



NAUKA – **TECHNIKA** – TECHNOLOGIA

S E R I A W Y D A W N I C Z A A G H

T O M 8



WYDAWNICTWA AGH

KRAKÓW 2023

NAUKA – TECHNIKA – TECHNOLOGIA

S E R I A W Y D A W N I C Z A A G H

T O M 8



WYDAWNICTWA AGH

KRAKÓW 2023

Wydawnictwa Akademii Górniczo-Hutniczej im. Stanisława Staszica w Krakowie

© Wydawnictwa AGH, Kraków 2023

ISBN 978-83-67427-74-6

Monografia elektroniczna dostępna online.

DOI: <https://doi.org/10.7494/978-83-67427-74-6>

Publikacja może być rozpowszechniana zgodnie z licencją Creative Commons CC BY 4.0

Komitet Redakcyjny serii „Nauka – Technika – Technologia”:

redaktor: *prof. dr hab. inż. Andrzej Gonet* (AGH Akademia Górniczo-Hutnicza, Wydział Wiertnictwa, Nafty i Gazu, Kraków)

sekretarz: *dr inż. Robert Rado* (AGH Akademia Górniczo-Hutnicza, Wydział Wiertnictwa, Nafty i Gazu, Kraków)

Zamieszczone w tomie publikacje zostały pozytywnie zaopiniowane przez recenzentów wyznaczonych przez Komitet Redakcyjny serii.

Opracowanie redakcyjne i korekta: *Joanna Ciągała, Magdalena Grzech*

Projekt okładki i strony tytułowej serii: *Paweł Sepielak*

Skład komputerowy: *Andre*

Wydawnictwa AGH

al. A. Mickiewicza 30, 30-059 Kraków

tel. 12 617 32 28, 12 636 40 38

e-mail: redakcja@wydawnictwoagh.pl

www.wydawnictwo.agh.edu.pl

Spis treści

Mikołaj Cichoń, Paweł Fitner, Mariusz Gibiec

Zawisza4000 – the design of the Polish first bi-liquid rocket's propulsion system	5
---	---

Krzysztof Duda, Szymon Barcentewicz, Rafał Burza, Dariusz Borkowski, Andrzej Bień, Wacław Gawędzki, Zbigniew Marszałek, Jerzy Nabielec, Paweł Turcza, Andrzej Wetula, Tomasz P. Zieliński

Frequency estimation in electric power system measurements	19
--	----

Michał Dudziński, Michał Waląg, Anna Kula

Analiza anizotropii właściwości mechanicznych stopu magnezu typu WE43	33
---	----

Zbigniew Fafara

Badanie bezpieczeństwa rurociągu ze względu na uderzenie hydrauliczne	43
---	----

Jakub Karbowski, Bartosz Bartoszewski, Daria Kokot, Michał Jan Kwiecień, Tymoteusz Turlej

Aerial SLAM algorithm with global alignment using satellite imagery	55
---	----

Robert Karpiński, Adam Kokosza

Charakterystyka mikrostruktury i własności materiałów stosowanych do wyrobu kling szermierczych	67
---	----

Karol Kuglarz, Grzegorz Michta

Kosmiczne materiały – platery	83
-------------------------------------	----

Antoni Ligeża, Weronika T. Adrian, Marek Adrian, Krzysztof Kluza, Krystian Jobczyk, Piotr Wiśniewski, Mateusz Ślażyński, Paweł Jemiolo, Dominik Sepiolo, Bernadetta Stachura-Terlecka, Mateusz Zaremba, Mateusz Szymkowski, Anna Suchenia, Tomasz Potempa

On graph models in knowledge engineering	97
--	----

Szymon Żołyńia, Sebastian Sobula

Symulacja numeryczna obróbki cieplnej odlewu ze staliwa i jej weryfikacja doświadczalna	113
---	-----

Zawisza4000 – the design of the Polish first bi-liquid rocket's propulsion system

Mikołaj Cichoń¹, Paweł Fitner², Mariusz Gibiec²

¹ AGH University of Krakow, Faculty of Physics and Applied Computer Science, AGH Space Systems, Krakow, Poland

² AGH University of Krakow, Faculty of Mechanical Engineering and Robotics, AGH Space Systems, Krakow, Poland

Abstract: AGH Space Systems stands as a student-driven organization situated within the AGH University of Krakow, Cracow, Poland. At present, the team is fully engrossed in the development of Poland's first bi-liquid propellant rocket. Named the "Turbulence" rocket, its propulsion is entrusted to the one-of-a-kind Z4000 engine, utilizing ethanol and nitrous oxide as its propellants. Drawing from their adeptness in hybrid rocket technology, the team ingeniously devised the Z4000 propulsion system, distinguished by its N₂O self-pressurization cycle, elegantly embodied through a piston-dependent pressure vessel solution. Additionally, the team incorporated a composite combustion chamber into the system's architecture. These pioneering elements, among others, firmly position this nitrous oxide-powered demonstrator as a standout contender for prospective applications within the space industry.

Keywords: rocket engine, nitrous oxide, bi-liquid, bi-propellant, design, simulation, self-pressurizing, ethanol, composite, nozzle

ZAWISZA4000 –
PROJEKT SYSTEMU NAPĘDOWEGO PIERWSZEJ POLSKIEJ RAKIETY
NA PALIWO CIEKŁE

Streszczenie: AGH Space Systems jest kołem naukowym działającym w Akademii Górniczo-Hutniczej w Krakowie. Obecnie zespół skupia swoje siły na opracowywaniu pierwszej w Polsce rakiety na paliwo ciekłe o nazwie „Turbulencja”. Rakieta ta jest wyposażona w wyjątkowy silnik Z4000, wykorzystujący etanol i podtlenek azotu jako materiały pędne. Dzięki doświadczeniu nabytemu podczas konstrukcji rakiet hybrydowych zespół stworzył innowacyjny system napędowy Z4000, wyróżniający się cyklem samodoprężania N₂O, który powstał w wyniku rozdzielania zbiornika ciśnieniowego za pomocą ruchomego tłoka. Dodatkowo do architektury systemu dodano kompozytową komorę spalania. Te i wiele innych rozwiązań decydują o innowacyjności opisywanej konstrukcji, która może konkurować z innymi kandydatami do potencjalnych zastosowań w przemyśle kosmicznym

Słowa kluczowe: silnik raketowy, podtlenek azotu, dwuskładnikowe paliwo ciekłe, projekt, symulacja, cykl samodoprężania, etanol, kompozyt, dysza

https://doi.org/10.7494/978-83-67427-74-6_1

1. Introduction

AGH Space Systems is an interdisciplinary, non-profit, student engineering team designing and constructing sounding rockets, planetary rovers and stratospheric balloons' payloads. The team operates at the AGH University of Krakow and associate active members from the wide range of faculties and scientific disciplines.

The team specializes in building hybrid and liquid propellant rocket engines which power their sounding rockets that compete for victory in rocketry competitions against other teams from around the world. Recently, the team acquired the 4th place with their N_2O based hybrid rocket 3-TTK+ on the Spaceport America Cup 2023, taking place in the desert in New Mexico, USA.

Now, AGH Space Systems works on developing the first polish bi-liquid rocket built by university students. The rocket – Turbulence – is designed to break the altitude record for liquid propellant student rockets with estimated apogee of 9 km. The launch is planned for 2024 – the months before the debut of the Turbulence are intended for intensive engine test campaign.

The Turbulence's propulsion system utilizes unique solutions based on the team traditional choice of N_2O as an oxidizer and its self-pressurization properties. The Z4000 engine being currently in development will be a demonstrator of the usefulness of the nitrous oxide based bi-liquid rocket propulsion for the space industry.

2. Z4000 propulsion system overview

The Z4000 propulsion system consists of the pressure vessel, venting and feed system and the engine itself (as seen on Fig. 1). The following description will unfold all its elements in detail.

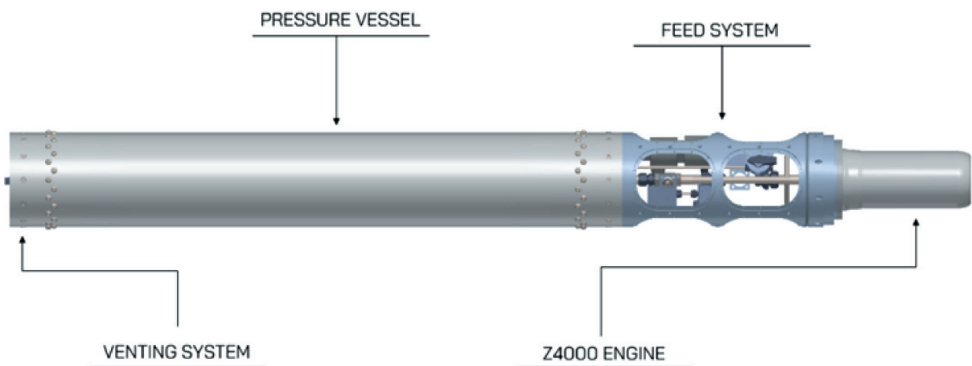


Fig. 1. Z4000 propulsion system schematic

2.1. Preliminary design

The Z4000 engine uses ethanol and nitrous oxide as the rocket propellants in a self-pressurization cycle. The propulsion system was designed to provide 4 kN of nominal thrust and 40 kNs of total impulse. The nitrous oxide has been chosen as an oxidizer – firstly because of the previous team’s experience of handling it in hybrid rocket propulsion systems and secondly because of its high vapor pressure, which is the core factor in a utilized pressurization system. Although N_2O is not a common choice for bi-liquid engines, it has some advantages against other oxidizers (such as liquid oxygen or hydrogen peroxide) mainly due to its low toxicity and ease of liquefying in room temperature, therefore, safe and easy storage.

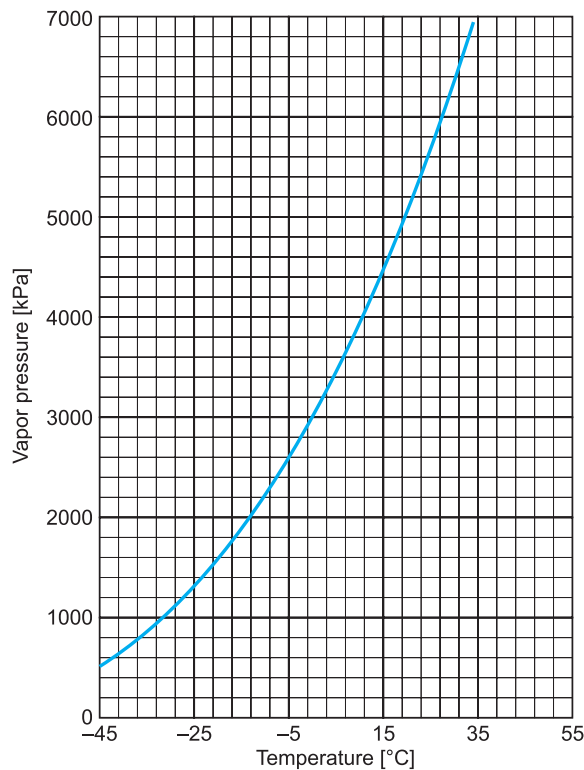


Fig. 2. Vapor pressure of nitrous oxide
Source: based on Palacz (2017)

The choice of ethanol as the fuel was dictated mainly by its price and availability. Since the oxidizer-to-fuel ratio usually oscillates around the factor of 4, the choice of the fuel becomes less crucial for engine overall performance. It was evaluated that various propellants mixed with N_2O give only slight changes in the specific impulse.

By using the NASA CEA software for thermodynamic calculations, it was estimated that the optimal O/F ratio for the Z4000 was 4.2. Due to the lack of any thermal insulation for the pressure tank containing the propellants, the nominal pressure of the system was decided using the team experience with launching rockets in the desert where the average temperature ranges between 35 to 40°C which corresponds to about 60 bars of nominal pressure of the nitrous oxide (see Fig. 2) which was chosen as a nominal value for the propulsion system calculations.

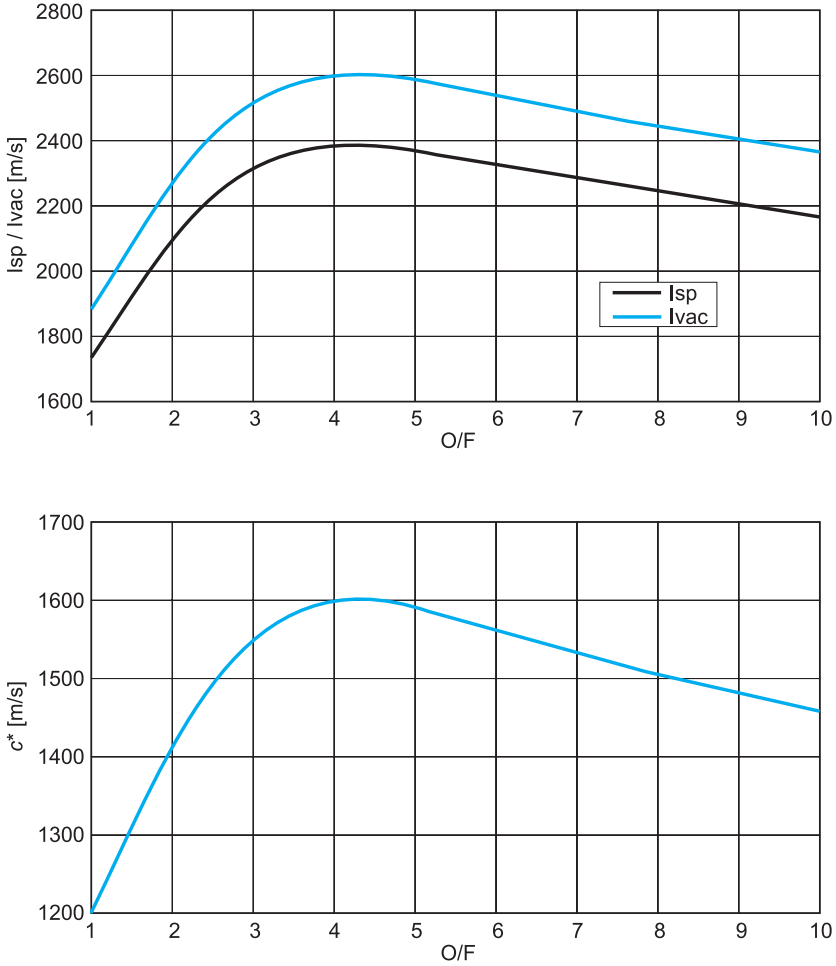


Fig. 3. Specific impulse and c^* as a function of oxidizer-to-fuel ratio

The nominal pressure of 40 bars in the combustion chamber was decided in order to provide sufficient conditions for stable combustion and to ensure a safe pressure drop between the combustion chamber and the feed system in the injector.

Using Equation (1) (Sutton and Biblarz 2001) and data gained from NASA CEA software:

$$I_s = \frac{c^* C_f}{g} \quad (1)$$

it is possible to determine that the specific impulse value is around 243 s. For the engine calculations, the efficiency was established to be 90% in order to give an additional margin for the other propulsion system parameters.

Using Equation (2) (Sutton and Biblarz 2001):

$$\dot{m} = \frac{F}{g I_s} \quad (2)$$

the required mass flow rate can be established to be 1.86 kg/s.

2.2. Pressure vessel

In the previous iterations of the Zawisza engine series developed in AGH Space Systems, where the nitrous oxide tank was pressurized using its self-pressurizing properties and the ethanol tanks used external pressurization with nitrogen, there was a significant issue concerning different pressure drops during engine operation. This led to the essential change in the O/F ratio and made the combustion fuel rich – with a loss on the performance in the outcome. In order to solve this problem and provide synchronous pressure curve, a movable sealed piston was implemented.

The Turbulence's pressure vessel contains both fuel and oxidizer in two compartments, which are separated with a movable piston (Fig. 4). The tank made out of 2017A aluminum alloy is enclosed with two domes screwed into the main tube. With the aim of not mixing those propellants in the tank, o-ring seals were provided on the piston. To prevent the skewing of the piston and to ensure smooth motion along the tank, special PTFE sleeves were designed (Fig. 5). The device itself uses two precise 316L stainless steel tubes going to the outlet of the tank (one is the fuel pipe and the second contains wires passing to the upper part of the rocket) as its guides. The pipes are covered with PTFE layer. Teflon is a great choice mainly because of its self-lubricating properties and excellent chemical compatibility with both nitrous oxide and ethanol.



