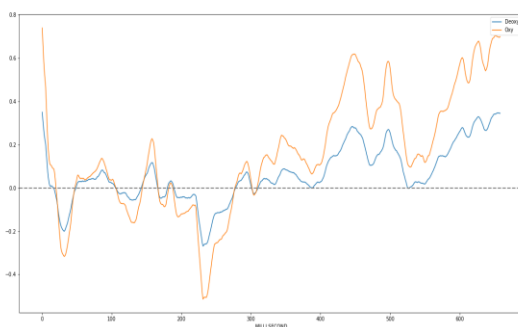
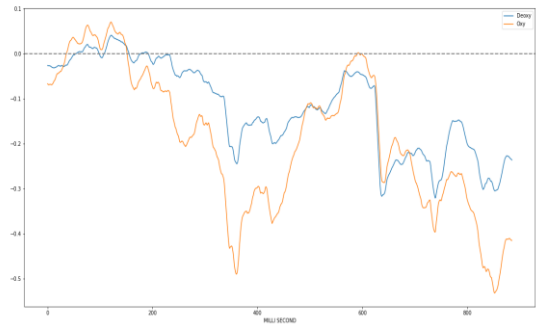
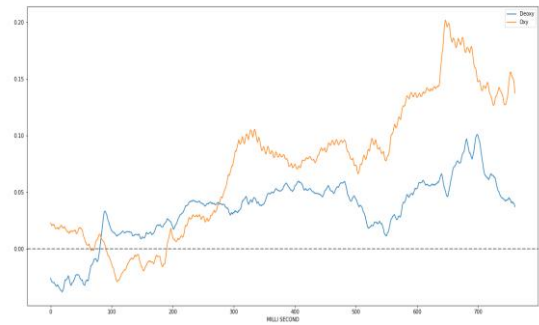
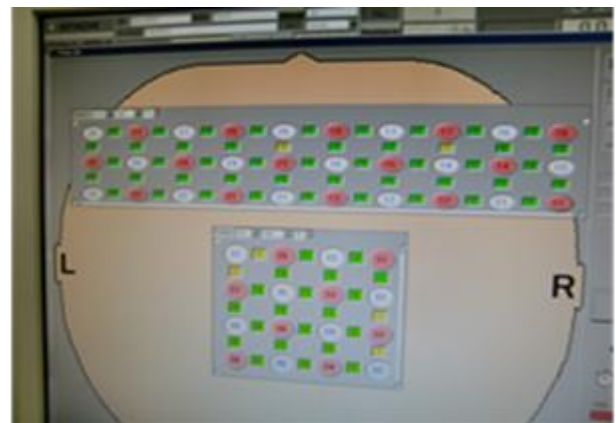


Fig2 Japanese 2

**Fig3 Chinese 1****Fig4 Chinese 2****Fig5 Korean 1****Fig6 Korean 2****Fig7 Vietnamese 1****Fig8 Vietnamese 2****Fig9 Broca's area location(green No.17)**

After about 6-7 seconds from the utterance(filler), cerebral blood flow in brain becomes increasing or peak. The relationship between fillers and the seconds of 8 participants are below although it is the researcher's visual check.

Japanese1: She uses fillers "etto" at 6 seconds and "ano" at 13 seconds. The increase and peak of oxygenated hemoglobin are shown from 10 to 20 seconds.

Japanese2: He uses fillers "ee" and "ano" at 32 seconds around. The increase and peak of oxygenated hemoglobin are shown from 37 to 40 seconds.

Chinese 1: She uses fillers "nanka" "mm" at 37 seconds around. The increase and peak of oxygenated hemoglobin are shown from 38 to 43 seconds.

Chinese 2: He uses filler "nanka" at 11 seconds around. The increase of oxygenated hemoglobin is shown from 10 to 17 seconds.

Korean 1: She uses filler “ettoo” at 3 seconds. The increase and peak of oxygenated hemoglobin are shown from 8 to 11 seconds.

Korean 2: He uses filler “nanka” at 41 seconds. The increase of oxygenated hemoglobin is shown from 47 to 49 seconds.

Vietnamese 1: She uses fillers “aa” “an” “an” from 1 to 7 seconds around. The increase and peak of oxygenated hemoglobin are shown from 3 to 16 seconds.

Vietnamese 2: He uses fillers “ee” “a” “nanka” “nanka” from 20 to 25 seconds around. The increase of oxygenated hemoglobin is shown from 25 to 32 seconds.

3.2. Results of hypotheses B and C

The symbol “○” in Fig.10 indicates the highest evaluation of politeness among the participants.

	A ano(o)				B saa, chotto				C ee		D maa		E hai		F eeto		G pause, filler		
	①	②	③	④	①	②	③	④	①	②	①	②	①	②	①	②	①	②	③
Japanese	○								○		○		○		○		○		
Korean			○						○	○	○		○		○	○			
Chinese			○						○	○	○		○		○	○			
Vietnamese	○	○							○	○	○		○		○	○			
A: ①no filler, ②ano(short, flat, no strength), ③anoo(long, flat, no strength), ④ ano(short, pitch, strength), ⑤ anoo(long, pitch, strength)																			
B: ①no filler, ②chotto, ③saa, ④saa, chotto					C: ①no filler, ②ee					D: no filler, ②hai									
E: ①no filler, ②eeto					G: ①no filler, ②pause are included, ③fillers are included														

Fig.10 The evaluation of the utterances

About acoustic analysis of the experiment A [1]
Concerning the acoustic analysis, only the experiment A is shown as the experiment A is especially connected to pitch, length and strength.

From the appearance of formant as the oral resonance parameters, Fig.15 is the most shining. Therefore, the sound of Fig.15 “anoo” is the strongest among 5 utterances.

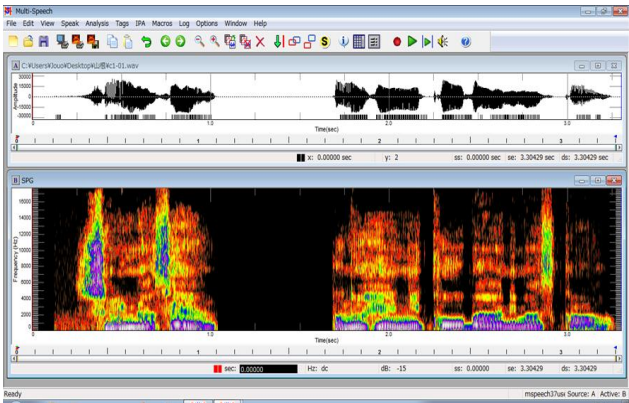


Fig.11 no filler

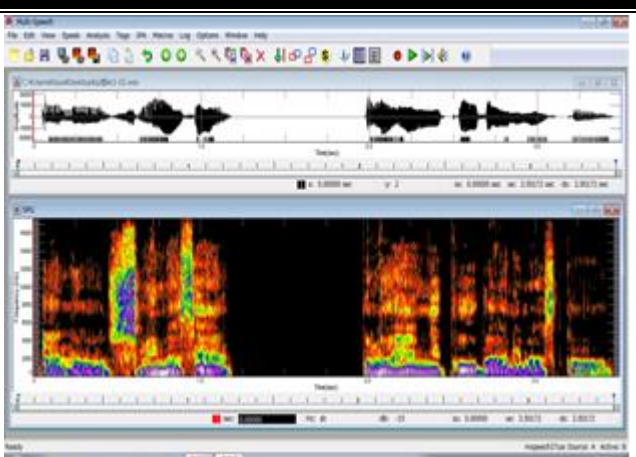


Fig.12 ano (short, flat, no strength)

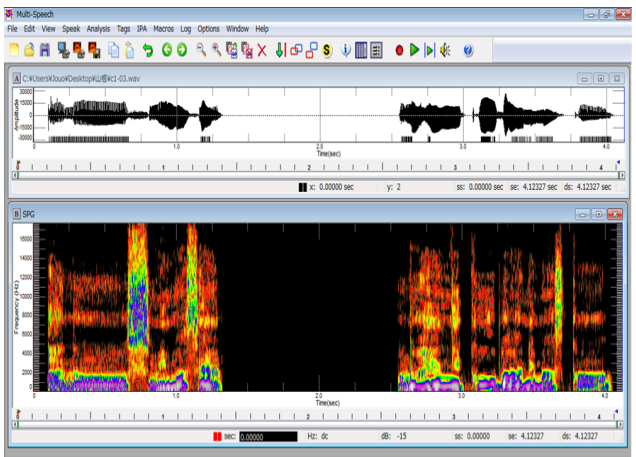


Fig.13 anoo (long, flat, no strength)

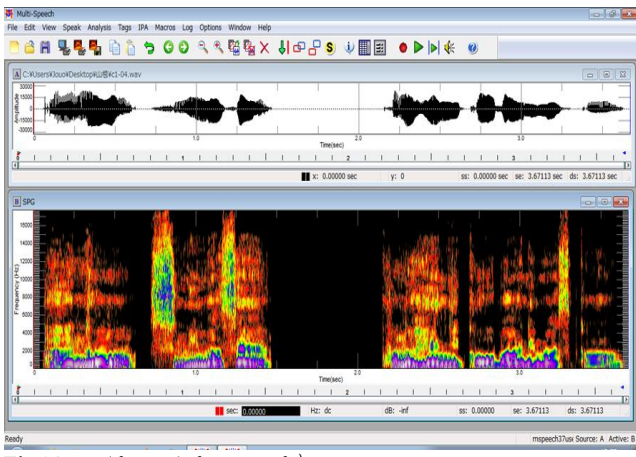


Fig.14 ano (short, pitch, strength)

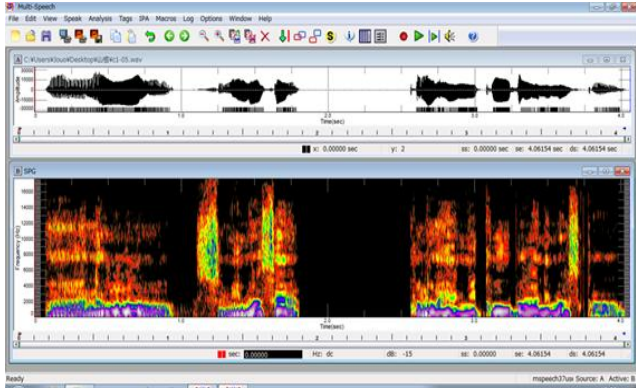


Fig.15 anoo (long, pitch, strength)

We also indicate the pitch analysis. According to the SD of analysis, Fig. 20 “anoo” is the highest. It means the extreme rise and down appears.