Fig. 4. A cross-sectional view of the pressure vessel in a CAD program

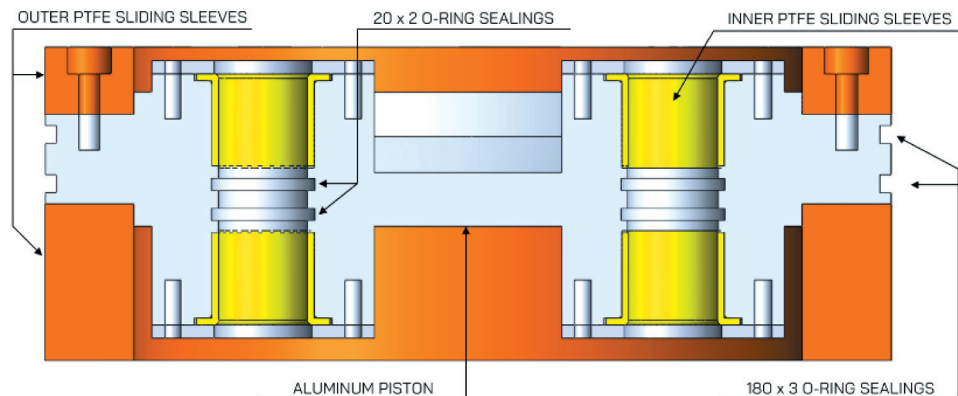


Fig. 5. A diagram of the sealed piston in a cross-sectional view

The pressurization system can be described as blowdown. In the bottom compartment, the nitrous oxide is being refueled. The refueling process starts with ethanol being pumped into the upper compartment. Then the nitrous oxide is being let into the lower compartment and because of its high pressure of around 40 bars it pushes the ethanol out of the tank. The refueling of the N_2O stops when the piston is located in the desired location in order to provide the correct O/F ratio. The process requires leaving a reservoir of vaporous nitrous oxide just beneath the piston. After the procedure is complete, the N_2O is being heat up to the temperature of 30 degrees Celsius using infrared radiators from the outside. The rise in temperature causes a raise in pressure to the nominal 60 bars. Because the gaseous oxidizer acts with this pressure on the piston, the piston itself acts with the same pressure on the ethanol, which provides the propulsion system with the synchronous pressurization of the oxidizer and fuel. Another advantage of the choice of nitrous oxide is that because of its high vapor pressure, the pressure drop during the engine operation is very slow during the majority of the engine burn time.

The position of the sealed piston is measured with a special array of 11 magnetic sensors located along the tank in one of the inner precise tubes. The piston itself is equipped with a small magnet. Its location is being determined by an algorithm that compares readings on all sensors and finds the maximum of the magnetic field strength in real time. The pressure in the tank is monitored by two pressure transducers located on the both sides of the vessel.

2.3. Feed system

The feed system of a rocket engine plays a critical role in ensuring the precise delivery and controlled mixing of propellants for efficient combustion. It was designed to ensure the 5 bars pressure drop across the piping. Five valves in the Turbulence

system that are controlled by servomechanisms are responsible for creating flow through the currently needed line.

The feed system of Turbulence consists of six lines (Fig. 6). The main oxidizer feed line delivers nitrous oxide to the manifold. At the end of the line, brass piston with EPDM o-ring seals is placed. This setup ensures load transfer through the feed system cage and distributes oxidizer above the back injector plate. The main fuel feed line supplies ethanol to the manifold. The valve is fixed in a brass piston with EPDM o-ring seals, connected to a pipe that splits into two curved pipes guided into the manifold. This distribution method supplies fuel to the manifold envelope that feeds the injector from the side.

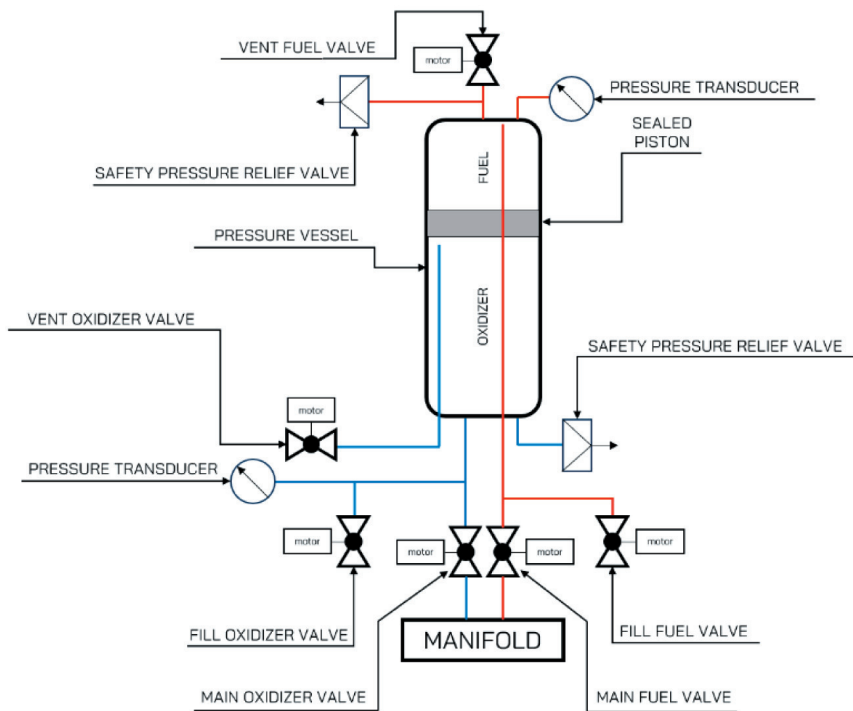


Fig. 6. A schematic of the Z4000 feed system

The fill oxidizer line refills nitrous oxide and the fill fuel line refills ethanol into the pressure vessel. Both of them start in their main propellant line in front of the valves. The lines then exit the feed system cage and terminate with a quick coupler.

The vent oxidizer line allows the excessive gaseous nitrous oxide to escape the system. A PTFE flexible pipe is introduced because vapor forms just under the sealed piston during rocket operation. One end is attached to the piston's base, the other is connected to the pressure vessel's dome. The gaseous nitrous oxide is then released outside the feed system cage. The vent fuel line purges the excess of gaseous ethanol

from the system. Unlike a nitrous oxide venting system, that one does not require a flexible solution because of the static nature of the problem. It starts at the orifice in the pressure vessel's upper dome and terminates with an outlet in the rocket's fuselage.

The valves are operated with steel gears designed with greater torque than is needed in order to open the valve under nominal pressure. The servomechanisms are connected to the flight computer through wires routed through the secondary pressure vessel's pipe, ending in the electronics bay.

The Turbulence hydraulic system is also equipped with two safety pressure relief valves. One valve is attached to the pressure vessel's lower dome, and the second is connected to the vent fuel line's tube using a pipe tee. Two pressure transducers are integrated into the design: one in the pressure vessel's upper dome and the other in the main oxidizer feed line's pipe.

The entire feed system is situated within an aluminum cage, exhibiting a structured configuration featuring eight apertures. Each of these apertures is fitted with a transparent polycarbonate fairing, seamlessly enclosing the cylindrical framework of the cage.

2.4. Z4000 engine

The Z4000 engine as all standard bi-liquid rocket engines has an injector placed in the manifold, combustion chamber and a nozzle (as seen on Figure 7). It is crucial that each of these components are calculated and designed carefully in order to obtain the nominal efficiency. The following is a description and explanation of the design of the Z4000 engine.

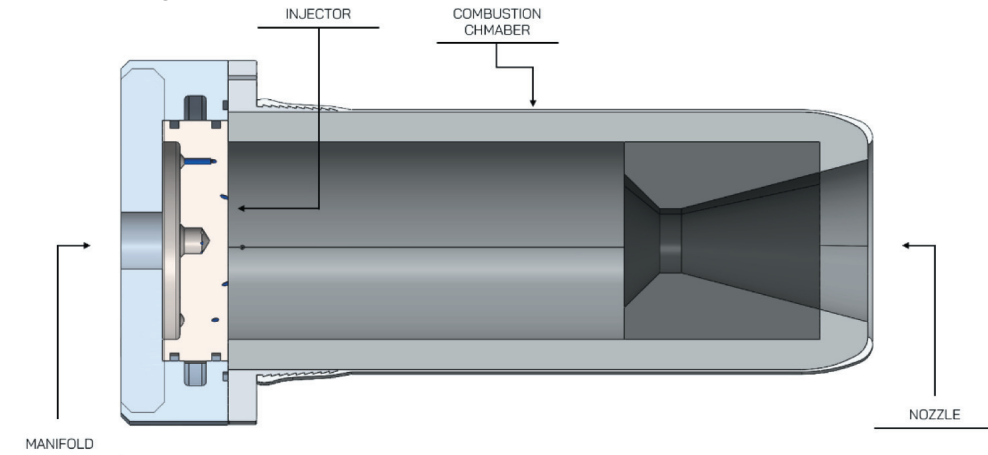


Fig. 7. Z4000 engine schematic

Injector

The Z4000 propulsion system features a cross-impinging injector disk, designed with the intention of achieving efficient mixing of propellants by intersecting oxidizer

and fuel streams. The choice to employ an impinging injector design is driven by its natural ability to enhance propellant mixing, which is especially relevant considering the dual-phase characteristics of nitrous oxide and the compressible single-phase properties exhibited by ethanol. The injector's geometric specifications were established through analytical calculations that encompassed the homogeneous equilibrium model for nitrous oxide and the compressible single-phase model for ethanol. This injector arrangement is characterized by a configuration of triplet ports, involving two oxidizer ports and one fuel port, meticulously arranged within the outer ring of seven ports. In addition, the inner ring accommodates seven quadruplet ports, each comprising three oxidizer ports and one fuel port (Fig. 8). This intricate layout contributes to an exhaust composition favoring a fuel-rich mixture in close proximity to the combustion chamber walls, effectively alleviating the thermal stress experienced by these integral components.

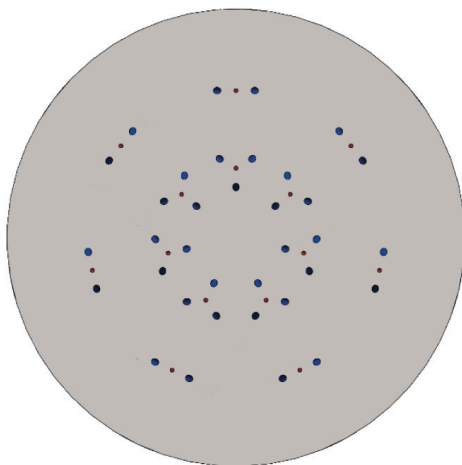


Fig. 8. The injector's plate

The injector, integral to the system's architecture, resides within the manifold. This positioning is reinforced through a joint interconnecting the manifold and the combustion chamber, the internal diameter of which is intentionally designed to be lesser than that of the injector itself, as elucidated in the Z4000 engine schematic. The injector functions as a receptacle for oxidizer, drawn from the upper section of the manifold, and fuel, sourced from lateral inlets. In order to ensure a robust and hermetic seal, the ingress points for the fuel and oxidizer are sealed through the implementation of a pair of PTFE o-ring seals, ensconced between the injector and the manifold.

The composition of the injector disk warrants attention, being precision-machined from aluminum. This decision is rooted in the team's empirical insights gleaned from prior experiences with aluminum injectors in the context of hybrid rockets.

Combustion chamber

In order to properly design the combustion chamber, both structural and thermal analysis is needed. The previous iteration of the Zawisza engine cycle utilized steel as the primary material for the combustion chamber due to its durability and easy access. The new combustion chamber consists of three main parts: aluminum flange, a thick layer of phenolic resin composite and a quite thin layer of carbon fiber with epoxy resin composite. During the burn, the ablative layer ablates and takes heat away from the structural layer. Additionally, an ablative layer serves as a nozzle extension of the nozzle. The carbon fiber-based composites were chosen for the structural layer due to their high strength, low mass, and ability to easily bond with the composite ablative layer. It is secured by the flange ring and inside of it there is a pressure sensor reading pressure from the combustion chamber during ground testing.

Nozzle

Nowadays, a use of bell-shaped CD nozzles in rocket engines is a common practice. Although those nozzles are designed in order to reduce the shockwaves, formed when the hot gas flow is passing by a flexion at the outlet of the throat, so the flow remains uniform and to direct the flow axially and to reduce overall mass – they are difficult and expensive to manufacture. In comparison, relative performance of the conical nozzle is estimated to be 0.98 of the thrust (with opening angle of 15°) and a 80% bell-shaped nozzle is 0.99. In respect to the costs, the trade-off between these two types takes advantage for the conical.

Gaining necessary data for nozzle calculations from NASA CEA thermodynamic calculations the geometry was established. Using Equation (3) (Sutton and Biblarz 2001):

$$\frac{1}{\varepsilon} = \frac{A_t}{A_2} = \left(\frac{\gamma+1}{2} \right)^{\frac{1}{\gamma-1}} \left(\frac{p_2}{p_1} \right)^{\frac{1}{\gamma}} \sqrt{\frac{\gamma+1}{\gamma-1} \left[1 - \left(\frac{p_2}{p_1} \right)^{\frac{\gamma-1}{\gamma}} \right]} \quad (3)$$

The nozzle area ratio was calculated to be 5.52 for the sea level ambient pressure. Using Equations (4) and (5) (Sutton and Biblarz 2001):

$$A_t = \frac{\dot{m}}{p_1} \sqrt{\frac{R'T_1}{\gamma \left(\frac{2}{\gamma+1} \right)^{\frac{\gamma+1}{\gamma-1}}}} \quad (4)$$

$$D_t = 2 \sqrt{\frac{A_t}{\pi}} \quad (5)$$

The critical diameter of the throat was set to 31 mm. The exit diameter, therefore, is 72.6 mm. The divergence and convergence angles were set to standard 15 and 45 degrees respectively (Huzel and Huang 1992). Nozzle throat length was designed in order to minimize any potential shockwaves and turbulences with length-to-diameter ratio of around 0.15 (Tolentino and Mirez 2022), that is 5 mm of length. In order to primarily evaluate the design of the nozzle a CFD simulation was performed and resulted in satisfactory velocity values (Fig. 9). Although the design needs to be evaluated in engine hot fire test.

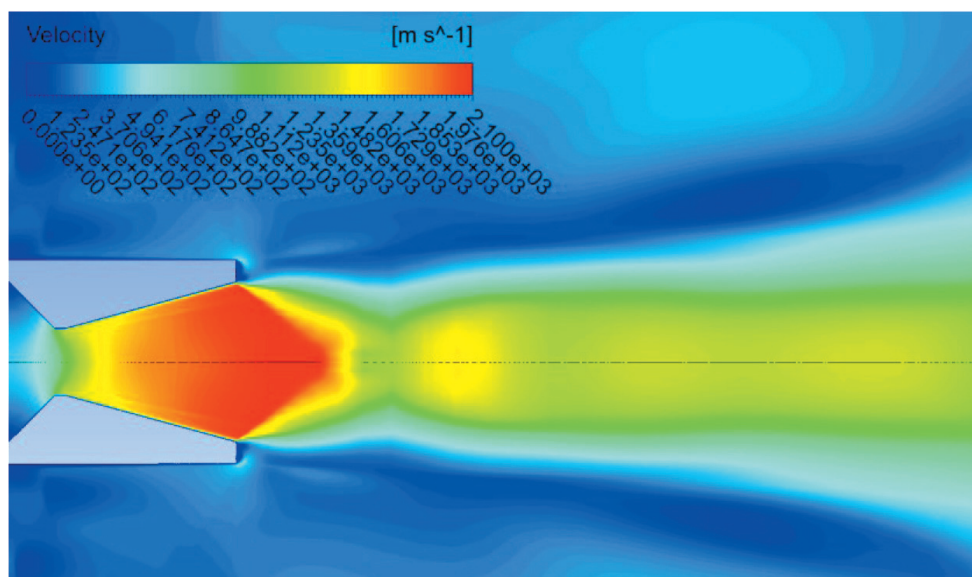


Fig. 9. CFD simulation of the Z4000 nozzle

Since the temperatures in the rocket nozzle – especially in the throat section – reach values up to 2500 K, which exceed most of the available materials' temperature strength in such duration of influence, a cooling solution shall be applied. It was decided that an ablative cooling system would be sufficient to withstand such extreme conditions. Ablative cooling consists in carbonization or graphitization of the coolant surfaces subjected to hot gas flow. For this purpose, a graphite nozzle insert was designed which rests on the nozzle extension in the combustion chamber.

Ignition system

The ignition system for this engine features a pyrotechnic mixture encased in a conical envelope, initiated using an e-match. This design characteristic facilitates a slow and sustained release of flames, which in turn serves as the ignition source for

the propellant vapors. The ignition assembly is inserted through the nozzle into the combustion chamber, utilizing an extending arm, while the arm itself is securely affixed to the launch platform.

3. Tests and development

Zawisza4000 engine is currently undergoing numerous tests for empirical verification of the design and analytical approach. There are two types of tests – cold flow and hot flow. The cold flow test verifies if the pressure drop in the injector and mass flow rate is correct. In order to ensure safety during tests, CO₂ and water are used as propellants since they are good substitutions for nitrous oxide and ethanol (Waxman et al. 2013). The hot flow tests essentially verify the engine operation, therefore, generated thrust, pressure in the combustion chamber and pressure in the pressure vessel. Those tests are usually conducted with stronger substitutes of the final structure of the rocket. Both of these tests are performed on a vertical test stand in order to satisfy gravitational based design of the pressurization system. The thrust is measured with tensometer on which the engine is hanging, and the pressure is monitored with mentioned before pressure transducers. Now, the propulsion system is undergoing hard-bitten iteration process in purpose of ensuring the greatest performance and reliability.

4. Conclusion

The Turbulence's propulsion system is a promising design which can demonstrate a nitrous oxide-based application for liquid propellant space propulsion. Although N₂O is not as dense and energetic oxidizer as other, commonly used in the space industry, it has a potential of finding its usefulness in low-cost rocket and space propulsion systems. Nowadays, a great popularity in this oxidizer among student rocketry constructions can be seen, which demonstrates the possibility of becoming an alternative for the space industry. Nevertheless, further researches have to be conducted in order to improve the database of the specific values for N₂O rocket engine designs and to designate a way of commercialization.

Acknowledgments

The authors would like to express their appreciation to the AGH Space Systems team, especially for the propulsion and mechanical teams involved in the development process of this engine, as well as the past AGH Space Systems members for their effort and extensive research conducted for the first Zawisza prototypes. This endeavor would not be possible without the financial support from the AGH University of Krakow.

Nomenclature

- A_2 – nozzle exit area [m^2]
 A_t – nozzle throat area [m^2]
 c^* – characteristic velocity [m/s]
 C_f – thrust coefficient
 D_t – nozzle throat diameter [m]
 F – nominal thrust [N]
 g – gravitational acceleration [m/s^2]
 I_s – specific impulse [s]
 $\dot{m}$ – mass flow rate [kg/s]
 p_1 – combustion chamber pressure [Pa]
 p_2 – nozzle exit pressure [Pa]
 R' – specific gas constant [$\text{J}/(\text{kg} \cdot \text{K})$]
 T_1 – chamber temperature [K]
 γ – specific heat ratio
 ε – nozzle area ratio

References

- Huzel D.K., Huang D.H. (eds.), 1992, *Modern Engineering for Design of Liquid-Propellant Rocket Engines*, American Institute of Aeronautics and Astronautics, Reston.
- Palacz T., 2017, Nitrous oxide application for low-thrust and low-coct liquid rocket engine, [in:] *7th European Conference for Aeronautics and Space SCIENCES (EUCASS)*. <https://doi.org/10.13009/EUCASS2017-474>.
- Sutton G.P., Biblarz O., 2001, *Rocket Propulsion Elements*, John Wiley & Sons, New York.
- Tolentino S.L., Mirez J., 2022, *Throat length effect on the flow patterns in off-design conical nozzles*, FME Transactions, vol. 50(2), pp. 271–280. <https://doi.org/10.5937/fme2201271T>.
- Waxman B.S., Zimmerman J.E., Cantwell B., Zilliac G., 2013, *Mass flow rate and isolation characteristics of injectors for use with self-pressurizing oxidizers in hybrid rockets*, [in:] *49th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit 2013, July 15–17, 2013, San Jose, CA. Volume 1*, American Institute for Aeronautics and Astronautics, pp. 383–414. <https://doi.org/10.2514/6.2013-3636>.

Frequency estimation in electric power system measurements

Krzysztof Duda¹ , Szymon Barcentewicz¹ , Rafał Burza¹ ,
Dariusz Borkowski¹ , Andrzej Bień¹ , Wacław Gawędzki¹ ,
Zbigniew Marszałek¹ , Jerzy Nabielec¹ , Paweł Turcza¹ ,
Andrzej Wetula¹ , Tomasz P. Zieliński² 

¹ AGH University of Krakow, Faculty of Electrical Engineering, Automatics, Computer Science and Biomedical Engineering, Krakow

² AGH University of Krakow, Faculty of Computer Science, Electronics and Telecommunications, Krakow

Abstract: The paper investigates the performance of recently proposed frequency estimation methods in the presence of typical power system measurements disturbances, i.e. sinusoids and additive noise. The tests are conducted for the polyphase Prony method, the polyphase smart DFT, and the discrete-time frequency-gain transducer. Considered methods are capable of unbiased one-cycle frequency estimation. The robustness against disturbances is obtained by the application of the proposed prefilter. Consider methods with the initially filtered measurement signal perform significantly better than the standard PMU (phasor measurement unit) P class and M class signal processing models.

Keywords: discrete Fourier transform, spectral filtration, smart DFT, Prony method, discrete-time frequency-gain transducer, least squares solution

ESTYMACJA CZĘSTOTLIWOŚCI W POMIARACH W SYSTEMIE ELEKTROENERGETYCZNYM

Streszczenie: W pracy przedstawiono zastosowanie nowoczesnych metod estymacji częstotliwości w przypadku obecności typowych zaburzeń pomiarowych występujących w systemie elektroenergetycznym, tj. sygnałów sinusoidalnych i szumu addytywnego. W badaniach uwzględniono polifazową metodę Prony'ego, polifazowe smart DFT oraz przetwornik częstotliwość-wzmocnienie z czasem dyskretnym. Powyższe metody umożliwiają nieobciążoną estymację częstotliwości na podstawie jednego okresu. Odporność na zakłócenia uzyskano dzięki zastosowaniu zaproponowanego filtra wstępnego. Badane metody działające na wstępnie przefiltrowanym sygnale wykazują znacznie mniejsze błędy pomiaru częstotliwości niż standardowe modele PMU (*phasor measurement unit*) w klasie P i w klasie M.

Słowa kluczowe: dyskretne przekształcenie Fouriera, filtracja częstotliwościowa, metoda Prony'ego, przetwornik częstotliwość-wzmocnienie, metoda najmniejszych kwadratów

1. Introduction

Estimation of a fundamental frequency is a basic problem in electric power system measurements. Many frequency estimation methods were designed so far but the problem is still under intense research. The design of frequency estimation method is difficult because it must not only be accurate, but also it must be robust against disturbances and computationally feasible. The most severe disturbances in electric power system are interfering sinusoids, i.e. harmonics, interharmonics, and subharmonics. Disturbances with frequency close to the system nominal frequency, i.e. the out-of-band interference, and the 2nd harmonic, are the most difficult to deal with. Also, the immunity to additive noise should be considered, although the levels of measured voltages and currents are high, and signal to noise ratio is moderate to high. Instantaneous frequency, amplitude and phase of a measured signal are time-varying, thus an estimation algorithm should have low latency, preferably several nominal cycles or less. It should also have a fast response to amplitude and phase steps. Above requirements are contradictory in many aspects, e.g. filtering out harmonic disturbance requires long impulse response that increases system latency. Available algorithms made different trade-offs, i.e. they sacrifice one property for another, e.g. in general a shorter time estimator is expected to be more “instantaneous” by the cost of less robustness against disturbances.

Different systematics can be applied for frequency estimation methods, and still some methods can belong to more than one category. In this writing we restrict our interest to the selected methods based on: the DFT (discrete Fourier transform), a FIR (finite impulse response) filter, and the LS (least squares) solution.

The DFT based frequency estimation is biased by the spectral leakage and the picket fence effect (Harris 1978) both occurring for noncoherent sampling. The impact of both phenomena can be mitigated on the stage of data acquisition by synchronous sampling, which is done by adjusting the sampling frequency to the signal frequency by the PLL (phase locked loop) (IEC 61000-4-7:2002), or by the software coherent resampling (Borkowski and Bień 2009). For non-coherently sampled signal the spectral leakage is reduced by time windows with high attenuation of the spectral side lobes, and the picket fence effect is overcome by the Interpolated DFT (IpDFT) algorithms. Application of a time window other than the rectangular one decreases frequency estimation bias by the cost of increasing estimation variance. Closed-form IpDFT formulas for a family of $\sin^\alpha(x)$ windows, that are generalization of the Rife-Vincent class I windows and are also referred to as the maximum sidelobe decay windows, are given in Duda and Barcentewicz (2014). For $\alpha = 0$ and $\alpha = 2$ the $\sin^\alpha(x)$ window is the rectangular window and the Hann window, respectively. For arbitrary window IpDFT formula can be derived by polynomial approximation (Duda 2011). IpDFT is also used for analysis of damped sinusoidal signals (Duda et al. 2011). The IpDFT algorithm proposed in Wang et al. (2021) is free of the spectral leakage, and enables frequency and damping estimation of a single sinusoid even from an observation shorter than one signal cycle.

The Smart DFT (SmDFT) (Yang and Liu 2000) uses the ratio of three DFT bins computed with one sample shift. The SmDFT is not biased by the spectral leakage nor

by the picket fence effect. In Duda and Zieliński (2021) it is shown that the downsampled (polyphase) SmDFT has higher noise immunity than the original SmDFT. Exemplary results of frequency estimation with the polyphase SmDFT are presented next.

FIR filter-based frequency estimation is used in the IEC/IEEE standard for synchrophasor measurements (IEC/IEEE 60255-118-1:2018). The phasor measurement unit (PMU) signal processing model contains quadrature demodulator working with the power system nominal frequency. The phasor is next obtained by lowpass filtration, and the frequency is computed as a time derivative of signal's phase. Applied lowpass filter determines measurement properties of the PMU, i.e., accuracy and disturbance immunity. Unfortunately, PMU with the standard M class lowpass filter is not compliant with the standard. Compliant PMU implementation can be obtained with the flat-top filters (Duda and Zieliński 2016, Duda et al. 2016).

Implementation of the Taylor–Fourier transform results in FIR filters for phasor, frequency and rate of change of frequency estimation (Platas-Garza and de la O Serna 2010, de la O Serna et al. 2021) with the advantages of negative image component and harmonics attenuation. The Taylor–Fourier transform basis contains not only phasor, as the DFT does, but also time derivatives of the dynamic phasor Taylor series expansion. The flat-top passband FIR Hilbert transformers, designed by the Taylor–Fourier transform, can also be designed with the flat-top filters (Duda et al. 2016) with the advantage of higher sidelobes attenuation (Duda et al. 2018).

FIR filter-based methods able for low latency estimation are the complex band-pass (CxBp) filters (Xu et al. 2021) and the discrete-time frequency-gain transducer (DTFGT) (Duda and Zieliński 2022a). Both, the CxBp filters and the DTFGT, can estimate frequency based on approximately one nominal cycle observation. Exemplary results of frequency estimation with the DTFGT are presented next.

The most popular frequency estimation method based on LS solution is the Prony method (Kay and Marple 1981, Zieliński and Duda 2011, Duda and Zieliński 2013). The original Prony method is sensitive to noise, however by introducing its polyphase implementation the noise immunity of the Prony method is significantly improved (Duda and Zieliński 2022b). Exemplary results of frequency estimation with the polyphase Prony method are presented next.

In this paper we investigate the properties of three low-latency unbiased frequency estimators namely: the polyphase SmDFT (Duda and Zieliński 2021), the DTFGT (Duda and Zieliński 2022a) and the polyphase Prony method (Duda and Zieliński 2022b). All above estimators work with approximately one nominal frequency period. The considered estimators are very close to the optimal one for sinusoidal signal disturbed only by additive noise. For sinusoidal interference their performance is improved by signal enhancing with a prefilter. As the harmonic interference is a very common disturbance, we propose the FIR prefilter with the impulse response being a self-convolution of the rectangular window, with length equal to the nominal period, and modulated to the system nominal frequency. This prefilter has zeros in frequency magnitude response at nominal harmonic frequencies. For reference the results are

compared with the standard PMU signal processing models in P class and M class (IEC/IEEE 60255-118-1:2018).

The paper is organized as follows. Section 2 defines the signal model. Section 3 describes investigated frequency estimation algorithms. Section 4 presents proposed prefilter for signal enhancing before frequency estimation. The results of frequency estimation for signal without disturbances, with additive noise and with sinusoidal interference are presented and discussed in Section 4. Section 5 concludes the paper.

2. Signal model

A continuous-time voltage or current signal in AC electric power grid is modelled by (IEC/IEEE 60255-118-1:2018):

$$x(t) = X_m(t) \cos[\Theta(t)] + D(t) \quad (1)$$

where: t is time in seconds, X_m is the peak magnitude of sinusoidal AC signal, $\Theta(t)$ is the angular position of the sinusoidal AC signal in radians, and $D(t)$ is a disturbance signal that contains additive contributions to the signal, including, but not limited to harmonics, noise, DC offset and out-of-band interference.

The angular position is the sum of the phase due to the nominal frequency f_0 in hertz, and instantaneous phase $\phi(t)$:

$$\Theta(t) = 2\pi f_0 t + \phi(t) \quad (2)$$

The complex value dynamic phasor of the signal (1) is given by:

$$X(t) = \frac{X_m(t)}{\sqrt{2}} e^{j\phi(t)} \quad (3)$$

Instantaneous frequency of the signal (1) is a time derivative of the cosine function argument:

$$f(t) = \frac{d\Theta(t)}{dt} = f_0 + \frac{1}{2\pi} \frac{d\phi(t)}{dt} \quad (4)$$

E.g. if the instantaneous phase is a linear function of time $\phi(t) = 2\pi\Delta f t + \phi_0$, then the instantaneous frequency deviates from the nominal frequency by Δf [Hz], i.e. $f(t) = f_0 + \Delta f$.

The parameters of the measurement signal (1), i.e., instantaneous frequency, amplitude and phase, are estimated based on its discrete-time representation obtained by periodic sampling with sampling frequency f_s :

$$x[n] = X_m[n] \cos\left(2\pi \frac{f_0}{f_s} n + \phi[n]\right) + D[n], \quad n = 0, 1, \dots, N-1 \quad (5)$$

where N is the number of signal samples. Nominal frequency ω_0 in radians, also referred to as pulsation or angular pulsation, is given by:

$$\omega_0 = 2\pi \frac{f_0}{f_s} = \frac{2\pi}{N_0}, \quad 0 < \omega_0 < \pi \quad (6)$$

where $N_0 = f_s/f_0$ is the number of samples in nominal period.

3. Considered frequency estimation methods

3.1. Polyphase Prony method

The polyphase Prony method was introduced for a single undamped sinusoidal signal in Duda and Zieliński (2021), and extended for a multifrequency damped sinusoidal signal in Duda and Zieliński (2022b). In comparison to the original Prony method the polyphase Prony method has significantly higher noise immunity. It exploits the LS solution of a set of equations build on autoregressive difference equation for sinusoidal signal. For an undamped single sinusoid defined as:

$$y[n] = A_0 \cos(\omega_0 n + \phi_0), \quad n = 0, 1, 2, \dots, N-1 \quad (7)$$

the current signal value $y[n]$ is determined by two previous values $y[n-L]$ and $y[n-2L]$:

$$y[n] = cy[n-L] - y[n-2L], \quad n = 2L, 2L+1, \dots, N-1 \quad (8)$$

where $c = 2\cos(\omega_0 L)$, and $L = 1, 2, \dots$, is a downsampling factor. It is shown in Duda and Zieliński (2021, 2022b) that the highest noise immunity is obtained for L selected such as to obtain 4 samples per signal cycle. For $L = 1$ we get the original Prony method. The frequency is computed as:

$$f[n] = \frac{f_s}{2\pi L} \cos^{-1}\left(\frac{c}{2}\right) \quad (9)$$

where the coefficient c is obtained by LS solution:

$$c = (\mathbf{E}^T \mathbf{E})^{-1} \mathbf{E}^T \mathbf{b} \quad (10)$$

where the vectors $\mathbf{E} = \begin{bmatrix} x_L \\ \vdots \\ x_{N-1-L} \end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix} x_0 + x_{2L} \\ \vdots \\ x_{N-1-2L} + x_{N-1} \end{bmatrix}$ are filled with the samples of the measured signal (5).