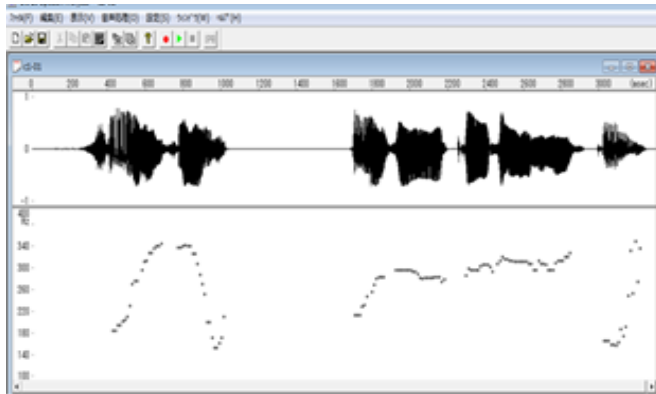


Fig.16 no filler (SD 54.65)

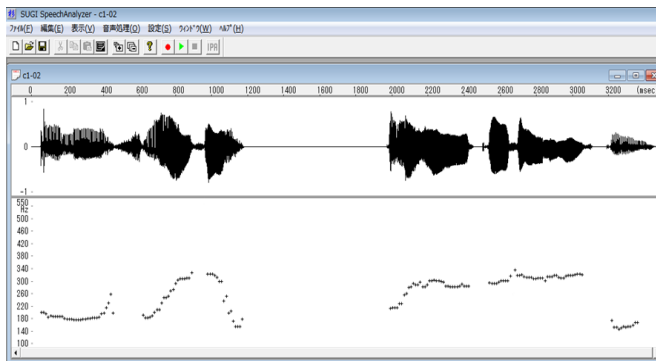


Fig.17 ano (short, flat, no strength) (SD 47.29)

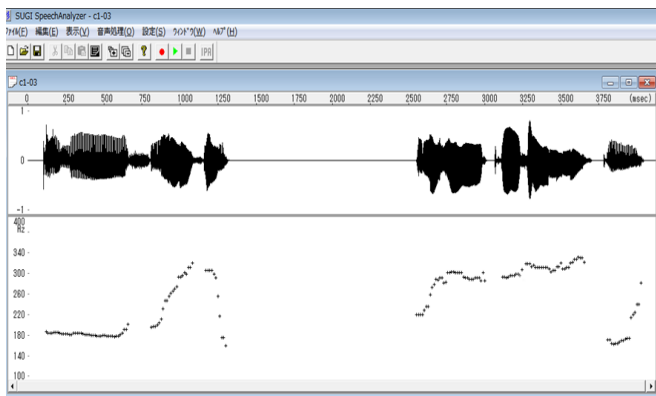


Fig.18 anoo (long, flat, no strength) (SD 42.4)

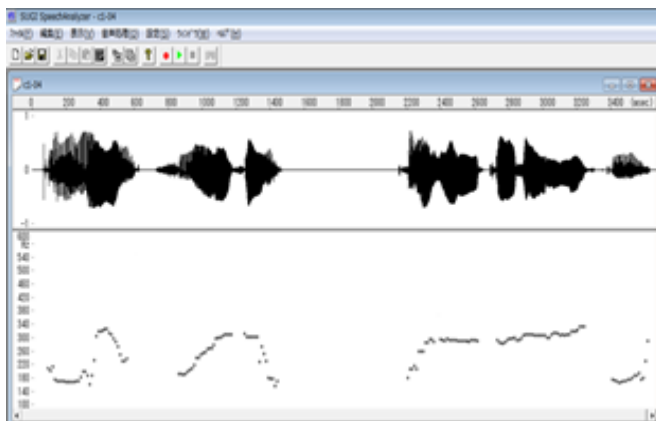


Fig.19 ano (short, pitch, strength) (SD 50.85)

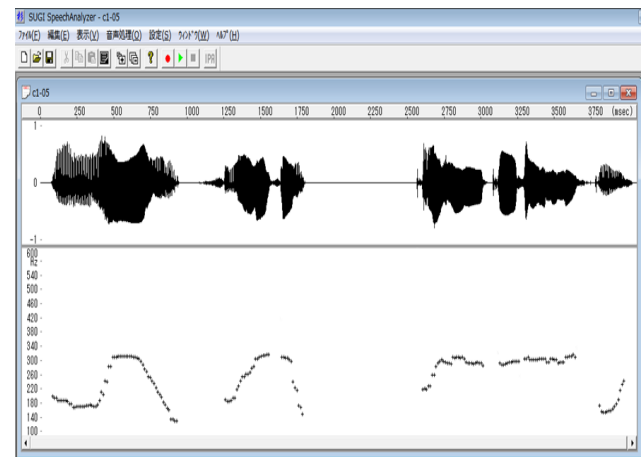


Fig.20 anoo (long, pitch, strength) (SD 55.73)

4. Discussion

About Broca's area, No.17 was analysed. About hypothesis A, all participant's cerebral blood flow increased when they used fillers. However, it was not every filler, not both happiness and sad utterances. About hypothesis B, there was no tendency about the difference of mother language kinds. However, there was a difference between native Japanese and non-native Japanese. Native Japanese indicate high evaluation to non-filler utterances. On the contrary, non-native Japanese indicate high evaluation about “anoo (flat plate intonation)” “saa chotto” “eeto”. About hypothesis C, “ano”, “anoo” pitch and strength of sounds were influenced to politeness. Flat plate intonation and soft sounds had high evaluation. In addition, the length of “noo” was related to politeness.

5. Conclusion

This study only analyzed each 2 participants of Japanese, Chinese, Korean, and Vietnamese by fNIRS. As data from 5-6 participants from each of the four countries has already been collected, we are planning to analyze this data and compare the results. Additionally, we will analyze the sound experiment survey, for which we collected questionnaires from other participants.

6. References

1. Joo, H. *An introduction of experimental phonetics*, Tokyo: Sun Educational (2008)
2. Sakatani, K. (General Ed.). *NIRS - foundations and clinics* -, Tokyo: Shinkoh Igaku Shuppan Co., Ltd (2012)
3. Yamane, C. *Fillers in Japanese discourse*, Tokyo: Kuroshio Publishers (2002)
4. Yamane, C. Para language and politeness, Ed. Ogino, T. *Encyclopedia of honorific language*, Tokyo: Asakura Publishing Co. Ltd. pp.147-154 (2022)
5. Yamane, C., & Takamura R. The influence of fillers to brains: the comparison between Japanese native speakers and Korean native speakers, The proceedings of The Japan Experimental Linguistics Society 16th conference, pp.1-6 (2023)

Acknowledgement

This study was supported by a Grant-in-Aid for Scientific Research (C) JP21K00561.

My gratitude also goes to Koda Masahide M.D., Ph.D., Co-learning Community Healthcare Re-innovation Office, Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University who has transcribed utterances of Chinese and Vietnamese.

ОБЕЩАНИЯ И ПРАКТИКИ ЗА ДИГИТАЛИЗАЦИЯ В БЪЛГАРСКИТЕ ДЪРЖАВНИ УЧИЛИЩА

Христова, М.¹, Нетов, Н.², Нетова, Д.³
 СУ „Св. Климент Охридски“, София, България^{1, 2, 3}
 marieta.filipova@gmail.com¹, nnetoff@feb.uni-sofia.bg², dianaht@feb.uni-sofia.bg³

Резюме: Нашето изследване разглежда характеристиките и разнообразието от инициативи за дигитализация в българските държавни училища. Анализиранията от нас информация беше събрана посредством анкетно проучване, приключило в края на месец април 2024 г., в което участваха 351 български средни училища. Респондентите отговориха на въпроси, свързани с ползите и предизвикателствата на дигитализацията, протичаща в техните училища. Изследването констатира, че през последните години предоставянето и използването на различни дигитални технологии в българските държавни средни училища нарастват, което може да се обясни както с големите държавни инвестиции в дигитализацията на училищата по време на COVID-19 пандемията, така и с по-доброто разбиране и възприемане на технологично подпомогнатото обучение. Приемането на различни ключови дигитални технологии, определящи за дигитализацията на българските държавни училища, не е еднозначно както в контекста на различните видове училища, така и по отношение на големината на населеното място. В този смисъл е препоръчително да бъдат финансирани допълнителни фокусирани обучения или практически занимания, така че ключовите дигитални технологии на бъдещето да бъдат позитивно възприемани от по-широк кръг потребители.

Ключови думи: информационни и комуникационни технологии; дигитално образование; дигитализация

THE PROMISES AND PRACTICES OF DIGITIZATION IN BULGARIAN PUBLIC SCHOOLS

Hristova, M.¹, Netov, N.², Netova, D.³
 Sofia University “St. Kliment Ohridski”, Sofia, Bulgaria^{1, 2, 3}
 marieta.filipova@gmail.com¹, nnetoff@feb.uni-sofia.bg², dianaht@feb.uni-sofia.bg³

Abstract: Our research examines the characteristics and diversity of digitization initiatives in Bulgarian public schools. The information analysed by us was collected through a survey completed at the end of April 2024, in which more than 350 Bulgarian secondary schools participated. Respondents answered questions related to the benefits and challenges of digitization taking place in their schools. The research found that in recent years the provision and use of various digital technologies in Bulgarian public secondary schools has increased, which can be explained both by the large state investments in digitalization of schools during the COVID-19 pandemic, and by better understanding and adoption of technology-assisted learning. The perception of various key digital technologies determining the digitization of Bulgarian public schools is not unambiguous both in the context of the different types of schools and in relation to the size of the settlement. In this sense, it is advisable to fund additional focused trainings or practical activities, so that the key digital technologies of the future start to be positively perceived by a wider range of users.

Keywords: information and communication technologies; digital education; digitization

1. Въведение

Изследванията, свързани с все по-широко разпространеното и бързо внедряване на информационни и комуникационни технологии (ИКТ) във всички области на общественото развитие - от икономиката през здравеопазването до сигурността и личния живот - показват, че един от секторите, в които въздействието на тези технологии е особено силно, е образованието. Решаващата му роля в развитието на обществения живот налага системни задълбочени изследвания на мястото, значението и подходите за внедряване на ИКТ в образователния процес. Въвеждането на дигиталните технологии в образованието е по същество нов инструмент, чрез който може да се промени/повлияе както качеството на образованието, така и начинът, по който знанията и специфичната информация се предоставят и интерпретират. Всеки един от тези елементи оказва силно влияние върху постигането на крайната цел – образованието на младите хора и подготовката им за пълноценно участие в обществените процеси и икономиката на дадена държава [1].

Актуален критичен литературен преглед [2] от 2024 г. използва COOC, HistCite и VOS viewer за систематичен преглед на 1562 образователни статии, публикувани в Computers in Human Behavior (CHB) от 2004 до 2022 г. Въз основа на библиометрични данни това проучване идентифицира важни тенденции, водещи изследователски проекти, ключови статии и изследователски теми. Резултатите, представени от неговите автори акцентират върху нарастващо влияние на технологиите с елементи на ИИ, разширената реалност и електронните учебни ресурси, умерено влияние на образователните инструменти и игри и забавено влияние на управлението на данни. В допълнение, те заключават, че

човешката психология и поведение в технологична среда остават важни за бъдещето теми. Редица престижни научни публикации [3], [4], [5], [6], [7], [8], [9], [10] определят няколко ключови технологии като определящи за дигитализацията на образователната система:

Големи данни (Big Data), извличане на данни/анализ на данни/използване на пространства от данни

В традиционната образователна система на българските държавни средни училища различните функционални образователни институции създават независими платформи за данни според нуждите си, което води до появата на плеяда от „острови“ с образователни данни. Ключово важно е да се ускорят процесите по интеграция и да се изгради единна система за услуги за отворени големи образователни данни, следващи утвърдените добри практики за защита на личните данни на всички заинтересовани страни.

Облачен компютинг (Cloud computing)

Като се отчита ролята на облачните изчислителни технологии в управлението и учебните практики на образователни институции, в [4] са открити следните шест основни предизвикателства:

1. Интеграция и сигурност
2. Съответствие с регулациите
3. Проблеми на управлението
4. Последици за ИТ персонала
5. Предимства за учителите
6. Сигурност на данните

Независимо от това облачният компютинг позволяват на образователните институции да съхраняват и осъществяват достъп до данни и приложения по интернет, а не на локални сървъри. Това им позволява по-голяма гъвкавост, мащабируемост и води до спестяване на разходи.

Мобилна технология (Mobile technologies)

От образователна гледна точка мобилната технология не може да се разглежда като единична и хомогенна технология, а по-скоро като набор от мобилни технологични устройства, които поддържат голямо количество софтуерни приложения. Поради тази причина е важно фокусът на изследванията относно мобилната технология като цяло да включва в себе си и анализ на дизайна и съдържанието на свързаните с нея софтуерни приложения. Да се потърси какви ползи за обучението ни дава комбинацията от мобилна технология и съответните ѝ мобилни софтуерни приложения [5].

Изкуствен интелект (AI)

През 2021 г. ЮНЕСКО публикува книгата „AI and education: guidance for policy-makers“. Целта на тази публикация е да насърчи общото разбиране за потенциалните ползи от ИИ в образованието, както и въздействието му върху необходимите умения, свързани с бързото навлизане на тази технология в икономическата и социалната практика на съвременните общества. През 2023 г. ЮНЕСКО публикува втора книга, със заглавие „Guidance for generative AI in education and research“. Несъмнено технологиите с елементи на ИИ и придобилите популярност големи езикови модели могат да осигурят ценна подкрепа и ресурси за образователния процес.

Интернет на нещата (Internet of things)

Интернет на нещата (IoT) е вече добре установена парадигма, която позволява комуникацията между цялостната образователна инфраструктура и разнообразни електронни устройства и сензори посредством интернет. Все по-популярен става и терминът „персонален интернет на нещата“ (PIoT) който описва дигитална екосистема, която взаимодейства с околната среда в контекста на човек и заобикалящата го среда, използвайки данни, събрани от сензори. Тези технологии ще намират все по-голямо приложение в процесите по дигитализацията на образователната система.

Блокчейн (Blockchain technologies)

Блокчейн е технология за разпределен цифров регистър, която революционизира бизнеса, индустриите и търговията, като елиминира необходимостта от централен орган за съхранение и контрол. Внедряването на базирани на блокчейн решения е изправено пред проблеми, свързани с ниска скалируемост, ограничена оперативна съвместимост, липса на стандартизация и недостиг на квалифицирани разработчици. Освен това на организационно ниво все още има липса на осведоменост, разбиране, липса на управление и регулаторни протоколи [9].

Социални медии (Social media)

Социалните медии с многобройните си функции могат да се използват не само като средство за комуникация, но и като инструмент в дигиталното обучение. Използването на социални медии като инструмент за представяне на дигитално образователно или учебно съдържание може да помогне за обогатяване на учебно съдържание, разширяване на учебните материали и предоставяне на възможности за лесно и увлекателно споделяне на информация [10].

2. Данни и методология

Проучването „Ключови технологии, определящи за дигитализацията на българските държавни училища“ беше реализирано в периода януари-април 2024 г. сред директори,

заместник-директори и ръководители на направление ИКТ в 351 български държавни училища посредством онлайн платформа за проучване ($n = 351$). Извадката включва 20 начални училища (5,7%), 121 основни училища (34,5%), 125 средни училища (35,6%), 14 обединени училища (4,0%), 63 професионални гимназии (17,9%) и 8 профилирани гимназии или друг вид училища (2,3%). Разпределението на включените в извадката училища по вид населено място е както следва: (1) столицата – 27 училища, или 7,7% от всички включени в изследването училища; (2) областен център – 109 училища, или 31%; (3) общински център или голям град – 69 училища, или 19,7%; (4) малък град – 65 училища, или 18,5%, и село – 81 училища, или 23,1%.

Целта на проучването е да установи оценката на респондентите за идентифицираните от нас ключови за дигитализацията на българските училища технологии и степента на тяхното възприемане. За въпросника е използвана платформа за онлайн проучване и резултатите са анализирани с помощта на софтуер SPSS (Statistical Package for the Social Sciences) за статистически анализ.

3. Резултати и дискусия

В табл. 1 са представени резултатите за степента на възприемане на технологиите за големи данни от българските държавни училища, обхванати от нашето анкетно проучване (Cramer's $V = 0,145$, Approximate Sig. = 0,076).

Таблица 1. Степен на възприемане на технологиите за големи данни

Вид училище	Не се използват	Използват се слабо	Донякъде се използват	Използват се често	Използват се в най-голяма степен
Начално	15.0%	25.0%	25.0%	20.0%	15.0%
Основно	28.9%	24.0%	32.2%	12.4%	2.5%
Средно	15.2%	17.6%	38.4%	23.2%	5.6%
Обединено	14.3%	35.7%	28.6%	21.4%	0.0%
Професионална гимназия	15.9%	19.0%	30.2%	28.6%	6.3%
Друг вид училище	12.5%	50.0%	12.5%	25.0%	0.0%
Общо	19.9%	21.9%	33.0%	20.2%	4.8%

Получените резултати показват ниска до средна степен на възприемане на тази технология. Близко половината (41,8%) от включените в изследването държавни училища не използват или използват слабо тази технология и само 25% я определят като често или много често използвана. Една трета от училищата определят технологията като такава, която донякъде се използва в тяхната практика. Технологията е най-добре възприемана в началните училища и е най-малко популярна в основните училища.