3.2. Polyphase Smart DFT

The SmDFT was introduced in Yang and Liu (2000). In Duda and Zieliński (2021) it was shown that by introducing downsampling factor L into the solution the noise immunity is significantly improved. The polyphase SmDFT is defined based on the autoregressive difference equation (8) and linearity of the DFT. The frequency is computed as:

$$f[n] = \frac{f_s}{2\pi L} \cos^{-1} \left(\frac{DFT_{k=1}\{b\}}{2DFT_{k=1}\{E\}} \right) \quad (11)$$

where L , $\mathbf{E}$, and $\mathbf{b}$ are the same as previously, and $DFT_{k=1}\{\cdot\}$ stands for the computation of the DFT bin with index $k = 1$ from the signal within the brackets, thus it is assumed that the observed signal spans approximately one signal cycle.

3.3. Discrete-time frequency-gain transducer

The DTGFT proposed in Duda and Zieliński (2022a) uses a FIR filter with a slope in frequency response. The frequency is computed based on the gain G of this filter:

$$f[n] = \frac{f_s}{2\pi L} \cos^{-1} (2G - 1) \quad (12)$$

where the gain is computed as:

$$G = \sqrt{\frac{h_{LP}[n]^* (h_0^L[n]^* x[n])^2}{h_{LP}[n]^* (x[n])^2}} \quad (13)$$

where the asterisk $*$ denotes a linear convolution. According to (13) measured signal $x[n]$ is firstly filtered by the slope filter $h_0^L[n]$:

$$h_0^L[n] = \begin{cases} h_0 \left[\frac{n}{L} \right], & n = 0, \pm L, \pm 2L \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

where $h_0[n] = [1/4, 1/2, 1/4]$, and next the filtered signal is squared and filtered by the lowpass filter $h_{LP}[n]$:

$$h_{LP}[n] = 1, \quad n = 0, \dots, N_{LP} - 1 \quad (15)$$

3.4. Standard PMU signal processing model

The standard PMU signal processing model (IEC/IEEE 60255-118-1:2018) exploits quadrature demodulator followed by the lowpass filter. In the P class a FIR filter with triangular impulse response is used, i.e., the convolution of two rectangular pulses each with the width equal to nominal cycle. This lowpass filter completely attenuates harmonics for nominal frequency. In the M class a FIR filter is designed by the “brick wall” method, also known as the window method (Oppenheim et al. 1999), with the Hamming window. The phasor is computed as the result of lowpass filtration after quadrature demodulation:

$$X[n] = (x[n]e^{-j\omega_0 n}) * h_{LP}[n] \quad (16)$$

where $h_{LP}[n]$ is an applied, P class or M class, lowpass filter. The frequency is next computed by approximation of time derivative:

$$f[n] = f_0 + \frac{f_s}{2\pi} \frac{\phi[n+1] - \phi[n-1]}{2} \quad (17)$$

where $\phi[n] = \text{angle}(X[n])$.

4. Prefilter for signal enhancement

The prefilter is used for the attenuation of the additive disturbance $D[n]$ in the measurement signal $x[n]$ (5). In the standard a lowpass filter is used for this purpose in the P class and in the M class PMUs. Frequency estimation methods under research, i.e. the polyphase Prony, the polyphase SmDFT, and the DTGFT do not apply quadrature demodulation, thus the prefilter is a bandpass filter around nominal frequency. For an even value of N_0 the prototype prefilter impulse response is defined as a rectangular pulse with the width equal to the period of nominal frequency ω_0 , modulated by a cosine with the nominal frequency:

$$h_0[n] = \frac{2}{N_0} \cos\left(\omega_0\left(n - \frac{N_0}{2} + 0,5\right)\right), \quad n = 0, \dots, N_0 - 1 \quad (18)$$

The self-convolved prefilter of order P is then obtained by linear convolution of the prefilter of order $P - 1$ and the prototype prefilter:

$$h_P[n] = h_{P-1}[n] * h_0[n], \quad P = 1, 2, \dots \quad (19)$$

Figure 1 depicts impulse responses and frequency magnitude responses for the prefilters of orders $P = 0, 1, \dots, 5$ for $N_0 = 16$. Due to iterated convolution (19) the length of the impulse response of the filter $h_p[n]$ elongates, in each iteration, by $N_0 - 1$ taps and equals $N = N_0(P + 1) - P$. The legend gives the length N of impulse response of each $h_p[n]$ filter, and the length of the filter expressed in the number of nominal cycles N/N_0 . It is observed that with the increase of order P the attenuation in the whole stopband, and not only at harmonic frequencies, significantly increases. Each $h_p[n]$ filter has a gain equal to 1 for the nominal 50 Hz frequency. With the frequency increase up to around 56.5 Hz the gain increases to a local maximum, e.g. the maximum gain equals to 1.03 for the $h_0[n]$ filter, and 1.22 for the $h_5[n]$ filter, this however has no effect on frequency estimation.

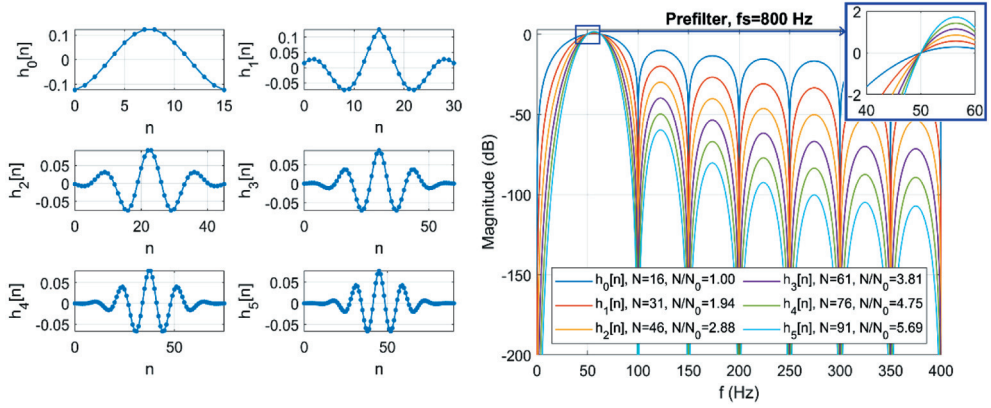


Fig. 1. Prefilter impulse responses (left) and magnitude responses (right) for orders $P = 0, 1, \dots, 5$

5. Results

The following sections present results of frequency estimation for a signal without disturbance, with additive zero-mean white Gaussian noise (WGN), sinusoidal interference, and multi-disturbance composed of both WGN and interfering sinusoid. The sampling frequency $f_s = 800$ Hz, the same as in the standard reference signal processing model (IEC/IEEE 60255-118-1:2018), is used in all simulations. The investigated frequency estimators are: the polyphase Prony method (denoted by Prony in all legends), the polyphase SmDFT (denoted by SmDFT), and the DTGFT. In all methods $L = 4$ is set, as this results in 4 samples per nominal period, which in turn guarantees the highest noise immunity in these methods. For each test the results are presented for: considered frequency estimators working alone, and considered frequency estimators working with the prefilter $h_0[n]$, with comparison to the standard P class PMU model

(IEC/IEEE 60255-118-1:2018), and working with the prefilter $h_7[n]$, with comparison to the standard M class PMU model (IEC/IEEE 60255-118-1:2018). The investigated estimators have the length of 1, 2, and 8.5 nominal cycles, the P class PMU model is 2 cycles estimator, and M class PMU model is 8.9 cycles estimator. The results are presented in terms of frequency error (FE) in hertz defined in (IEC/IEEE 60255-118-1:2018) as the difference between the measured frequency and the reference frequency. Further, the maximum absolute FE or the root-mean-square error (RMSE) of FE are presented.

5.1. Steady state frequency test

Steady state frequency compliance test is defined in (IEC/IEEE 60255-118-1:2018). In the P class PMU the range of fundamental frequency is 48–52 Hz, and in the M class PMU it is 45–55 Hz. In both classes the maximum absolute frequency error must not exceed 0.005 Hz. Obtained results are presented in Figure 2. It is observed that the methods under investigation, with and without the prefilter, are practically unbiased. The P class PMU and the M class PMU both are not compliant with its own standard (IEC/IEEE 60255-118-1:2018), as the error exceeds imposed limits.

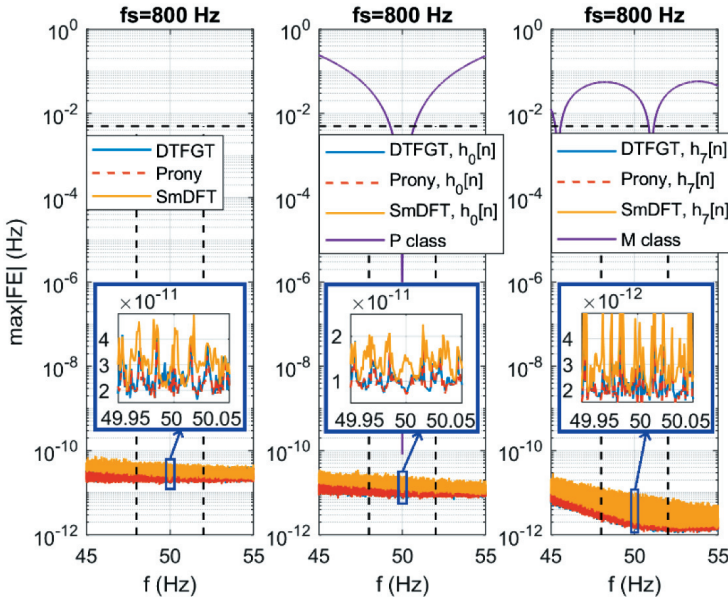


Fig. 2. Maximum absolute FE as a function of fundamental frequency: frequency estimators alone – 1 cycle estimation (left), frequency estimators with the prefilter $h_0[n]$ and the standard P class estimator – 2 cycles estimation (middle), and frequency estimators with the prefilter $h_7[n]$ – 8.5 cycles estimation and the standard M class estimator – 8.9 cycles estimation (right). Black dotted lines show the standard limits in the P class and the M class.

Errors for the DTFGT (blue) are overlapped by the Prony method (dashed red)

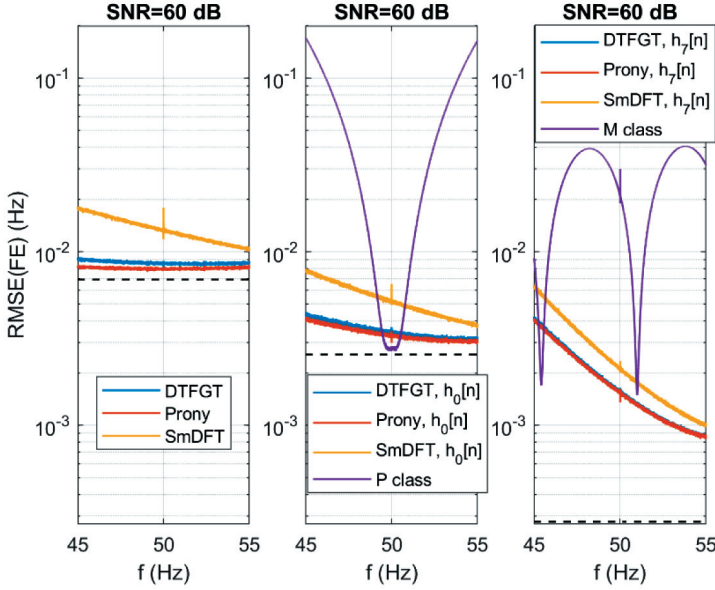


Fig. 3. RMSE of FE as a function of fundamental frequency for a signal with 60 dB SNR. Denotations as in Figure 2. Black dotted lines show the CRLB

5.2. Noise

Noise propagation through a frequency estimation algorithm is a fundamental property. Statistically efficient estimator is unbiased and its variance is determined by the Cramér–Rao lower bound (CRLB) (Kay 1993):

$$\text{var}(\omega) = \frac{12}{N(N^2 - 1)SNR}, \quad SNR = \frac{X_m^2}{2\sigma^2} \quad (20)$$

where ω is in radians as in (5) and (6), SNR stands for the signal-to-noise ratio and σ^2 is a variance of additive WGN. The variance of frequency estimation (20) is inversely proportional to the third power of the number of samples N thus noise immunity increases very fast with the number of samples. Figure 3 presents the RMSE of FE in comparison to the CRLB expressed as the standard deviation in hertz. It is observed that without the prefilter the DTFGT and the polyphase Prony are close to optimal frequency estimator, as they closely track the CRLB, however application of the prefilter deteriorates this optimality, and the longer the prefilter is the further the results are from the CRLB. The P class PMU is close to the CRLB only in a very narrow band of approximately 49.6–50.4 Hz, and the M class PMU is generally sensitive to additive noise.

5.3. Sinusoidal interference

In the standard (IEC/IEEE 60255-118-1:2018) sinusoidal interference is separately treated in the harmonic distortion test and the out-of-band interference test. In this study the frequency of interfering sinusoid is swept, with the step of 0.1 Hz, in the range from 75 to 400 Hz (i.e. the half of the sampling rate), what includes out-of-band interference and harmonics, but also interharmonics. Figure 4 depicts maximum absolute FE as a function of a 10% interfering sinusoid frequency for a signal with nominal 50 Hz frequency. From the investigated methods without a prefilter only the polyphase SmDFT rejects even harmonics except the 2nd one. Application of the rectangular prefilter $h_0[n]$ benefits in an ideal rejection of all nominal harmonics. By increasing the order of the prefilter also interharmonics are strongly attenuated. The P class PMU performs better than the polyphase SmDFT with $h_0[n]$ prefilter, but slightly worse than the DTFGT, and the polyphase Prony with $h_0[n]$ prefilter, and the M class PMU is the clear outsider.

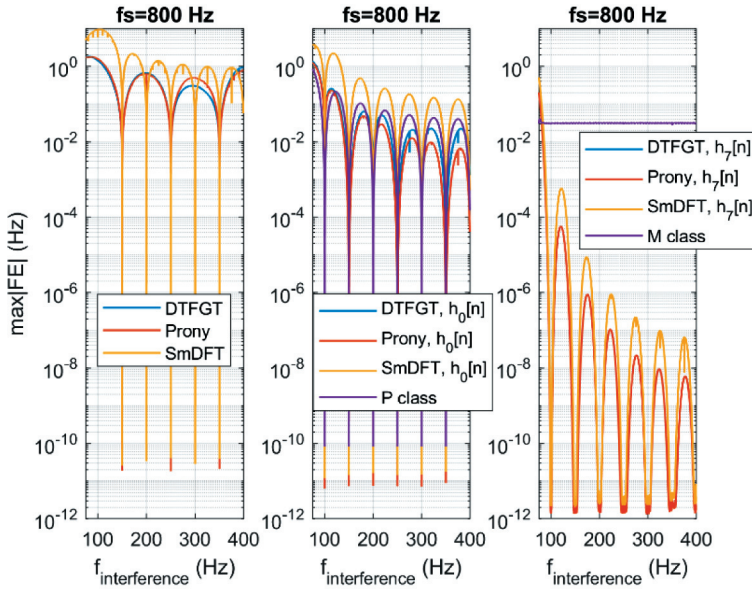


Fig. 4. Maximum absolute FE as a function of a 10% interfering sinusoid frequency for a signal with nominal 50 Hz frequency. Denotations as in Figure 2

5.4. Multi-disturbance

In practice the measurement signal is simultaneously affected by different disturbances. A frequency estimation algorithm often compromises sensitivity to some disturbances to obtain robustness against other disturbances.

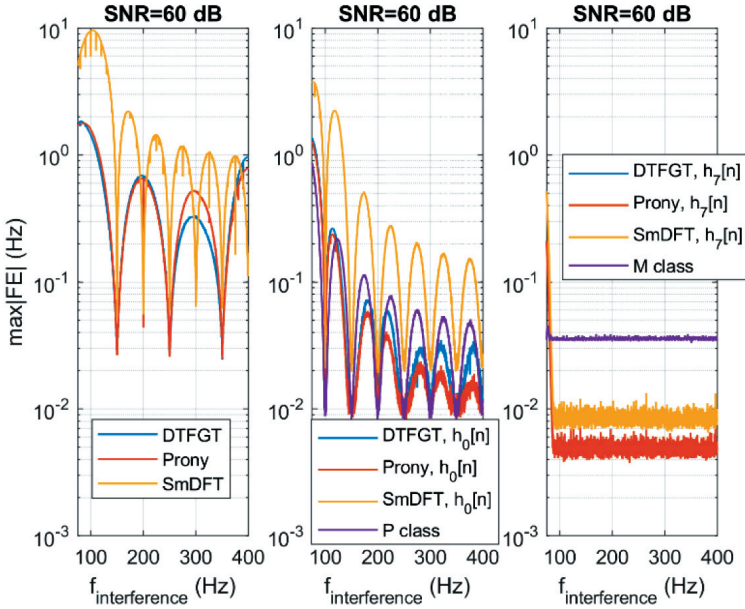


Fig. 5. Maximum absolute FE as a function of interfering sinusoid frequency for signal with 60 dB SNR. Denotations as in Figure 2.
Errors for the DTFGT (blue) are overlapped by the Prony method (red)

Figure 5 depicts maximum absolute FE for a signal simultaneously disturbed by interfering sinusoid and additive WGN. It is observed that the DTFGT and the polyphase Prony method are the best performers in this test. The polyphase SmDFT without the prefilter and with $h_0[n]$ prefilter performs worse. The M class PMU has the highest error except the range from 75 to 83 Hz where it has the smallest error on the level of 0.04 Hz.