Според нас се изграждането на единна система за услуги за отворени големи образователни данни, следващи утвърдените добри практики за защита на личните данни на всички заинтересовани страни, значително ще ускори процесите по възприемането на тази технология от българските държавни училища.

В табл. 2 са представени резултатите за степента на възприемане на технологиите за извличане на данни/анализ на данни/използване на пространства от данни от българските държавни училища, обхванати от нашето анкетно проучване (Cramer's $V = 0.146$, Approximate Sig. = 0.073).

Получените резултати показват сравнително висока степен на възприемане на тази технология. Близко една четвърт (27,3%) от включените в изследването държавни училища не използват или използват слабо тази технология, докато цели 40,2% я

определят като често или много често използвана. Една трета от училищата определят технологията като такава, която донякъде се използва в тяхната практика. Технологиите за извличане, анализ на данни и използване на пространства от данни са най-добре възприемани в началните и средните училища, докато остават най-малко популярни в групата на профилираните гимназии, основните или обединените училища.

Таблица 2. Степен на възприемане на технологиите за извличане на данни/анализ на данни/използване на пространства от данни

Вид училище	Не се използват	Използват се слабо	Донякъде се използват	Използват се често	Използват се в най-голяма степен
Начално	10.0%	10.0%	30.0%	25.0%	25.0%
Основно	16.5%	22.3%	31.4%	24.8%	5.0%
Средно	7.2%	11.2%	35.2%	33.6%	12.8%
Обединено	7.1%	21.4%	42.9%	28.6%	0.0%
Професионална гимназия	6.3%	17.5%	27.0%	41.3%	7.9%
Друг вид училище	12.5%	25.0%	37.5%	25.0%	0.0%
Общо	10.5%	16.8%	32.5%	31.1%	9.1%

В табл. 3 са представени резултатите за степента на възприемане на технологиите за облачни услуги/изчислителен облак (Cloud computing) от българските държавни училища, обхванати от нашето анкетно проучване (Cramer's V = 0.153, Approximate Sig. = 0.036).

Таблица 3. Степен на възприемане на технологиите за облачни услуги/изчислителен облак (Cloud computing)

Вид училище	Не се използват	Използват се слабо	Донякъде се използват	Използват се често	Използват се в най-голяма степен
Начално	15.0%	5.0%	25.0%	25.0%	30.0%
Основно	16.5%	16.5%	36.4%	19.8%	10.7%
Средно	4.0%	14.4%	32.0%	29.6%	20.0%
Обединено	14.3%	21.4%	42.9%	14.3%	7.1%
Професионална гимназия	7.9%	23.8%	17.5%	28.6%	22.2%
Друг вид училище	12.5%	25.0%	37.5%	25.0%	0.0%
Общо	10.3%	16.8%	31.1%	25.1%	16.8%

Получените резултати показват средна степен на възприемане на тази технология. Малко над една четвърт (27,1%) от включените в изследването държавни училища не използват или използват слабо тази технология, докато 41,9% я определят като често или много често използвана. Близко една трета от училищата (31,1%) определят технологията като такава, която донякъде се използва в тяхната практика.

Облачните услуги се възприемат най-добре от началните и средните училища, както и от професионалните гимназии, докато основните и обединените училища остават скептично настроени към тази технология.

В табл. 4 са представени резултатите за степента на възприемане на мобилните технологии в зависимост от големината на населеното място, в което попадат училищата, включени в нашия анализ (Cramer's V = 0.130 , Approximate Sig. = 0.095).

Получените резултати показват висока степен на възприемане на тази технология. Едва 15,7 % от включените в изследването държавни училища не използват или използват

слабо тази технология, докато 57,9% от тях я определят като често или много често използвана. По-малко от една трета от училищата, включени в нашия анализ, (26.5 %) определят технологията като такава, която само донякъде се използва в тяхната практика.

Таблица 4. Степен на възприемане на мобилните технологии

Вид училище	Не се използват	Използват се слабо	Донякъде се използват	Използват се често	Използват се в най-голяма степен
Столица	7.4%	18.5%	40.7%	22.2%	11.1%
Областен център	2.8%	11.9%	24.8%	39.4%	21.1%
Общински център или голям град	0.0%	7.2%	29.0%	36.2%	27.5%
Малък град	1.5%	15.4%	16.9%	46.2%	20.0%
Село	7.4%	12.3%	29.6%	37.0%	13.6%
Общо	3.4%	12.3%	26.5%	38.2%	19.7%

Мобилните технологии се възприемат добре от държавните училища, като най-скептично настроени към тези технологии остават училищата от София.

В табл. 5 са представени резултатите за степента на възприемане на технологиите с елементи на ИИ в зависимост от големината на населеното място, в което попадат училищата, включени в нашия анализ (Cramer's V = 0.150 , Approximate Sig. = 0.011).

Получените резултати показват много ниска степен на възприемане на тази технология. Много над половината (65,2 %) от включените в изследването държавни училища не използват или използват слабо тази технология, докато едва 10,0% от тях я определят като често използвана, а само 2,8% – като много често използвана. По-малко от една четвърт от училищата, включени в нашия анализ, (21,7 %) определят технологията като такава, която само донякъде се използва в тяхната практика.

Таблица 5. Степен на възприемане на технологиите с елементи на ИИ

Вид училище	Не се използват	Използват се слабо	Донякъде се използват	Използват се често	Използват се в най-голяма степен
Столица	40.7%	29.6%	22.2%	7.4%	0.0%
Областен център	31.2%	26.6%	27.5%	11.0%	3.7%
Общински център или голям град	30.4%	26.1%	20.3%	18.8%	4.3%
Малък град	29.2%	38.5%	20.0%	7.7%	4.6%
Село	56.8%	23.5%	16.0%	3.7%	0.0%
Общо	37.3%	28.2%	21.7%	10.0%	2.8%

Технологиите с елементи на ИИ се разглеждат с най-голям интерес от държавните училища, разположени в не много големите градове от тип общински център, като най-скептично настроени към тази технология остават както училищата, разположени в селата, така и училищата от София.

В табл. 6 са представени резултатите за степента на възприемане на технологиите, свързани с интернет на нещата (Internet of things), в зависимост от големината на населеното място, в което попадат училищата, включени в нашия анализ (Cramer's V = 0.138 , Approximate Sig. = 0.046).

Получените резултати показват ниска степен на възприемане на тази технология. По-малко от половината (40,4%) от включените в изследването държавни училища не използват или използват слабо тази технология, докато по-

малко от една четвърт (22,2%) от тях я определят като често използвана, а 18,2% – като много често използвана. По-малко от една четвърт от училищата, включени в нашия анализ, (22,8%) определят технологията като такава, която само донякъде се използва в тяхната практика.

Таблица 6. *Степен на възприемане на технологиите, свързани с интернет на нещата (Internet of Things)*

Вид училище	Не се използват	Използват се слабо	Донякъде се използват	Използват се често	Използват се в най-голяма степен
Столица	25.9%	29.6%	18.5%	14.8%	11.1%
Областен център	11.0%	19.3%	25.7%	25.7%	18.3%
Общински център или голям град	13.0%	18.8%	33.3%	13.0%	21.7%
Малък град	10.8%	27.7%	18.5%	23.1%	20.0%
Село	25.9%	16.0%	14.8%	27.2%	16.0%
Общо	16.0%	20.8%	22.8%	22.2%	18.2%

Технологиите с интернет на нещата (Internet of things) се разглеждат с най-голям интерес от държавните училища, разположени в малки градове или общински центрове, като най-скептично настроени към тази технология остават както училищата, разположени в селата, така и училищата от София.

В табл. 7 са представени резултатите за степента на възприемане на социалните медии като инструмент в дигиталното обучение в зависимост от големината на населеното място, в което попадат училищата, включени в нашия анализ (Cramer's V = 0.155 , Approximate Sig.= 0.005).

Таблица 7. *Степен на възприемане на социалните медии като инструмент в дигиталното обучение*

Вид училище	Не се използват	Използват се слабо	Донякъде се използват	Използват се често	Използват се в най-голяма степен
Столица	11.1%	22.2%	18.5%	37.0%	11.1%
Областен център	3.7%	11.0%	24.8%	32.1%	28.4%
Общински център или голям град	1.4%	8.7%	17.4%	36.2%	36.2%
Малък град	7.7%	15.4%	12.3%	36.9%	27.7%
Село	14.8%	13.6%	28.4%	32.1%	11.1%
Общо	7.1%	12.8%	21.4%	34.2%	24.5%

Получените резултати показват много висока степен на възприемане на тази технология. Под една пета (19,9%) от включените в изследването държавни училища не използват или използват слабо тази технология, докато над половината (58,7%) от тях я определят като често използвана или много често използвана. Малко над една пета от училищата, включени в нашия анализ, (21,4%) определят технологията като такава, която само донякъде се използва в тяхната практика. Възприемане на социалните медии като инструмент в дигиталното обучение е най-позитивно в държавните училища, разположени в малките градове или по-големите градове от тип общински или областен център, като най-скептично настроени към тази технология остават както училищата, разположени в селата, така и училищата от София.

4. Изводи и заключения

През последните години се наблюдава нарастващо влияние на технологиите с елементи на ИИ, разширената реалност и електронните учебни ресурси, умерено влияние на

образователните инструменти и игри и забавено влияние на управлението на данни [2].

Ние анализирахме резултатите от реализирано в периода януари-април 2024 г. сред директори, заместник-директори и ръководители на направление ИКТ в 351 български държавни училища онлайн анкетно проучване, оценяващо степента на възприемане на различните ключови технологии, определящи за дигитализацията на българските държавни училища. Получените от нас резултати показват значително влияние на мобилните технологии и социалните медии, умерено влияние на технологиите, свързани с големи данни (Big Data), извличане на данни/анализ на данни/използване на пространства от данни, изчислителен облак (Cloud computing), и забавено влияние на технологиите с елементи на ИИ и на технологиите, използващи интернет на нещата (Internet of things).

Възприемането на разнообразните ключови дигитални технологии не е еднозначно както в контекста на различните видове училища, така и по отношение на големината на населеното място. В този смисъл е препоръчително да бъдат финансирани допълнителни фокусирани обучения или практически занимания, така че ключовите дигитални технологии на бъдещето да започнат да бъдат възприемани позитивно от по-широк кръг потребители.

5. Благодарности

Публикацията е във връзка с резултатите от работата по научен проект с Договор № 80-10-151 от 17.04.2024 г.

6. Използвана литература

- [1] Varbanova T., Netov N., Vutsova A., "Digitization of the education in Bulgaria: state and general trends," *STRATEGIES FOR POLICY IN SCIENCE AND EDUCATION*, vol. 30, no. 6, pp. 577-599, 2022.
- [2] Jun Liu, Zile Liu, Cong Wang, Yanhua Xu, "Identification and evaluation of educational technology trends from 2004 to 2022: Evidence based on computers in human behavior and horizon report," *Heliyon*, vol. 10, no. 2, p. e24277, 2024.
- [3] Xiaomei Bai, Fuli Zhang, Jinzhou Li, Teng Guo, Abdul Aziz, Aijing Jin, Feng Xia, "Educational Big Data: Predictions, Applications and Challenges," *Big Data Research*, vol. 26, p. 100270, 2021.
- [4] Mohd Naved, Domenic T. Sanchez, Adeline P. Dela Cruz, Larry B. Peconcillo, Emerson D. Peteros, Janine Joy L. Tenerife, "Identifying the role of cloud computing technology in management of educational institutions," *Materials Today: Proceedings*, vol. 51, no. 8, pp. 2309-2312, 2022.
- [5] Marta Gómez Domingo, Antoni Badia Garganté, "Exploring the use of educational technology in primary education: Teachers' perception of mobile technology learning impacts and applications' use in the classroom," *Computers in Human Behavior*, vol. 56, pp. 21-28, 2016.
- [6] UNESCO, AI and education: guidance for policy-makers, 2021.
- [7] UNESCO, Guidance for generative AI in education and research, 2023.
- [8] Bisma Gulzar, Shabir Ahmad Sofi, Sahil Sholla, "Exploring Personalized Internet of Things (PIoT), social connectivity, and Artificial Social Intelligence (ASI): A survey," *High-Confidence Computing*, p. 100242, 2024.
- [9] Gautami Tripathi, Mohd Abdul Ahad, Gabriella Casalino, "A comprehensive review of blockchain technology: Underlying principles and historical background with future challenges," *Decision Analytics Journal*, vol. 9, p. 100344, 2023.
- [10] N. T. Martoredjo, "Social media as a learning tool in the digital age: A review," *Procedia Computer Science*, vol. 227, pp. 534-539, 2023.

Energy efficiency measures to reduce the carbon footprint: A case study of Tirana airport.

Artan Hoxha^{1,*}, Ardit Gjeta¹, Majlinda Alcani¹, Elona Caushi², Albert Shira¹
Polytechnic University of Tirana, Faculty of Mechanical Engineering¹, Environment Manager, Tirana Airport²
ahoxha@fim.edu.al

Abstract: One of the biggest challenges of our century is climate change. The aviation industry is a major contributor to global carbon emissions, and airports are responsible for a significant portion of this impact. To reduce the carbon footprint of airports, it is necessary to implement energy efficiency measures that can help reduce energy consumption and greenhouse gas emissions.

This study analyzes the impact of the use of photovoltaic panels (PV) and the replacement of boiler burners with oil, on energy efficiency and the reduction of carbon emissions at the airport of Tirana. The electricity produced by these panels is used to power airport buildings, lighting systems and other infrastructure, reducing the need for traditional energy sources.

The reviewed period covers the year 2023, because during this year important investments were made with the installation of photovoltaic panels and the replacement of oil boiler burners with LPG burner, in addition to the improvement of the lighting system and the replacement of diesel vehicles with electric vehicles. Energy consumption data is used to calculate the carbon footprint and key energy performance indicators. The results showed that the use of photovoltaic panels and the change of burners of oil boilers brought a significant reduction of the carbon footprint and a significant economic benefit.

Keywords: AIRPORT, CARBON EMISSIONS, ENERGY EFFICIENCY, CASE STUDY, TIRANA AIRPORT

1. Introduction

Airports are not exempt from the effects of climate change and like other sectors, airports must implement measures to mitigate its impacts. Recognized as nationally critical infrastructure, airports play a pivotal role in enabling both mobility and economic development. The World Bank [1] asserts that the carbon footprint is a vital economic and financial gauge. By reducing the carbon footprint, societies can alleviate the expenses linked to climate change, enhance their ability to adapt to its effects, and stimulate investment in renewable energy and other low-carbon technologies. This reduction not only boosts return on investment but also mitigates risks and generates cost savings. Airports, much like other economic entities, have a responsibility to meet international obligations in addressing climate change while also striving to lower their operational expenses. Contemporary airports are dedicated to optimizing resource and energy usage per passenger served, aiming to enhance their competitiveness and minimize energy consumption, carbon dioxide emissions, and waste. Airports serve as vital hubs within the global air transport network. As airport infrastructures expand, there is often concurrent development of businesses, commercial areas, residential zones, and spatial infrastructure in the airport vicinity. These developments are interconnected by surface transportation infrastructure [2]. Airports play a pivotal role as key stakeholders in the value chain of the global air transport industry. They provide essential landside and airfield infrastructure that facilitates interactions between ground-based transportation and air transport modes for both passengers and air cargo shippers.

Compared to conventional civil buildings, airport terminals are distinguished by their vast, open spaces primarily enclosed with glazing systems. These structures accommodate a large volume of passengers and operate for extended durations daily. These factors collectively contribute to exceptionally high energy consumption for HVAC (heating, ventilation, and air conditioning) systems [3]. Energy consumed by airports can be broadly categorized into two main areas: the energy used for airside activities and the energy utilized in providing landside area activities. The airside refers to the movement area at an airport, including adjacent terrain, buildings, and infrastructure, where access is restricted. Conversely, landside refers to parts of an airport, along with adjacent terrain and buildings, that are not within the airside precinct [4]. The dynamic nature of airports, with their round-the-clock operations and emphasis on safe flight operations, as well as the provision of comfortable conditions for passengers within terminal buildings, often leads to heightened energy consumption. Consequently, energy management has become a crucial priority for airports. They are actively pursuing continuous enhancement of energy

performance, given their status as sites with high and intensive energy usage [5].

Furthermore, airport managers recognize the imperative to not only reduce energy consumption but also optimize its efficiency. Additionally, airports face mounting pressure to adopt energy-efficient systems while ensuring compliance with progressively stringent regulations [6]. Crucially, airports worldwide are transitioning towards the adoption of clean energy technologies and implementing practices to mitigate local emissions. This environmental strategy involves replacing fossil fuel-based operations with electricity-based alternatives at airports [7]. Indeed, in recent years, numerous new energy technologies have been developed and implemented as energy sources for airports. These technologies encompass solar photovoltaic (PV), concentrating solar power, wind power, oil and natural gas extraction, steam-generated power production, and electricity transmission [8]. The primary areas of energy consumption at an airport encompass heating, cooling, lighting, and the energy needed for operating various facilities and systems [9]. Additionally, fuel is utilized for the airport's heating boiler systems and emergency generators [10]. As per [11], reducing the carbon footprint presents a substantial business opportunity for enterprises. By minimizing their carbon footprint, businesses can elevate their return on investment, mitigate risks, and realize cost savings. In a separate study, the International Monetary Fund (2021) [12], scrutinizes the potential repercussions of climate change on the global economy. According to their findings, climate change has the potential to exert adverse effects on the global economy, impeding economic growth. Wongsapai et al. [13] analyze the effects of Thailand's long-term low-carbon energy efficiency and renewable energy plan on energy consumption and CO₂ emissions. Their findings indicate that implementing this plan will decrease energy demand and reduce CO₂ emissions. Additionally, the study suggests that the energy efficiency and renewable energy plan will aid Thailand in attaining its energy and climate objectives. Busu [14] endeavors to assess renewable energy efficiency and its influence on CO₂ emissions across the European Union. The study's results reveal an enhancement in renewable energy efficiency within the EU, correlating with a favorable reduction in CO₂ emissions. These findings underscore the significance of renewable energy and energy efficiency measures in mitigating CO₂ emissions. The overall energy consumption stemming from air transport operations results in the emission of greenhouse gases, including CO₂, NO_x, SO₂, H₂O, and others. The total greenhouse gas emissions generated at airports can be