6. Conclusions

The paper presents the performance of the following short-cycle frequency estimators: the polyphase Prony method, the polyphase SmDFT, and the DTGFT, in the presence of typical electric power system disturbances, i.e. a sinusoid and an additive noise. It is shown that these frequency estimators outperform PMU signal processing models defined in IEC/IEEE 60255-118-1:2018 and the robustness against disturbances in frequency estimation methods can be obtained by enhancing the signal with a prefilter. Thus the accuracy can be increased by the cost of higher latency, with the advantage that only signal preprocessing stage is changed and frequency estimation

part remains the same. IEC/IEEE 60255-118-1:2018 defines also dynamic testing not considered in this paper, but it is known from Duda and Zieliński (2022a), that e.g. the DTFGT can be used for fully compliant PMU implementation.

References

- Borkowski D., Biń A., 2009, *Improvement of accuracy of power system spectral analysis by coherent resampling*, IEEE Transactions on Power Delivery, vol. 24(3), pp. 1004–1013. <https://doi.org/10.1109/TPWRD.2009.2013662>.
- Duda K., 2011, *DFT interpolation algorithm for Kaiser-Bessel and Dolph-Chebyshev windows*, IEEE Transactions on Instrumentation and Measurement, vol. 60(3), pp. 784–790. <https://doi.org/10.1109/TIM.2010.2046594>.
- Duda K., Barcentewicz S., 2014, *Interpolated DFT for $\sin^a(x)$ windows*, IEEE Transactions on Instrumentation and Measurement, vol. 63(4), pp. 754–760. <https://doi.org/10.1109/TIM.2013.2285795>.
- Duda K., Zieliński T.P., 2013, *Efficacy of the frequency and damping estimation of real-value sinusoid*, Instrumentation & Measurement Magazine, pp. 48–58.
- Duda K., Zieliński T.P., 2016, *FIR filters compliant with the IEEE standard for M class PMU*, Metrology and Measurement Systems, vol. 23(4), pp. 623–636. <https://doi.org/10.1515/mms-2016-0055>.
- Duda K., Zieliński T.P., 2021, *Fast one-cycle frequency estimation of a single sinusoid in noise using downsampled linear prediction model*, Metrology and Measurement Systems, vol. 28(4), pp. 661–672. <https://doi.org/10.24425/mms.2021.137701>.
- Duda K., Zieliński T.P., 2022a, *P class and M class compliant PMU based on discrete-time frequency-gain transducer*, IEEE Transactions on Power Delivery, vol. 37(2), pp. 1058–1067. <https://doi.org/10.1109/TPWRD.2021.3076831>.
- Duda K., Zieliński T.P., 2022b, *The polyphase Prony method*, IEEE Signal Processing Magazine, vol. 39(3), pp. 115–120. <https://doi.org/10.1109/MSP.2022.3148712>.
- Duda K., Zieliński T.P., Magalas L.B., Majewski M., 2011, *DFT-based estimation of damped oscillation's parameters in low-frequency mechanical spectroscopy*, IEEE Transactions on Instrumentation and Measurement, vol. 60(11), pp. 3608–3618. <https://doi.org/10.1109/TIM.2011.2113124>.
- Duda K., Zieliński T.P., Barcentewicz S., 2016, *Perfectly flat-top and equiripple flat-top cosine windows*, IEEE Transactions on Instrumentation and Measurement, vol. 65(7), pp. 1558–1567. <https://doi.org/10.1109/TIM.2016.2534398>.
- Duda K., Turcza P., Zieliński T., Marszałek Z., 2018, *Flat-top bandpass FIR Hilbert transformers*, [in:] Machowski W., Stępień J. (eds.), *ICSES 2018: 2018 International Conference on Signals and Electronic Systems: September 10–12, 2018, Kraków, Poland: Proceedings*, IEEE, s. 143–146. <https://doi.org/10.1109/ICSES.2018.8507307>.

- Harris F.J., 1978, *On the use of windows for harmonic analysis with the discrete Fourier transform*, Proceedings of the IEEE, vol. 66(1), pp. 51–83. <https://doi.org/10.1109/PROC.1978.10837>.
- IEC 61000-4-7:2002, *Electromagnetic compatibility (EMC) – Part 4-7: Testing and measurement techniques – General guide on harmonics and interharmonics measurements and instrumentation, for power supply systems and equipment connected thereto*, 2nd ed.
- IEC/IEEE 60255-118-1:2018, *Synchrophasor for power systems – Measurements*, ed. 1.0.
- Kay S.M., 1993, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, Englewood Cliffs, NY.
- Kay S.M., Marple S.L., 1981, *Spectrum analysis – A modern perspective*, Proceedings of the IEEE, vol. 69(11), pp. 1380–1419. <https://doi.org/10.1109/PROC.1981.12184>.
- Oppenheim A.V., Schaffer R.W., Buck J.R., 1999, *Discrete-Time Signal Processing*, 2nd ed., Prentice-Hall, Englewood Cliffs, NY.
- O Serna J.A., de la, Paternina M.R.A., Zamora-Mendez A., 2021, *Assessing synchrophasor estimates of an event captured by a phasor measurement unit*, IEEE Transactions on Power Delivery, vol. 36(5), pp. 3109–3117. <https://doi.org/10.1109/TPWRD.2020.3033755>.
- Platas-Garza M.A., de la O Serna J.A., 2010, *Dynamic phasor and frequency estimates through maximally flat differentiators*, IEEE Transactions on Instrumentation and Measurement, vol. 59(7), pp. 1803–1811. <https://doi.org/10.1109/TIM.2009.2030921>.
- Wang K., Wen H., Xu L., Wang L., 2021, *Two points interpolated DFT algorithm for accurate estimation of damping factor and frequency*, IEEE Signal Processing Letters, vol. 28, pp. 499–502. <https://doi.org/10.1109/LSP.2021.3059364>.
- Xu S., Liu H., Bi T., 2021, *A novel frequency estimation method based on complex band-pass filters for P-class PMUs with short reporting latency*, IEEE Transactions on Power Delivery, vol. 36(6), pp. 3318–3328. <https://doi.org/10.1109/TPWRD.2020.3038703>.
- Yang J.Z., Liu C.W., 2000, *A precise calculation of power system frequency and phasor*, IEEE Transactions on Power Delivery, vol. 15(2), pp. 494–499. <https://doi.org/10.1109/61.852974>.
- Zieliński T.P., Duda K., 2011, *Frequency and damping estimation methods – an overview*, Metrology and Measurement Systems, vol. 18(4), pp. 505–528. <https://doi.org/10.2478/v10178-011-0051-y>.

Analiza anizotropii właściwości mechanicznych stopu magnezu typu WE43

Michał Dudziński, Michał Wałag , Anna Kula 

AGH Akademia Górniczo-Hutnicza, Wydział Metali Nieżelaznych, Kraków

Streszczenie: Artykuł przedstawia prace eksperymentalne dotyczące aspektów anizotropii właściwości mechanicznych stopu WE43 oraz technicznie czystego magnezu (99,95%) uzyskanych w procesie wyciskania na gorąco. Prace eksperymentalne obejmowały charakterystykę strukturalną i fazową materiałów w stanie wyjściowym oraz próby jednoosiowego ściskania prowadzone równolegle lub prostopadle do kierunku wyciskania. Komplementarnie przeprowadzono badania makrotekstury materiałów w stanie wyjściowym oraz po deformacji plastycznej. Na podstawie uzyskanych wyników określono stopień anizotropii właściwości mechanicznych badanych materiałów oraz wskazano na teksturę jako ważny czynnik warunkujący sposób deformacji plastycznej materiałów po wyciskaniu.

Słowa kluczowe: stop WE43, anizotropia, tekstura, próba ściskania, mikrostruktura

ANALYSIS OF THE MECHANICAL PROPERTIES ANISOTROPY IN WE43 MAGNESIUM ALLOY

Abstract: The article presents experimental work on the anisotropy aspects of the mechanical properties of WE43 alloy and technically pure magnesium (99.95%) obtained by hot extrusion. The experimental work includes structural and phase characterisation of as-extruded materials and uniaxial compression tests conducted parallel and perpendicular to the extrusion direction. Complementary studies of the macrotexture of materials before and after plastic deformation were carried out. Based on received results the degree of anisotropy of the mechanical properties of the tested materials was determined. The texture was identified as an important factor determining the activation of particular deformation mode of the materials after extrusion.

Keywords: WE43 alloy, anisotropy, texture, compression test, microstructure

https://doi.org/10.7494/978-83-67427-74-6_3

1. Wprowadzenie

Prężnie rozwijający się przemysł transportowy oraz lotniczy wymaga od materiałów stosowanych w tych gałęziach równie dynamicznego rozwoju. Od nowoczesnych materiałów konstrukcyjnych wymaga się zmniejszenia wagi przy zachowaniu tych samych bądź lepszych własności mechanicznych. Ponadto konsekwentne wprowadzanie idei gospodarki obiegu zamkniętego wymusza projektowanie materiałów łatwych do recyklingu. Dążenie do „odchudzenia” konstrukcji, głównie w przemyśle lotniczym i samochodowym, spowodowało zwrócenie uwagi inżynierów na stopy magnezu. Zastąpienie dotychczas stosowanych elementów aluminiowych magnezem, którego gęstość wynosi 1740 kg/m^3 , pozwala zredukować masę gotowego wyrobu o 30% (Bohlen i in. 2007). W takim świetle stopy na osnowie magnezu ukazują się jako przyszłościowe materiały konstrukcyjne. W tej grupie szczególną rolę odgrywają stopy magnezu z pierwiastkami ziem rzadkich (*rare earth elements*, REE). Zainteresowanie metalami ziem rzadkich jako dodatkami do stopów magnezu wynika głównie z ich korzystnego wpływu na właściwości wytrzymałościowe, plastyczność oraz teksturę (Kula 2019). Efektywność oddziaływania REE przejawia się głównie w zmodyfikowanej teksturze stopów magnezu, która znacząco ułatwia proces kształtowania wyrobów podczas przeróbki plastycznej (Dziadoń 2012). W procesach przeróbki plastycznej materiałów heksagonalnych dochodzi do powstawania bardzo silnej tekstury odkształceniowej. Wyróżniona orientacja powoduje anizotropię własności, co oznacza, że np. materiał stekstrowany będzie wykazywał różnice np. w granicy plastyczności w zależności od kierunku przyłożonej siły (Robson 2015). W niniejszej pracy podjęto próbę określenia stopnia anizotropii własności mechanicznych komercyjnego stopu WE43 oraz czystego magnezu (99,95%).

2. Charakterystyka materiału badawczego oraz metod badawczych

Materiał do badań stanowił stop magnezu WE43 o składzie chemicznym podanym w tabeli 1.

Tabela 1
Skład chemiczny materiału badanego [% wag.]

Y	Nd	Pr	Zn	Zr	Mg
4,38	2,84	0,5	<0,01	0,5	reszta

Stop został wyciśnięty współbieżnie w temperaturze 370°C z prędkością ruchu stempla 1 mm/s oraz stopieniem redukcji przekroju $\lambda \sim 14$. Z otrzymanego pręta wycięto próbki okrągłe o średnicy $d_o = 6 \text{ mm}$ i wysokości $h_o = 8 \text{ mm}$. Próbki z pręta

wycinano za pomocą elektrodrążarki drutowej, tak aby zminimalizować stopień deformacji materiału. Próbki wycinano w następujący sposób:

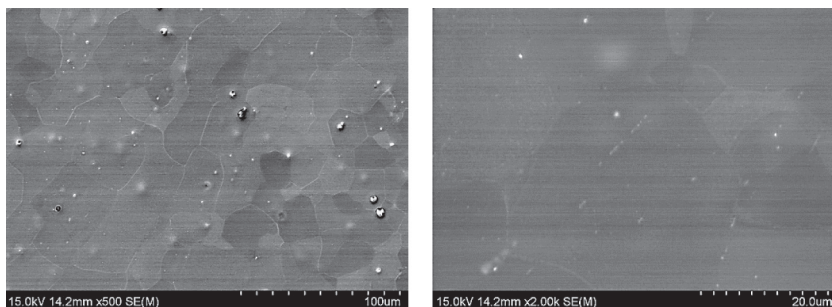
- oś próbki równoległa do kierunku wyciskania – **oznaczenie 0st**; w takim układzie podczas deformacji kierunek ściskania jest równoległy do kierunku wyciskania ($K\acute{S} \parallel ED$);
- oś próbki prostopadła do kierunku wyciskania – **oznaczenie 90st**; w takim układzie podczas deformacji kierunek ściskania jest prostopadły do kierunku wyciskania ($K\acute{S} \perp ED$).

Próbki odkształcano w warunkach jednoosiowego ściskania w temperaturze otoczenia na uniwersalnej maszynie wytrzymałościowej typu ZwickRoell z prędkością odkształcania $8 \cdot 10^{-3} \text{ s}^{-1}$. Próbki stopu WE43 poddano również badaniom mikrostruktury przy wykorzystaniu skaningowego mikroskopu elektronowego typu Hitachi SU-70. Próbki do badań przygotowano z wykorzystaniem klasycznych technik szlifowania i polerowania. Badania tekstury materiałów w stanie przed deformacją i po deformacji przeprowadzono za pomocą dyfraktometru rentgenowskiego typu Bruker D8 Advance/Discover operującego promieniowaniem typu Cu o długości fali $\lambda = 1.54 \text{ \AA}$. Dla badanych próbek rejestrowano niekompletne figury biegunowe, które w następnej kolejności przeliczano, wykorzystując program MTEX operujący w środowisku MATLAB. Dyfraktometr D8 Advance/Discover zastosowano również do rentgenowskiej analizy fazowej stopu WE43 w stanie wyjściowym po wyciskaniu. Dla celów porównawczych zestawowi podobnych badań poddano próbkę magnezu o czystości 99,95% uzyskanego w procesie wyciskania w temperaturze 300°C przy zachowaniu ruchu stempla 1 mm/s oraz $\lambda \sim 14$ tak jak w przypadku stopu WE43.

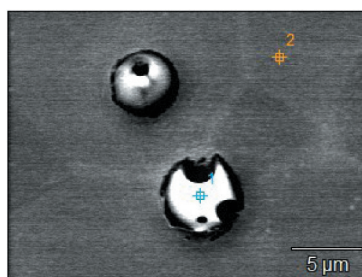
3. Wyniki badań

3.1. Wyniki obserwacji mikrostruktury materiału wyjściowego za pomocą skaningowego mikroskopu elektronowego

Na rysunku 1 pokazano mikrostrukturę czystego magnezu. Zdjęcia SEM potwierdzają, że magnez po wyciskaniu na gorąco wykazuje cechy zrekrystalizowanego materiału o charakterystycznych równoosiowych ziarnach. W mikrostrukturze wyróżniono drobne cząstki, wzbogacone w żelazo oraz krzem (rys. 2, tab. 2), stanowiące typowe zanieczyszczenie technicznie czystego magnezu. Na rysunku 3 przedstawiono charakterystyczne cechy mikrostruktury stopu WE43 obserwowane przy użyciu SEM. Z przeprowadzonych obserwacji wynika, że w mikrostrukturze stopu WE43 znajduje się duża ilość drobnych cząstek wzbogaconych w pierwiastki ziem rzadkich (rys. 4, tab. 3). Rentgenowska analiza fazowa wskazuje, że są to głównie cząstki faz międzymetalicznych typu Mg_{12}Nd oraz Mg_{24}Y (rys. 5).



Rys. 1. Mikrostruktura czystego magnezu uzyskana po procesie wyciskania na gorąco, obserwowana za pomocą skaningowego mikroskopu elektronowego

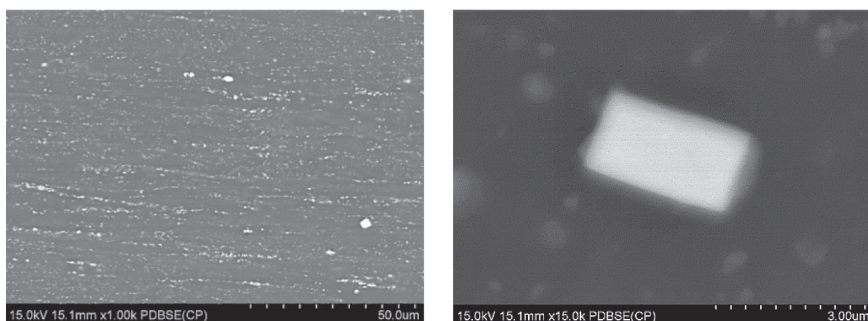


Rys. 2. Mikrostruktura magnezu wraz z zaznaczonymi punktami, dla których wykonano analizę EDS

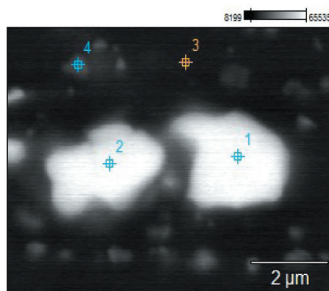
Tabela 2

Wyniki analizy składu chemicznego [% wag.] przeprowadzonej w punktach zaznaczonych na rysunku 2

	O	Mg	Si	Fe
Pkt 1	30,6	44,9	2,0	21,6
Pkt 2	—	100,0	—	—



Rys. 3. Mikrostruktura stopu WE43 uzyskana po procesie wyciskania na gorąco, obserwowana za pomocą skaningowego mikroskopu elektronowego

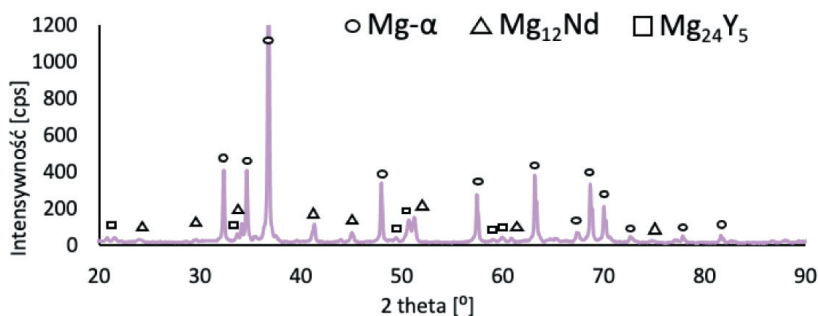


Rys. 4. Mikrostruktura stopu WE43 wraz z zaznaczonymi punktami, dla których wykonano analizę EDS

Tabela 3

Wyniki analizy składu chemicznego [% wag.] przeprowadzonej w punktach zaznaczonych na rysunku 4

	Mg	Y	Nd
Pkt 1	71,85	2,82	25,33
Pkt 2	72,56	2,99	24,46
Pkt 3	94,38	3,70	1,91
Pkt 4	90,06	4,63	5,31

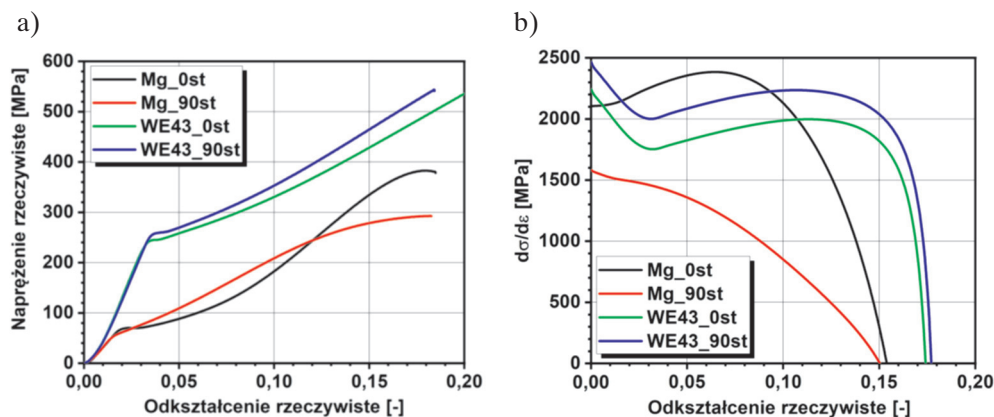


Rys. 5. Dyfraktogram rentgenowski stopu magnezu WE43 w stanie po wyciskaniu

3.2. Wyniki statycznej próby ściskania

Na rysunku 6a przedstawiono przebieg krzywych ściskania, uzyskanych dla próbki czystego magnezu oraz stopu WE43 odkształcanych w wariacie 0st i 90st. Z obserwacji przebiegu charakterystyk σ - ϵ wynika, że w przypadku stopu WE43, niezależnie od wariantu odkształcenia, krzywe wskazują na wyraźne przejście ze stanu sprężystego w stan plastyczny. Po przejściu granicy plastyczności na krzywych ściskania obserwuje

się bardzo krótkie „plateau”, po którym następuje praktycznie liniowy wzrost naprężenia, aż do pęknięcia. Dane eksperymentalne wskazują, że stop odkształcany w wariacie 90st wykazuje lekko wyższą granicę plastyczności.



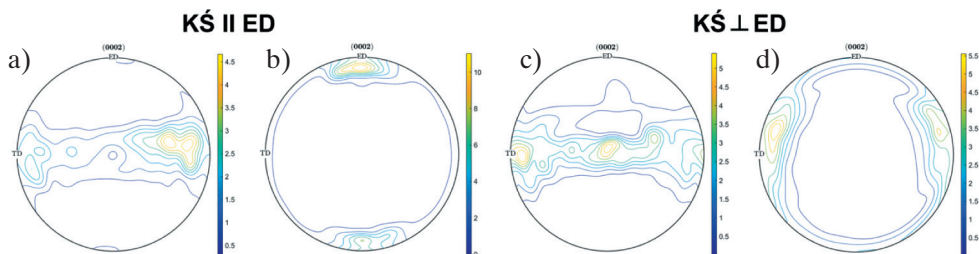
Rys. 6. Zestawienie krzywych: naprężenie rzeczywiste w funkcji odkształcenia rzeczywistego σ_{rz} vs. ϵ_{rz} (a) oraz prędkości umocnienia $d\sigma/d\epsilon$ w funkcji odkształcenia rzeczywistego ϵ_{rz} (b) dla próbek czystego magnezu i stopu WE43 odkształcanych w wariacie 0st oraz 90st

W przypadku czystego magnezu odkształcanego w wariacie 0st proces płynięcia plastycznego jest zdefiniowany przez wyraźniejszą granicę plastyczności niż w wariacie 90st. Charakterystyka ściskania magnezu odkształcanego w kierunku równoległym do kierunku wyciskania wykazuje sigmoidalny kształt z charakterystycznym zakresem „plateau”, odpowiadającym niskiej wartości umocnienia związanej z intensywnym bliźniakowaniem dominującym na tym etapie odkształcenia (rys. 6a). Krzywe prędkości umocnienia (rys. 6b) uzyskane w próbie ściskania stopu WE43 wykazują jakościowo podobny przebieg niezależnie od przeprowadzonego wariantu odkształcenia. Początkowy spadek $d\sigma/d\epsilon$ związany jest z zakresem sprężystego odkształcenia stopów. Po przekroczeniu granicy plastyczności krzywe prędkości umocnienia wykazują tendencję wzrostową aż do uzyskania pewnego maksimum, po którym następuje ciągły spadek krzywych prędkości umocnienia (rys. 6b). W przypadku czystego magnezu odkształcenie w wariacie 0st charakteryzuje ciągły wzrost krzywej prędkości umocnienia od momentu przekroczenia granicy plastyczności aż do maksimum ustalonego na poziomie około 0,07 odkształcenia rzeczywistego, po czym następuje ciągły spadek krzywej $d\sigma/d\epsilon$.

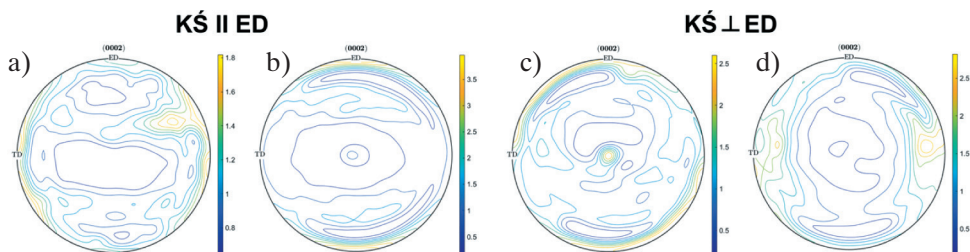
3.3. Wyniki badań tekstury

Na rysunku 7 przedstawiono teksturę czystego magnezu w stanie po wyciskaniu na gorąco oraz po maksymalnej deformacji plastycznej. Analiza figur biegunowych wskazuje, że magnez w stanie wyjściowym po wyciskaniu wykazuje silną teksturę bazalną.

Po maksymalnej deformacji w próbie jednoosiowego ściskania zaobserwowano silną reorientację ziarn w stosunku do wyjściowej tekstury (rys. 7). Taka sytuacja może mieć miejsce tylko wtedy, gdy w czasie deformacji dochodzi do mechanicznego bliźniakowania. Na rysunku 7d przedstawiono figurę biegunową czystego magnezu odkształconego w wariacie 90st, gdzie kierunek ściskania jest prostopadły do kierunku wyciskania ($K\acute{S} \perp ED$). Z badań tekstury wynika, że odkształcanie magnezu w konwencji 90st prowadzi do częściowej zmiany wyjściowej tekstury, zwłaszcza ziaren o orientacjach skoncentrowanych wokół kierunku TD. W tym przypadku następuje reorientacja ziarn w stronę kierunku wyciskania ED wywołana bliźniakowaniem. Ponadto po deformacji obserwuje się również pewną frakcję ziarn o orientacji zlokalizowanej wokół kierunku ND (rys. 8b). Na rysunkach 8a i c przedstawiono figurę biegunową płaszczyzny (0002) dla stopu WE43 w stanie wyjściowym po wyciskaniu na gorąco. W tym przypadku większość ziarn jest zorientowana osią c komórki heksagonalnej prostopadłe do kierunku wyciskania ED. Po deformacji plastycznej w wariacie 0st początkowa tekstura jest przetransformowana do nowej bliźniaczej orientacji, gdzie oś c komórki heksagonalnej jest równoległa do kierunku ściskania (rys. 8b).



Rys. 7. Figury biegunowe typu (0002) dla czystego magnezu w stanie wyjściowym (a, c) oraz po maksymalnej deformacji w procesie ściskania (b, d), gdzie kierunek ściskania jest równoległy do kierunku wyciskania ($K\acute{S} \parallel ED$ – wariant 0st) (b) oraz prostopadły do kierunku wyciskania ($K\acute{S} \perp ED$ – wariant 90st) (d)



Rys. 8. Figury biegunowe typu (0002) dla stopu WE43 w stanie wyjściowym (a, c) oraz po maksymalnej deformacji w procesie ściskania (b, d), gdzie kierunek ściskania jest równoległy do kierunku wyciskania ($K\acute{S} \parallel ED$ – wariant 0st) (b) oraz prostopadły do kierunku wyciskania ($K\acute{S} \perp ED$ – wariant 90st) (d)

Deformacja stopu w wariancie 90st prowadzi do reorientacji początkowej tekstury reprezentowanej przez orientację wokół kierunku TD. Transformacja tej orientacji związana jest z procesem bliźniakowania mechanicznego, a dokładniej – aktywacją tzw. bliźniaków rozciągania. Warto również zwrócić uwagę na to, że pewna frakcja ziarn po maksymalnym odkształceniu ściskającym wykazuje teksturę podobną do wyjściowej (rys. 8d).

4. Analiza wyników

Eksperyment badawczy wskazuje, że badany stop WE43 charakteryzuje się około pięciokrotnie większą granicą plastyczności w porównaniu z czystym magnezem. Wyznaczone wartości granicy plastyczności magnezu oraz stopu WE43 to odpowiednio ~ 55 MPa oraz ~ 255 MPa (rys. 6). Pozytywny wpływ dodatku pierwiastków ziem rzadkich odnotowano również, badając teksturę materiałów wyjściowych. Czysty magnez wykazuje bardzo silną bazalną teksturę, o znacznie większej intensywności niż stop WE43. Ponadto stop WE43 charakteryzuje tekstura o większym stopniu randomizacji bazalnych płaszczyzn w kierunku równoległym do kierunku wyciskania (rys. 8). Tekstura materiałów wyjściowych bardzo silnie wpływa na przebieg krzywych odkształcenia σ - ϵ oraz krzywych prędkości umocnienia $d\sigma/d\epsilon$ (rys. 6). Po procesie wyciskania badane materiały wykazują orientację, w której oś c komórki heksagonalnej jest prostopadła do kierunku wyciskania ED. Taka orientacja bardzo silnie warunkuje mechanizmy deformacji, które mogą być aktywowane podczas jednoosiowego ściskania. Gdy oś ściskania jest równoległa do kierunku wyciskania ($K\bar{S} \parallel ED$), wszystkie ziarna zorientowane osią c komórki heksagonalnej prostopadle do $K\bar{S}$ będą doświadczały odkształcenia spowodowanego bliźniakowaniem mechanicznym typu $\{10\bar{1}2\} < \bar{1}011 >$. Dokładnie tak dzieje się w przypadku odkształcania czystego magnezu oraz stopu WE43 w wariancie 0st. Przy tak zadanym kierunku ściskania dominującym systemem odkształcenia będzie bliźniakowanie, zwłaszcza w początkowym stadium deformacji.

O aktywacji bliźniakowania świadczy:

- kształt krzywych σ - ϵ (rys. 6) – zwłaszcza „plateau” w początkowym stadium odkształcenia;
- szybki wzrost krzywych prędkości umocnienia przy niskich odkształceniach spowodowany pojawieniem się granic bliźniaczych w podstrukturze badanych materiałów, które skutecznie blokują ruch dyslokacji, wpływając na wyższą efektywność akumulacji defektów;
- silna reorientacja ziarn, która może być wywołana tylko bliźniakowaniem.

Bliźniakowanie typu $\{10\bar{1}2\} < \bar{1}011 >$ opisuje kąt dezorientacji $86,3^\circ$ pomiędzy płaszczyzną bazalną osnowy i bliźniaka (Dziadoń 2012), stąd też transformacja większości ziarn po deformacji o około 90° , tak jak ma to miejsce w przypadku badanych materiałów w wariacie 0st, wywołana jest aktywnością bliźniakowania. Modyfikacja wyjściowej tekstury stopu WE43 przez jej osłabienie wpływa na bardziej zbalansowany udział bliźniakowania w globalnym procesie deformacji przez ściskanie. W przypadku wariantu odkształcenia 90st, tj. gdy oś ściskania jest prostopadła do kierunku wyciskania ($K\dot{S} \perp ED$), obserwuje się pewną frakcję ziarn zorientowanych osią c komórki heksagonalnej prostopadłe do $K\dot{S}$ oraz równoległe do $K\dot{S}$. Zatem w tym przypadku można oczekiwać, że frakcja ziarn zorientowana osią c komórki heksagonalnej prostopadłe do $K\dot{S}$ będzie preferować odkształcenie przez bliźniakowanie typu $\{10\bar{1}2\} < \bar{1}011 >$, zwłaszcza na początkowym etapie odkształcenia. W przypadku ziarn zorientowanych osią c komórki heksagonalnej równoległe do $K\dot{S}$ bliźniakowanie typu $\{10\bar{1}2\} < \bar{1}011 >$ będzie zablokowane ze względu na polarną naturę bliźniakowania – dla tej frakcji ziarn dominującym systemem odkształcenia będzie poślizg. Można zatem wnioskować, że odkształcenie plastyczne w wariacie 90st zależy od udziału objętościowego poszczególnych frakcji ziarn, co ma bezpośredni wpływ na:

- charakter krzywych σ - ϵ ,
- kinetykę umocnienia badanych materiałów i efektywność akumulacji defektów w podstrukturze,
- teksturę po odkształceniu.