specifically attributed to and measured separately for various activities occurring in the airside area, landside area, and supporting activities. Within the landside area of the airport, greenhouse gas emissions typically arise from passenger and air cargo servicing activities conducted in the terminals. Additionally, emissions result from the provision of heating, cooling, and lighting in these and other airport buildings [15]. To diminish airports' carbon footprints, there is a growing imperative for airport power systems to transition towards greater utilization of renewable energy sources [16]. Mirza et al. [17] investigate the influence of energy efficiency on CO₂ emissions in developing nations. Their research reveals a notable and adverse impact of energy efficiency on CO₂ emissions, with particularly pronounced effects observed in countries with lower CO₂ emissions. This highlights the effectiveness of energy efficiency initiatives in mitigating CO₂ emissions, especially within developing countries. Additionally, the study identifies a similarly significant negative effect of renewable energy consumption on CO₂ emissions, emphasizing the crucial roles of both energy efficiency and renewable energy in reducing CO₂ emissions.

In their study, Wang et al. [18] explore the comparative effectiveness of energy efficiency, renewable energy, and natural gas in mitigating environmental degradation. Their findings unmistakably emphasize energy efficiency as the most powerful option. Improving energy efficiency reduces environmental degradation by decreasing energy consumption and greenhouse gas emissions. Moreover, the research underscores the effectiveness of renewable energy sources in mitigating environmental harm, highlighting their lesser negative impacts compared to fossil fuel-based energy sources. In recent years, there has been an increasing movement towards airports striving to achieve "carbon neutrality" [19].

The carbon footprint quantifies the environmental impact of a firm's activities on climate change. It encompasses all greenhouse gases emitted by the firm, typically measured in metric tons or kilograms of carbon dioxide (CO₂). It's essential to recognize that a firm's carbon footprint exclusively assesses CO₂ emissions directly and indirectly generated by its activities [20]. The heightened attention on carbon dioxide (CO₂) emissions stems from the significant carbon footprint generated by the global aviation industry [21].

Environmental sustainability in air transport has garnered increased attention recently, given its significant impact on climate change and the environment. Within the global airline industry, the emissions of carbon dioxide (CO₂) from aircraft engines have garnered significant attention due to their detrimental impact on the environment [22]. As awareness of climate change grows, the concept of embodied carbon dioxide (CO₂) emissions, often referred to as the "carbon footprint," has become a common environmental performance indicator for products and production activities [23].

Carbon dioxide (CO₂) emissions are directly linked to jet fuel consumption, as each kilogram of fuel consumed by aircraft produces more than three kilograms of carbon dioxide (CO₂) [24].

European airports are highly engaged in the Airport Carbon Accreditation program, with Tirana Airport being among the participants. According to data from 2022 [25], the airports enrolled in the program account for 73.1% of the total air passenger traffic in Europe.

Multiple factors, including weather conditions, building envelope efficiency, energy systems, and occupancy behavior, influence building energy demand [26]. Solar energy systems, encompassing technologies like solar thermal collectors and photovoltaic (PV) modules, along with wind energy, have emerged as pivotal solutions for clean and sustainable energy production on a global scale. Lee and Baek [27] pioneered a method to compute solar collector efficiency based on heating load, proposing several parameters to enhance the assessment of collector efficiency.

Solar heating presents numerous advantages, such as decreased fuel consumption, diminished pollution, and minimal maintenance needs [28]. Hansen and Vad Mathiesen [29] examined the influence

of solar thermal energy on carbon dioxide emissions resulting from both technological advancements and fuel substitution. Their research revealed that solar heating yields greater carbon savings, particularly during the summer months, owing to its higher solar ratio. Integrating solar thermal systems can curtail overall fuel consumption, thereby reducing reliance on this heavily sought-after energy source in the future.

2. The case study

Tirana International Airport, also known as Mother Teresa International Airport (TIA), is the primary airport of Albania and serves as the main air gateway to the country. It is located approximately 17 kilometers northwest of the capital city, Tirana. The airport officially opened in 1958 and has undergone several phases of modernization and expansion to accommodate the growing air traffic and to meet international standards. The airport features a large passenger terminal offering a range of services and amenities for travelers, including duty-free shops, restaurants, cafes, banking services, and currency exchange points. The terminal is equipped with modern baggage handling technology and has the capacity to manage a high volume of passengers. Future plans include further expansion of the terminal and runway infrastructure to improve capacity and meet the increasing demands of travelers and airlines. Tirana International Airport continues to play a crucial role in connecting Albania with the world and in promoting the country's economic and tourism development.

TIA is now the airport with the highest flow in the region, with over 2.8 million passengers only during the first 4 months of 2024. This is also thanks to the new lines and the possibilities for connections to countries that were not easily accessible before. In 2023, Tirana International Airport has welcomed 40% more passengers, reaching a figure of 7,257,574 passengers.

Tirana International Airport (TIA) employs an integrated environmental, energy, and quality management system in accordance with international standards ISO 14001, ISO 50001, and ISO 9001. This means that the company meticulously follows a comprehensive process that encompasses everything from policy statements to improvement actions.

TIA has continued implementing its Energy and Carbon Reduction Action Plan to achieve objectives and targets for reducing energy consumption and increasing energy efficiency.

As part of the European airport industry's response to the challenge of climate change and in recognition of its efforts to optimize energy management and reduce CO₂ emissions, TIA has been a participant in the CO₂ movement since 2011/2012.

TIA achieved Level II Reduction accreditation from Airports Council International Europe, which was last renewed in January 2020. TIA has reviewed its Carbon Management Plan with the aim of increasing energy efficiency to best practice benchmarking. This includes improving processes and meter readings, investing in new energy-efficient technologies such as LED lighting, and evaluating the use of alternative renewable energy sources. TIA also emphasizes strong collaboration and leadership with all airport stakeholders to raise awareness and implement common initiatives, setting a long-term action plan for Energy and Carbon Reduction.

TIA considers taking action to reduce its carbon footprint especially important, regardless of the size of its contribution. As part of the airport industry's response to the challenge of climate change, TIA has been part of the initiative since 2011. Currently, TIA holds a Level 2 - Reduction airport carbon accreditation certificate, recognizing its efforts to manage and reduce CO₂ emissions.

TIA can produce both direct emissions (Scope 1) and indirect emissions (Scope 2). However, together, these sources account for a small part of the total emissions at the airport. Most emissions fall within Scope 3, primarily attributable to aircraft. In 2022, TIA's total calculated and validated carbon footprint was 1,682 tons of CO₂.

The emissions for the years 2011–2022 are given in Figure 1&2.

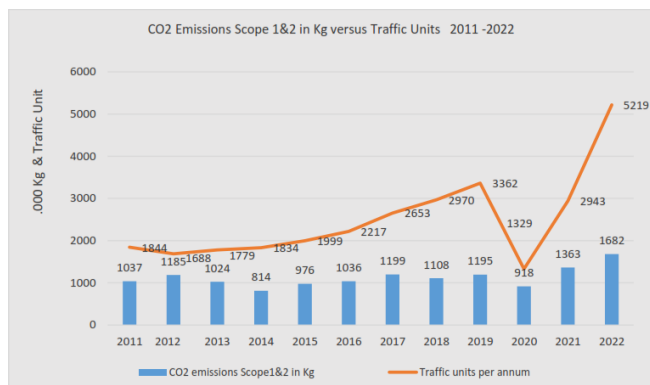
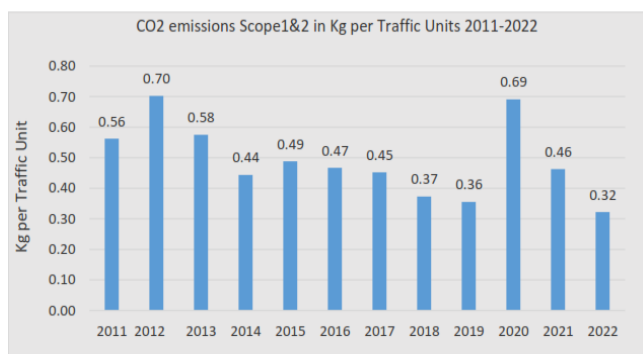
Fig 1. Total CO₂ emissions, 2011-2022

Fig 2. Emissions per traffic unit, 2011-2022

The largest CO₂ emission sources at TIA are Scope 1, which is owned and controlled sources of direct emissions. These include heating boilers, TIA ground support equipment, emergency power, and fleet vehicles, accounting for 96 percent of the total emissions. In comparison, Scope 2 emissions, which are indirect emissions such as purchased electricity, account for just four percent, primarily due to the very low emission factor of hydropower electricity in Albania.