Z makroskopowej analizy tekstury materiałów wyjściowych oraz przebiegu krzywych odkształcenia i prędkości umocnienia można wnioskować, że czysty magnez przy deformacji 90st wykazuje nieco większy udział objętościowy ziarn o orientacji, która doświadcza deformacji przez poślizg. Zmodyfikowana tekstura stopu WE43 i większy udział orientacji ziarn, w której płaszczyzna bazalna leży w płaszczyźnie TD (rys. 8), powoduje, że bliźniakowanie w wariacie 90st może być bardziej aktywne niż w przypadku czystego magnezu. W celu potwierdzenia powyższych spostrzeżeń konieczna jest dalsza i bardziej szczegółowa analiza tekstury i udziałów objętościowych ziarn o poszczególnych orientacjach. Analizując wyniki badań własnych, można zauważyć, że współczynnik asymetrii, który można tutaj zdefiniować jako stosunek granic plastyczności wyznaczonych w wariacie 0st i 90st, $\sigma_{y0st}/\sigma_{y90st}$, dla czystego magnezu wynosi 1,160, zaś dla stopu WE43 1,019. Można zatem wnioskować, że badane materiały wykazują niewielki stopień asymetrii granicy plastyczności. Wyraźniejsza anizotropia właściwości pojawia się przy większych odkształceniach, zwłaszcza w przypadku czystego magnezu (rys. 6).

3. Wnioski

Na podstawie przeprowadzonych badań można sformułować następujące wnioski:

1. Wprowadzenie pierwiastków ziem rzadkich do magnezu znacząco wpływa na zwiększenie jego wytrzymałości. W odniesieniu do granicy plastyczności σ_y zaobserwowano jej wzrost o około 200 MPa w stosunku do granicy plastyczności technicznie czystego magnezu. Zaobserwowano również pozytywny wpływ dodatku metali ziem rzadkich na teksturę materiałów wyjściowych.
2. Zarówno technicznie czysty magnez, jak i stop WE43 wykazują niską anizotropię granicy plastyczności podczas ściskania w badanych wariantach deformacji. Większa asymetria, zwłaszcza w przypadku czystego magnezu, pojawia się dopiero podczas zaawansowanego odkształcenia plastycznego.
3. Mechanizm odkształcenia plastycznego czystego magnezu oraz stopu WE43 podczas ściskania silnie zależy od tekstury badanych materiałów. W przypadku odkształcania badanych materiałów w wariancie, w którym oś ściskania jest równoległa do kierunku wyciskania ($K\bar{S} \parallel ED$), dominującym systemem deformacji, zwłaszcza na początkowym etapie odkształcenia, jest bliźniakowanie typu $\{10\bar{1}2\} < \bar{1}011 >$. Gdy oś ściskania jest prostopadła do kierunku wyciskania ($K\bar{S} \perp ED$), aktywowane mechanizmy odkształcenia są silnie uwarunkowane orientacją indywidulanych ziarn, dla których proces deformacji może być realizowany przez poślizg i/lub bliźniakowanie typu $\{10\bar{1}2\} < \bar{1}011 >$.

Literatura

- Bohlen J., Letzig D., Kainer K.U., 2007, *New perspectives for wrought magnesium alloys*, Materials Science Forum, vol. 546–549, s. 1–10. <https://doi.org/10.4028/www.scientific.net/MSF.546-549.1>.
- Dziadoń A., 2012, *Magnez i jego stopy*, Samodzielna Sekcja „Wydawnictwo Politechniki Świętokrzyskiej”, Kielce.
- Kula A., 2019, *Mechanizmy i parametry aktywacyjne odkształcenia plastycznego stopów Mg-Gd i Mg-Y zachodzącego w warunkach silnego ograniczenia łatwego poślizgu dyslokacji*, Wydawnictwo Naukowe Akapit, Kraków.
- Robson J.D., 2015, *Critical Assessment 9: Wrought magnesium alloys*, Materials Science and Technology, vol. 31(3), s. 257–264. <https://doi.org/10.1179/1743284714Y.00-00000683>.

Badanie bezpieczeństwa rurociągu ze względu na uderzenie hydrauliczne

Zbigniew Fąfara 

AGH Akademia Górniczo-Hutnicza, Wydział Wiertnictwa, Nafty i Gazu, Kraków

Streszczenie: W pracy przedstawiono wyniki badania wpływu uderzeń hydraulicznych na nagły wzrost ciśnienia wewnętrznego w rurociągu przesyłowym „Družba” transportującym ropę naftową. Przedmiotem badania jest fragment jednej z nitek rurociągu o długości 103,5 km. Podczas analizy wzięto pod uwagę dziewięć wybranych zdarzeń, w trakcie których została całkowicie zatrzymana jedna ze stacji pomp. Praktyka pokazuje, że skutki uderzeń hydraulicznych powstających w rurociągowych systemach transportujących media ciekłe są w ponad 80% przypadków przyczyną ich awarii. Wyniki bezpośrednich pomiarów zmiany ciśnienia wewnątrz rurociągu towarzyszącej całkowitemu zatrzymaniu przepływu ropy naftowej wskazują nawet na trzykrotny wzrost ciśnienia transportowanego medium na wejściu zatrzymanej stacji pomp, co jest szczególnie niebezpieczne i grozi rozszczelnieniem rurociągu „Družba”. Mimo ostatnich zmian na rynku dostaw ropy naftowej do Europy i ograniczenia jej importu z Rosji z powodu agresji na Ukrainę rurociąg magistralowy „Družba” stanowi obecnie ciągle podstawową instalację dostarczającą ropę naftową do Ukrainy, Słowacji, Czech i na Węgry, a w Polsce jego część jest wykorzystywana do przesyłu ropy do rafinerii znajdujących się we wschodniej części Niemiec. Rurociąg magistralowy „Družba” został wybudowany 60–70 lat temu i obecnie jego stan techniczny może pozostawiać wiele do życzenia, dlatego bezpieczeństwo jego użytkowania stanowi bardzo istotną kwestię.

Słowa kluczowe: rurociąg przesyłowy „Družba”, transport ropy naftowej, całkowite zatrzymanie przepływu medium, uderzenie hydrauliczne, wzrost ciśnienia wewnętrznego w rurociągu

PIPELINE SAFETY INVESTIGATION DUE TO HYDRAULIC HAMMER

Abstract: The paper presents the results of a study of the effect of hydraulic hammer on the sudden increase in internal pressure in the “Družba” trunk pipeline transporting crude oil. The subject of the investigation is a path of one section of the pipelines with a length of 103.5 km. During the analysis, nine selected events were taken into account, during which one of the pump stations was completely stopped. Practice shows that the effects of hydraulic hammer occurring in pipeline systems transporting liquid media are the cause of their failure in more than 80% of cases. The results of direct measurements of pressure changes inside the pipeline accompanying a complete stoppage of the flow of crude oil indicate an up to three-fold increase in the pressure of the transported medium at the entrance of the stopped pump station, which is particularly dangerous and may result in leakage of the “Družba” pipeline. Despite the recent changes in the market for crude oil supplies to Europe and the reduction of its import from Russia due to the aggression against Ukraine, the trunk pipeline “Družba” is currently still the primary installation supplying oil to Ukraine, Slovakia, the Czech Republic and Hungary, while in Poland its part is used to transport oil to refineries located in the eastern part of Germany. The main pipeline “Družba” was built 60–70 years ago and currently its technical condition may leave much to be desired, so its safety is a very important issue.

Keywords: transmission pipeline “Družba”, crude oil transport, complete stoppage of medium flow, hydraulic hammer, increase in internal pressure in the pipeline

1. Wprowadzenie

Każda zmiana hydrodynamicznych parametrów tłoczenia (w szczególności prędkości przepływu) cieczy w rurociągu prowadzi do zmian ciśnienia panującego w jego wnętrzu. Pojawiające się w efekcie stany nieustalone przepływu nazywane są uderzeniem hydraulicznym. Rozróżnia się dodatnie uderzenie hydrauliczne, następujące w momencie wzrostu ciśnienia wewnątrz rurociągu oraz ujemne uderzenie hydrauliczne – w momencie jego spadku. Pojęcie uderzenia hydraulicznego pojawiło się w literaturze w XIX wieku (Thorley 2004, Niełacny 2005, Mambretti 2014), z podziałem na proste i złożone uderzenie hydrauliczne. Proste uderzenie hydrauliczne następuje w sytuacji, gdy czas działania czynnika powodującego zmianę parametrów tłoczenia cieczy (np. czasu zamykania zaworu) jest dużo krótszy od czasu powrotu odbitej fali ciśnienia. Złożone uderzenie hydrauliczne dotyczy sytuacji przeciwnej, gdy czas działania czynnika powodującego zmianę parametrów tłoczenia cieczy jest dłuższy od czasu powrotu odbitej fali ciśnienia.

Do najczęstszych bezpośrednich czynników sprzyjających powstaniu uderzenia hydraulicznego należą: gwałtowne włączanie i wyłączanie układów pompowych, uruchamianie i awarie szybkostrzałnej armatury odcinającej oraz działanie systemów automatycznej regulacji przepływu. Duża liczba czynników inicjujących uderzenie hydrauliczne oraz oddziałujących na przebieg tego zjawiska utrudnia opracowanie jego modelu matematycznego. Uderzenie hydrauliczne jest zjawiskiem niepożądanym we wszystkich przewodach ciśnieniowych, ponieważ może wytworzyć falę uderzeniową o dużej wartości wzrostu ciśnienia wewnątrz przewodu oraz dużej częstotliwości, co sprzyja powstaniu awarii przewodu lub jego armatury oraz zaburza proces transportu cieczy. Dane literaturowe pokazują (Thorley 2004, Niełacny 2005, Kodura 2018), że ponad 80% awarii ciśnieniowych systemów wodociągowych w sieciach wywołuje zjawisko uderzenia hydraulicznego i jego niekorzystne oddziaływanie.

2. Model matematyczny

W praktyce inżynierskiej do badania uderzenia hydraulicznego wykorzystuje się jedną z najstarszych zależności dotyczących opisu tego zjawiska, opracowaną w XIX wieku, tzw. wzór Żukowskiego–Allieiego (Mambretti 2014, Thorley 2004, Kodura 2018). Równanie to określa pierwszy maksymalny przyrost ciśnienia w rurociągu w wyniku prostego uderzenia hydraulicznego:

$$\Delta p = \Delta v \cdot \rho \cdot a \quad (1)$$

gdzie:

- Δp – pierwszy maksymalny przyrost ciśnienia,
- Δv – zmiana (spadek) prędkości przepływającej cieczy,
- ρ – gęstość właściwa przepływającej cieczy,
- a – prędkość przemieszczania się fali uderzeniowej ciśnienia.

Maksymalną wartość przyrostu ciśnienia w rurowości uzyskuje się przy całkowitym zatrzymaniu przepływającej cieczy. Prędkość rozchodzenia się fali uderzeniowej ciśnienia w rurowości można określić na podstawie czasu powrotu odbitej fali ciśnienia do punktu rejestracji wzrostu ciśnienia lub czasu dojścia fali ciśnienia do innego punktu na trasie rurowości. W pierwszym przypadku równanie ma postać:

$$a = \frac{2 \cdot L}{T_R} \quad (2)$$

gdzie:

L – odległość miejsca powstania uderzenia hydraulicznego od przegrody odbijającej falę ciśnienia,

T_R – czas powrotu fali ciśnienia.

W drugim przypadku wzór na prędkość rozchodzenia się uderzeniowej fali ciśnienia jest następujący:

$$a = \frac{D}{T_A} \quad (3)$$

gdzie:

D – odległość punktu rejestracji ciśnienia od miejsca powstania uderzenia hydraulicznego,

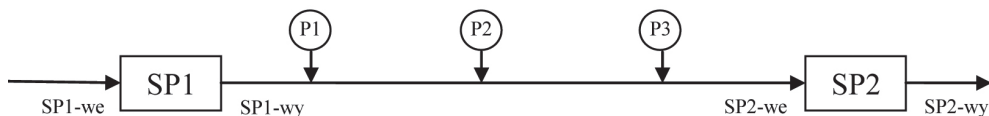
T_A – czas dojścia fali ciśnienia do punktu rejestracji.

3. Przedmiot badań

W trakcie budowy rurowości „Drużba” w latach 50. i 60. XX wieku rury przewodowe wykonywane były ze zwykłej stali budowlanej o minimalnej wytrzymałości na rozzerwanie $R_m = 200\text{--}350$ MPa (Michałowski i Trzop 2006). W tamtym czasie nie były jeszcze dostępne na rynku gatunki stali o podwyższonej wytrzymałości. Dopiero w latach 70. XX wieku zaczęto na skalę przemysłową rozwijać produkcję stali o zwiększonej wytrzymałości na potrzeby dynamicznie rozwijającego się działu związanego z budową rurowości. Na początku była to tzw. stal niskostopowa, a później bezstopowa (Michałowski i Trzop 2006, Magda 2018). W rezultacie tych prac obecnie na rynku są powszechnie dostępne gatunki stali o minimalnej wytrzymałości na rozzerwanie $R_m > 700$ MPa, a w laboratoriach pracuje się nad wytwarzaniem stali o $R_m > 1000$ MPa. Z powyższych względów w rurowości „Drużba” nie można stosować zbyt wysokich ciśnień tłoczenia ropy naftowej i generalnie maksymalne ciśnienie eksploatacyjne nie przekracza w nim wartości $p_{\text{exp,max}} = 4,7$ MPa. Dla przykładu w nowo budowanych rurowościach magistralowych lądowych niejednokrotnie wykorzystywane jest ciśnienie tłoczenia $p_{\text{exp,max}} > 10$ MPa, a w rurowościach magistralowych podmorskich – nawet $p_{\text{exp,max}} > 20$ MPa (Michałowski i Trzop 2006, Magda 2018).

Stabilność parametrów wytrzymałościowych stali w dłuższym horyzoncie czasowym sprawia, że zasadniczym czynnikiem decydującym o wielkości dopuszczalnego ciśnienia wewnątrz rurociągu jest grubość ścianki, która z czasem może ulec zmniejszeniu, najczęściej na skutek procesów korozji. Obecnie na zewnętrzne powierzchnie rur przewodowych nanoszone są wielowarstwowe (trójwarstwowe lub nawet czterowarstwowe) powłoki z tworzyw sztucznych w celu ochrony rur przed korozją. W latach 50. i 60. XX wieku takie technologie nie były jeszcze znane i powszechnie pokrywano rury przewodowe półpłynnymi masami bitumicznymi, zawijając je następnie spiralnie papą. Najstarsze odcinki rurociągu „Družba” mają już siedemdziesiąt lat, a nitki znajdujące się na obszarze Polski na pewno ukończyły już sześćdziesiąt lat. Z tego powodu należy się liczyć z możliwością obniżenia pierwotnej grubości ścianek rur przewodowych w wyniku procesów korozyjnych, a w konsekwencji – ze spadkiem wytrzymałości rur na rozerwanie. Do lat 80. XX wieku w Polsce wykorzystywano do wymiarowania rur przewodowych rurociągów przesyłowych dalekosiężnych radziecką metodę Łubińskiego, która zakładała bardzo duży margines bezpieczeństwa. W rezultacie tego maksymalne ciśnienie eksploatacyjne w rurociągu stanowiło około 40% minimalnego ciśnienia rozrywającego (Michałowski i Trzop 2006). Trzeba jednak pamiętać, że na przestrzeni kilkudziesięciu lat wykorzystywania rurociągu „Družba” grubość ścianek mogła ulec zmniejszeniu, co doprowadziłoby do obniżenia wartości minimalnego wewnętrznego ciśnienia rozrywającego. Z tych względów dla bezawaryjnej dalszej pracy rurociągu „Družba” bardzo istotne są wszystkie sytuacje, które mogą doprowadzić do niekontrolowanego wzrostu ciśnienia wewnątrz rurociągu powyżej ciśnienia eksploatacyjnego. Jedną z takich sytuacji, powszechnie występującą w rurociągach, jest uderzenie hydrauliczne – szczególnie niebezpieczne w przypadku całkowitego zatrzymania tłoczenia ropy naftowej ze względu na największy przyrost ciśnienia. Z tej przyczyny w niniejszej pracy dokonano analizy zmian ciśnienia w rurociągu „Družba” wywołanych pojawieniem się uderzenia hydraulicznego.

Przedmiotem analizy jest fragment rurociągu „Družba” o długości 103,5 km, który schematycznie został przedstawiony na rysunku 1. Jest on ograniczony dwoma kolejnymi stacjami pomp na trasie rurociągu. Analizowany fragment rurociągu zaczyna się stacją pomp SP1 i kończy stacją pomp SP2. W rurociągu pomiędzy stacjami pomp umieszczone są czujniki ciśnienia wewnętrznego oznaczone symbolami P1, P2 i P3. Czujniki ciśnienia znajdują się także na wejściu (SP1-we, SP2-we) i wyjściu (SP1-wy, SP2-wy) każdej stacji pomp. Czujniki te rejestrują w sposób ciągły, co sekundę, ciśnienie wewnątrz rurociągu. Lokalizacja punktów pomiaru ciśnienia na trasie rurociągu podana została w tabeli 1.