TIA is currently revising its Carbon Management Plan, aiming to set targets of an average of 0.35 kg CO₂ per traffic unit, for TIA scope 1 and 2 emissions. This is a significant improvement compared to the baseline year of 2011, when the first carbon footprint calculation took place and recorded 0.56 kg of CO₂ per traffic unit. The graphs show that the trend in the Airport's total footprint in the last three years is in the range of 9,00–1,700 tons and the carbon footprint per traffic unit is in the range of 0.69–0.32 kg. Emissions per traffic unit average 0.49 kg CO₂ per passenger, 20-40% reduction with the base year. Emissions according to scope are given below.

Table 1 Emissions according to scope 1&2

Airport Operator Source Break Down:			
Mobile Sources	Scope 1	(t CO ₂ e)	905.0
Stationary Sources	Scope 1	(t CO ₂ e)	414.0
Process Emissions	Scope 1	(t CO ₂ e)	-
Process Emissions	Scope 2	(t CO ₂ e)	44.2
Total Scopes 1+2		(t CO₂e)	1,363.2

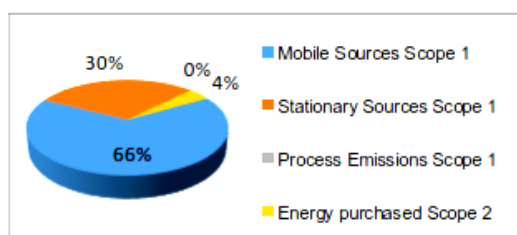


Fig 3. Emissions according to scope 1&2

Table 2 Emissions according to scope 3

Scope 3 Source Group Break Down:		
Aircraft LTO cycle, APU usage and engine testing	(t CO ₂ e)	20,451.4
Landside ground access	(t CO ₂ e)	12,214.6
Staff Commute	(t CO ₂ e)	385.5
Waste Municipality	(t CO ₂ e)	2,165.4
Construction works	(t CO ₂ e)	317.7
Third party electricity	(t CO ₂ e)	2.9

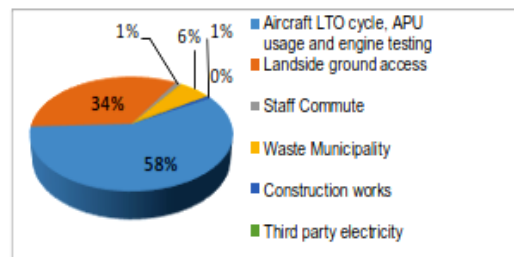


Fig 4. Emissions according to scope 3

Table 3 Total emissions

Airport Total Scope Break Down:		
Airport Operator Scope 1	(t CO ₂ e)	1,405
Airport Operator Scope 2	(t CO ₂ e)	44
Airport Scope 3	(t CO ₂ e)	35,539
Total Gross Emissions	(t CO₂e)	36,988

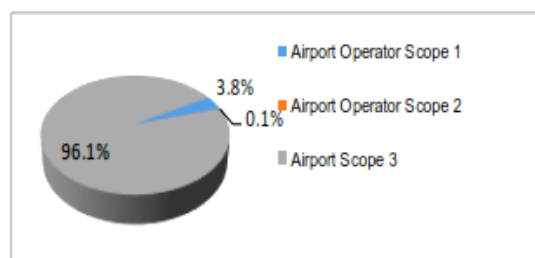


Fig 5. Total emissions

TIA has continued with implementation of its Energy and Carbon Reduction Action Plan for achievement of objectives and targets for energy consumption reduction and energy efficiency.

Energy measures for the scope 1&2 are: Reducing the consumption of conventional fuels, reduction of energy consumption (thermal insulation in buildings, layers of film for protection from the sun, double doors at the arrivals, air curtains at all entrances of the passenger terminal), reduction of electricity for lighting (LED lighting installations, Sensor), design of new buildings in category B (new passenger terminal, panels with hydro and thermal insulation, windows for natural lighting with triplex thermal glass), energy control identification program by consumption units (installation of meters for sub-consumers and third parties), production of renewable energy (installation of photovoltaic panels). Energy measures for the scope 3 are: expansion of urban transport (adding new destinations and frequencies), bus transportation service (for TIA staff and third parties), waste minimization program and separation at source, installation of bins with separation in public areas and waste collection points.

2. Results and discussion

One of the most important measures for improving energy efficiency and reducing the carbon footprint is the installation of photovoltaic panels. In 2023, Tirana Airport installed two sets of photovoltaic panels: one in the airport parking lot with a capacity of

519.42 KWp and another on the terminal roof with a capacity of 594 KWp.

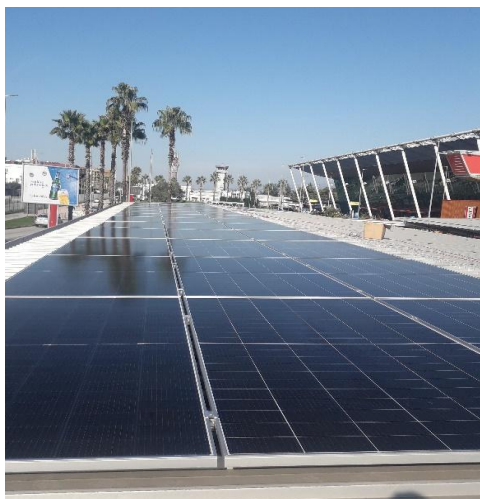


Fig 6. Photovoltaic panels at Tirana International Airport

The annual output of these panels and economic benefit varies based on when they were commissioned, as shown in Table 4&5.

During this year, two sets of photovoltaic panels (PV) with a power of 1 MW will be installed in the airport of Tirana in the new parking lots of the airport, which will affect the further reduction of the carbon footprint and increase the economic benefit.

Table 4 Production and annual profit of PV in the parking lot (T1)

Energy production, T1-Parking 519.42 KWp		Value
2023		
	KWh	ALL
January		0
February	77008.86272	1447766.619
March	116725.4797	2194439.018
April	148043.6785	2783221.156
May	156173.9426	2936070.122
June	148930.2337	2799888.393
July	92092.61081	1731341.083
August	76880.02713	1445344.51
September	59253.09701	1113958.224
October	44063	828384.4
November	21150	397620
December	16842	316629.6
TOTAL	957,162.9322	17,994,663.13

Table 5 Production and annual profit of PV in the terminal roof (T2)

Energy production, T2-Terminal 594 KWp		Value
2023		
	KWh	ALL
January		0
February	0	0
March	0	0
April	0	0
May	0	0
June	7182.445216	135029.9701
July	108593.1803	2041551.789
August	90421.50427	1699924.28
September	69314.84541	1303119.094
October	52043	978408.4
November	25070	471316
December	24413	458964.4
TOTAL	377,037.9752	7,088,313.933

The amount of CO₂ that will be emitted by a diesel power plant to produce a certain amount of electricity will depend on the efficiency of the power plant and the characteristics of the oil used. However, according to the US Environmental Protection Agency (EPA), the average CO₂ emission from the production of one MWh of electricity from power plants that use oil as an energy source is about 623.7 kg of CO₂ and power plants that use natural gas is 400 kg of CO₂.

According to this, a diesel power plant producing 957.16293 MWh of energy will emit about 596,982.52 kg (or about 596.983 tons) of CO₂ into the atmosphere and a natural gas power plant will emit 382,865.17 kg (382.865 tons) of CO₂ into the atmosphere.

A diesel power plant producing 377.03797 MWh of energy will emit about 235,158.8 kg (or about 235.159 tons) of CO₂ into the atmosphere and a natural gas power plant will emit 150,815.19 kg (150.815 tons) of CO₂ into the atmosphere.

In total a diesel power plant will emit 832,141.1 kg (832.14 tons) of CO₂ into the atmosphere and a natural gas power plant will emit 533,680.36 kg (533.68 tons) of CO₂ into the atmosphere.

The use of clean and sustainable energy sources, such as the energy produced by photovoltaic panels, helps reduce greenhouse gas emissions such as CO₂, reduces the carbon footprint and provides an economic benefit.

The heating-cooling system for the terminal is based on: Heating during the winter season is done with three boilers, Hover brand 2x410 Kw and 1 boiler 360 Kw. During the summer season, cooling is done with 2 chillers. The ventilation system also works during the 2 seasons.

The heating-cooling system for the new terminal extension will be with heat pumps, for heating and cooling with air. There will also be AHU air handling units for the ventilation system. This system is under construction.

One of the measures to improve energy efficiency and reduce the footprint at Tirana airport is the replacement of oil boiler burners with LPG (liquefied petroleum gas) burners.

From the data of 2023, oil boilers have consumed the amount of oil according to the following table:

Table 6 Annual oil consumption in boilers

Month	Consumption (l/month)
October	Ventilation
November	9300
December	18800
January	28900
February	32900
March	19600
April	15200
May	9000
Total	133700 liters

Based on these data, the amount of CO₂ emitted during the winter season from the operation of the boilers will be: From calculations, 1 kg of oil produces approximately 3.15 kg of CO₂, and 1 liter of oil produces 2.62 kg of CO₂.

In our case during the heating period, we will have the production: 133700 liters of oil x 2.62 kg CO₂/liter = 350.254 kg CO₂ = 350.24 tons of CO₂.

LPG has a composition of 40% propane and 60% butane. Propane (C₃H₈): Amount of CO₂ produced by burning 1 liter of propane: about 1.51 kg CO₂. Butane (C₄H₁₀): Amount of CO₂ produced by burning 1 liter of butane: about 1.71 kg CO₂.

The amount of CO₂ for each LPG component: - For propane: 40% of 1 liter of LPG is propane: 0.40 * 1.51 kg CO₂ = 0.604 kg CO₂; For butane: 60% of 1 liter of LPG is butane: 0.60 * 1.71 kg CO₂ = 1.026 kg CO₂. So, burning 1 liter of LPG with 40% propane and 60% butane produces about 1.63 kilograms of CO₂.

In the case of LPG, the total CO₂ output based on the boiler manufacturer's data is 137,396 liters x 1.63 kg CO₂/liter = 223,955.48 kg CO₂ (223,955 tons CO₂).

The reduction of carbon emissions because of replacing the burners will be: $350.24 - 223.955 = 126.285$ tons of CO₂.

Economic benefit:

In Albania, 1 liter of oil costs 184 ALL and 1 liter of LPG costs 62 ALL.

Total value of oil 24,600,800 ALL. The total value of LPG based on the data of the boilers received from the manufacturer is 8,518,540.84 ALL. The economic benefit of replacing the boiler burner, regardless of the initial cost, will be approximately 16,082,259.2 ALL (160,822.59 €).

CONCLUSIONS

This study investigated energy efficiency measures at Tirana International Airport (TIA) to reduce the carbon footprint and realize an economic benefit.

Tirana International Airport (TIA) uses an integrated environmental, energy and quality management system in accordance with international standards ISO 14001, ISO 50001 and ISO 9001.

TIA has achieved Level II Reduction accreditation from the International Council of Airports in Europe. The results show that the trend in the total carbon footprint of the airport in the last three years is in the range of 9,00-1,700 tons and the carbon footprint per traffic unit is in the range of 0.69-0.32 kg. Emissions per traffic unit average 0.49 kg CO₂ per passenger, 20-40% lower compared to the base year.

During the study period, two sets of photovoltaic panels with a capacity of 519.42 KWp and 594 KWp were installed at Tirana airport. As a result, the reduction of greenhouse gas emissions was 832.14 tons of CO₂ in the atmosphere, if the energy will be produced by the diesel power plant and 533.68 tons of CO₂ if the energy will be produced by the natural gas power plant with an economic profit of approximately 250,000 euros.

Replacing the oil burners with an LPG burner at the three boilers reduced 126.28 tons of carbon emissions and the economic profit was about 160,822 euros.

References

- World Bank. (2022). The Economic and Financial Implications of Carbon Footprints. World Bank Group. Accessed October 19, 2023. <https://openknowledge.worldbank.org/handle/10986/36704>.
- Ferrulli, P. (2016). Green airport design evaluation (GrADE) – Methods and tools improving infrastructure planning. *Transportation Research Procedia*, 14, 37813790. <https://doi.org/10.1016/j.trpro.2016.05.463>
- Kotopoulos A, Nikolopoulou M. Evaluation of comfort conditions in airport terminal buildings. *Building and Environment*. 2017; 130:162-178. <https://doi.org/10.1016/j.buildenv.2017.12.031>.
- Rossi Dal Pozzo, F. EU Legal Framework for Safeguarding Air Passenger Rights; Springer International Publishing: Cham, Switzerland, 2015.
- Akyuz, M.K., Altuntas, O., Sogut, M.Z., & Karakoc, T.H. (2019). Energy management at the airports. In T. Karakoc, C. Colpan, O. Altuntas & Y. Sohret (Eds.), *Sustainable aviation* (pp. 9-36). Cham, Switzerland: Springer International Publishing.
- Taber, C., & Steele, B. A. (2020). Impact of HVLS fans on airplane hangar air destratification. *ASHRAE Journal*, 62(4), 26-32.
- Sajed Sadati, S.M., Cetin, K., Ceylan, H., Sassani, A., & Kim, S. (2018). Energy and thermal performance evaluation of an automated snow and ice removal system at airports using numerical modeling and field measurements. *Sustainable Cities and Society*, 43, 238250. <https://doi.org/10.1016/j.scs.2018.08.021>
- Barrett, S., Devita, P., Ho, C., & Miller, B (2014). Energy technologies' compatibility with airports and airspace: Guidance for aviation and energy planners. *Journal of Airport Management*, 8(4), 318-326.
- Radomska, M., Chernyak, L., & Samsoniuk, O. (2018). The improvement of energy-saving performance at Ukrainian airports. In T.H. Karakoç, C. Ozgur Colpan & Y. Söhret (Eds.), *Advances in sustainable aviation* (pp. 189-203). Cham, Switzerland: Springer International Publishing.
- Ortega Alba, S., & Manana, M. (2016). Energy research in airports: A review. *Energies*, 9(5), 349. <https://doi.org/10.3390/en9050349>
- McKinsey & Company. (2021). The Business Case for Reducing Carbon Footprints. New York: McKinsey & Company. Accessed October 19, 2023. <https://www.mckinsey.com/capabilities/sustainability/our-insights/the-net-zero-transition-what-it-would-cost-what-it-could-bring>
- International Monetary Fund. (2021). The Economic Impacts of Climate Change. Washington, D.C: International Monetary Fund. Accessed October 19, 2023. <https://www.imf.org/en/Topics/climate-change/climate-and-the-economy>.
- Wongsapai, W., C. Ritkrerkkrai, and J. Pongthanaisawan. 2016. "Integrated Model for Energy and CO2 Emissions Analysis from Thailand's Long-Term Low Carbon Energy Efficiency and Renewable Energy Plan." *Energy Procedia* 100:492–495. <https://doi.org/10.1016/j.egypro.2016.10.208>
- Busu, M. 2019. "Measuring the Renewable Energy Efficiency at the European Union Level and Its Impact on CO2 Emissions." *Processes* 7 (12): 923. <https://doi.org/10.3390/pr7120923>.
- Janić, M. *Greening Airports: Advanced Technology and Operations*; Springer: London, UK, 2011.
- Zhou, Y. (2022). Low-carbon transition in smart city with sustainable airport energy ecosystems and hydrogen-based renewable-grid-storage-flexibility. *Energy Rev.* 1, 100001. doi: 10.1016/j.enrev.2022.100001
- Mirza, F. M., A. Sinha, J. R. Khan, O. A. Kalugina, and M. W. Zafar. 2022. "Impact of Energy Efficiency on CO2 Emissions: Empirical Evidence from Developing Countries." *Gondwana Research* 106:64–77. <https://doi.org/10.1016/j.gr.2021.11.017>
- Wang, Q., Z. Sun, J. Guo, and R. Li. 2023. "The More Effective Option to Combat Environmental Degradation: Energy Efficiency Vs. Renewable Energy Vs. Natural Gas?" *Energy* 283:128512. <https://doi.org/10.1016/j.energy.2023.128512>.
- Falk, M, and Hagsten, E. (2020) Time for carbon neutrality and other emission reduction measures at European airports. *Business Strategy and the Environment*, 29(3), 1448-1464. Retrieved from: <https://doi.org/10.1002/bse.2444>.
- Wiedmann, T. and Minx J. (2007) A definition of 'carbon footprint'. In: Pertsova, C.C. (Ed.). *Ecological Economics Research Trends*. New York: Nova Science Publishers, p. 1-11.
- Filimonau, V., Mika, M., & Pawlusinski, R. (2018). Public attitudes to biofuel use in aviation: Evidence from an emerging tourist market. *Journal of Cleaner Production*, 172, 3102-3110. <https://doi.org/10.5194/acp-13-429-2013>
- Olsen, S.C., Wuebbles, D.J., & Owen, B. (2013). Comparison of global 3-D aviation emissions datasets. *Atmosphere Chemistry and Physics*, 13, 429–441.
- Laurent, A, Olsen, S.I., & Hauschild, M.Z. (2010). Carbon footprint as environmental performance indicator for the manufacturing industry. *CIRP Annals*, 59(1), 37-40. <https://doi.org/10.1016/j.cirp.2010.03.008>
- Mrazova, M. (2014). Sustainable development – The key for green aviation. *INCAS Bulletin*, 6(1), 109-122.
- Airport Carbon Accreditation. (2022). "Airports addressing their CO2 emissions," <https://airportco2.org/airportsacross-the-world.html>.
- Yang T, Bandyopadhyay A, O'Neill Z, et al. (2022). From occupants to occupants: A review of the occupant information understanding for building HVAC occupant-centric control. *Building Simulation*, 15, 913–932.
- Lee KH, Baek NC. A modified efficiency equation of solar collectors. *Energy Procedia* 2014; 48:145–9. <https://doi.org/10.1016/j.egypro.2014.02.018>.
- Cristofari C et al (2015) Innovative alternative solar thermal solutions for housing in conservation-area sites listed as national heritage assets. *Energy Build* 89:123–131. <https://doi.org/10.1016/j.enbui.2014.12.038>
- Hansen K, Vad Mathiesen B (2018) Comprehensive assessment of the role and potential for solar thermal in future energy systems. *Sol Energy* 169:144–152. <https://doi.org/10.1016/j.solener.2018.04.039>
- Ahamd M. State of the Art Compendium of Macro and Micro Energies. *Advances in Science and Technology Research Journal*. 2019; 13, 88–109.