Rys. 1. Schemat rozważanego w analizie fragmentu rurociągu „Družba”

Tabela 1

Lokalizacja punktów pomiaru ciśnienia w analizowanym fragmencie rurociągu „Družba”

Nazwa punktu	Odległość punktu od początku fragmentu rurociągu [km]
SP1-wy	0,0
P1	31,5
P2	57,0
P3	71,5
SP2-we	103,5

Spośród wszystkich zdarzeń towarzyszących pracy analizowanego fragmentu rurociągu w jednym roku kalendarzowym wybrano dziewięć sytuacji opisanych w tabeli 2. Pięć z nich dotyczy kontrolowanego wyłączenia stacji pomp SP2 w celu redukcji ilości tłoczonej ropy naftowej ze względów technicznych (przepływ medium mógł być kontynuowany drugą niezależną nitką rurociągu). Pozostałe cztery sytuacje odnoszą się do awaryjnego wyłączenia stacji pomp SP2 na skutek przekroczenia maksymalnej dopuszczalnej temperatury w łożysku jednej z pracujących pomp (jeden przypadek) oraz niezależnego od operatora rurociągu odłączenia zasilania linii energetycznej (trzy przypadki). Wraz z wyłączeniem stacji pomp SP2 w sposób automatyczny zamykane były zasuwu na jej wejściu i wyjściu. W rozważanych sytuacjach wraz z zatrzymaniem stacji pomp SP2 niejednokrotnie wyłączanie były inne stacje pomp w tym samym czasie lub z opóźnieniem. Zamknięcie zasuw przy stacji pomp SP2 sprawiało jednak w takiej sytuacji, że żadne zaburzenie wywołane zdarzeniami poniżej SP2 nie oddziaływało na analizowany fragment rurociągu. Podobnie w każdym z dziewięciu rozważanych przypadków po wyłączeniu stacji pomp SP2 stacja SP1 pracowała jeszcze przez pewien czas i zaburzenie wywołane jej późniejszym zatrzymaniem nie mogło wpływać na gwałtowny wzrost ciśnienia wewnętrznego powstałego w miejscu wytworzenia uderzenia hydraulicznego (SP2-we).

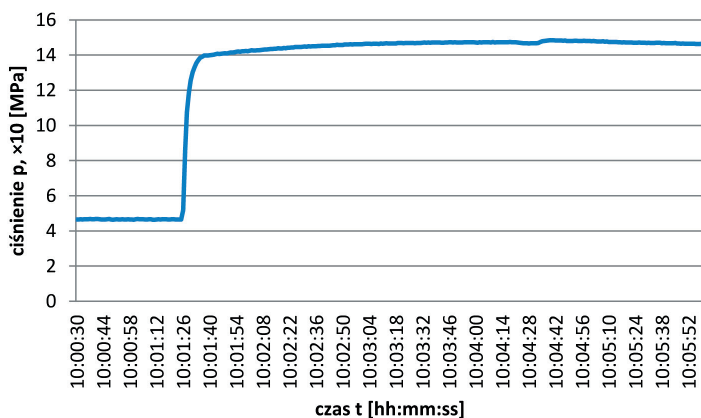
Tabela 2

Lista analizowanych zdarzeń w rozważanym fragmencie rurociągu „Družba”

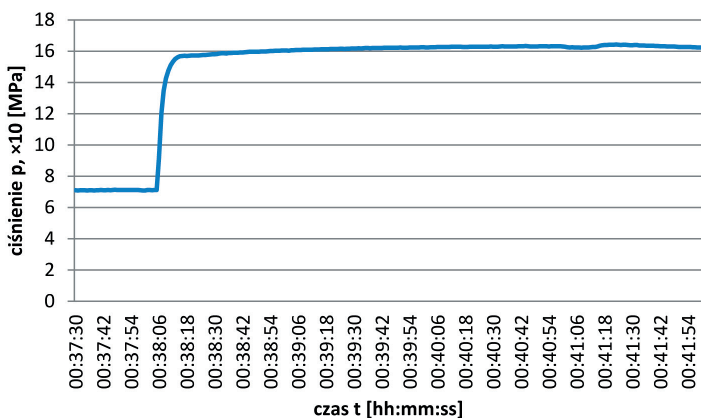
Lp.	Data	Zdarzenie
1	16.03	Zatrzymanie SP2 z przyczyn technicznych – redukcja ilości tłoczonej ropy naftowej
2	06.04	Zatrzymanie SP2 z przyczyn technicznych – redukcja ilości tłoczonej ropy naftowej
3	26.04	Awaryjne zatrzymanie SP2 – przekroczenie dopuszczalnej temperatury łożysk
4	07.06	Zatrzymanie SP2 z przyczyn technicznych – redukcja ilości tłoczonej ropy naftowej
5	20.06	Awaryjne zatrzymanie SP2 – przerwa w zasilaniu energią elektryczną
6	08.07	Zatrzymanie SP2 z przyczyn technicznych – redukcja ilości tłoczonej ropy naftowej
7	11.07	Awaryjne zatrzymanie SP2 – przerwa w zasilaniu energią elektryczną
8	26.07	Zatrzymanie SP2 z przyczyn technicznych – redukcja ilości tłoczonej ropy naftowej
9	23.08	Awaryjne zatrzymanie SP2 – przerwa w zasilaniu energią elektryczną

4. Wyniki pomiarów

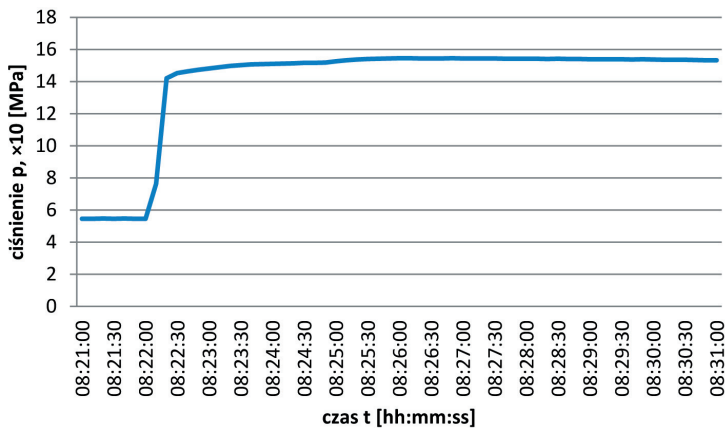
Na dziewięciu poniższych wykresach liniowych (rys. 2–10) przedstawiono zmierzono zmiany ciśnienia w czasie wewnątrz rurociągu na wejściu do stacji pomp SP2, spowodowane wyłączeniem stacji pomp. Rozważane w badaniach czujniki ciśnienia rejestrowały ciśnienie wewnętrzne w rurociągu w sposób ciągły (co sekundę) przez całą dobę. W prowadzonej analizie wykorzystano tylko fragment całego zarejestrowanego pliku obejmujący przedział czasu zazwyczaj o długości 4–5 min, rozpoczynający się około jednej minuty przed wyłączeniem stacji pomp SP2. Oś pozioma wykresów jest osią czasu rzeczywistego przedstawionego w standardowym formacie: godzina : minuta : sekunda. Na osi pionowej wykresów znajduje się ciśnienie wewnętrzne panujące w rurociągu, a jego wartości wyrażone w megapaskalach zostały pomnożone przez 10 dla zwiększenia czytelności zestawienia. Oznacza to, że wartości liczbowe ciśnienia na wykresach są wyrażone w barach.



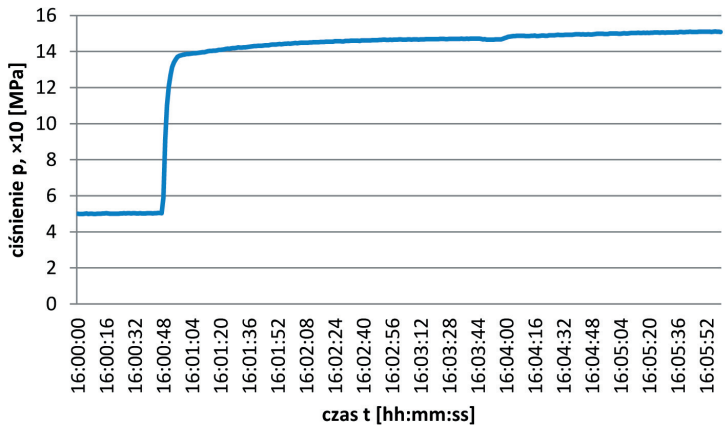
Rys. 2. Zmiany ciśnienia w punkcie SP2-we (pomiar 16 marca)



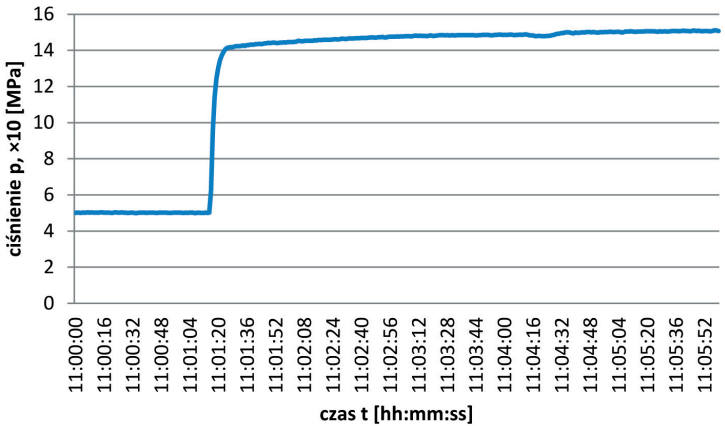
Rys. 3. Zmiany ciśnienia w punkcie SP2-we (pomiar 6 kwietnia)



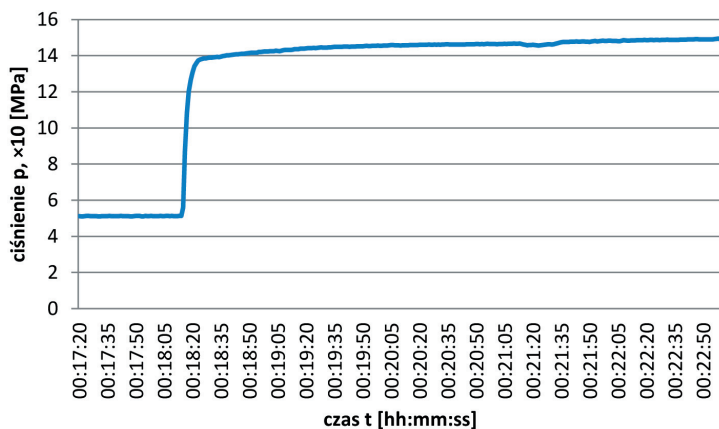
Rys. 4. Zmiany ciśnienia w punkcie SP2-we (pomiar 26 kwietnia)



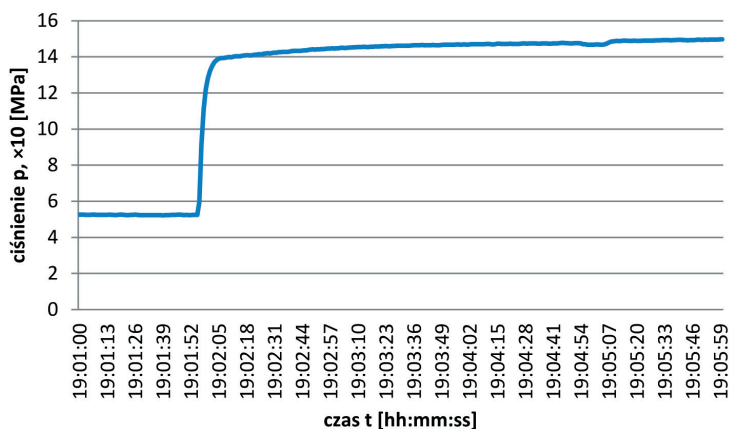
Rys. 5. Zmiany ciśnienia w punkcie SP2-we (pomiar 7 czerwca)



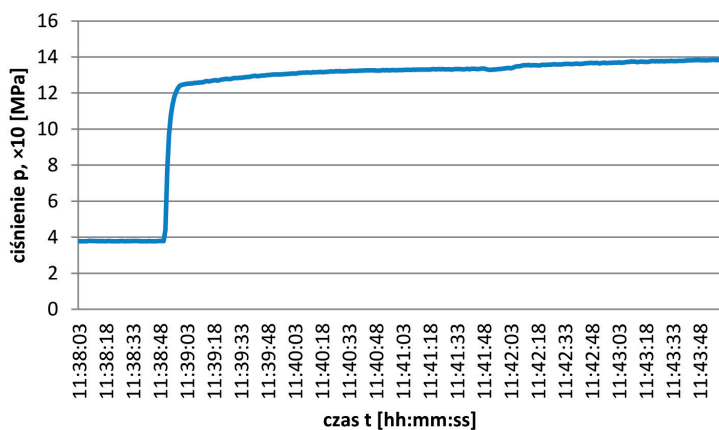
Rys. 6. Zmiany ciśnienia w punkcie SP2-we (pomiar 20 czerwca)



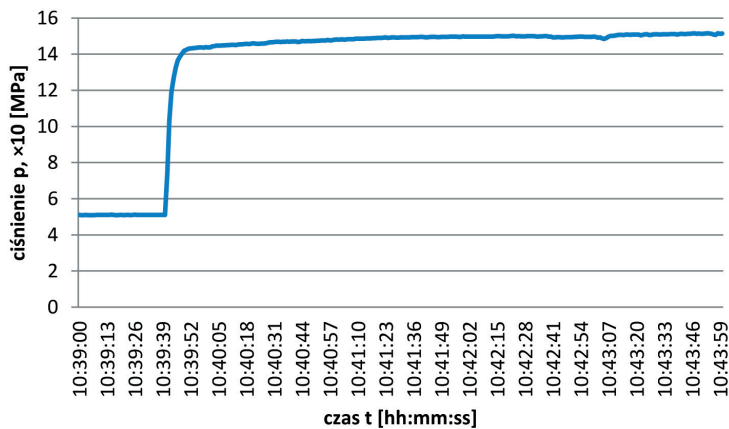
Rys. 7. Zmiany ciśnienia w punkcie SP2-we (pomiar 8 lipca)



Rys. 8. Zmiany ciśnienia w punkcie SP2-we (pomiar 11 lipca)



Rys. 9. Zmiany ciśnienia w punkcie SP2-we (pomiar 26 lipca)



Rys. 10. Zmiany ciśnienia w punkcie SP2-we (pomiar 23 sierpnia)

Przedstawione zależności czasowe ciśnienia wewnętrznego w rurociągu mają bardzo zbliżony przebieg. W pewnym momencie następuje wyłączenie stacji pomp SP2 i zaczyna się zamykać zasuwa, co powoduje gwałtowny wzrost ciśnienia na wejściu stacji pomp aż do chwili całkowitego jej zamknięcia. Czas zamykania zasuwy wynosi nieco ponad 10 s. Po tym czasie ciśnienie wewnątrz rurociągu nadal rośnie, ale dużo wolniej, aż do osiągnięcia ciśnienia maksymalnego, co następuje w rozważanych przypadkach po 2–3 min. W dalszej perspektywie czasowej ciśnienie wewnętrzne będzie nieznacznie fluktuować na skutek powrotu odbitej fali ciśnienia. W rozważanym fragmencie ropociągu „Družba” fala wzrostu ciśnienia wywołanego uderzeniem hydraulicznym odbija się od zamkniętej zasuwy znajdującej się na wyjściu stacji pomp SP1, która oddalona jest o 103,5 km. Pokonanie takiego odcinka tam i z powrotem zajmuje fali ciśnienia około 200 s. Moment dojścia odbitej fali ciśnienia widoczny jest na zamieszczonych wykresach jako niewielkie wahania ciśnienia w końcowym przebiegu krzywej. W tabeli 3 zamieszczono zbiorcze zestawienie wartości gwałtownego wzrostu ciśnienia na wejściu stacji pomp SP2 spowodowanego pojawiającym się uderzeniem hydraulicznym na skutek jej wyłączenia.

Tabela 3
Charakterystyka analizowanych uderzeń hydraulicznych

Data	16.03	6.04	26.04	7.06	20.06	8.07	11.07	26.07	23.08
Zatrzymanie stacji pomp [hh:mm:ss]	10:01:25	00:38:05	08:22:00	16:00:47	11:01:15	00:18:14	19:01:55	11:38:50	10:39:40
Ciśnienie początkowe ×10 [MPa]	4,653	7,116	5,461	5,029	5,014	5,127	5,249	3,794	5,108

Tabela 3 cd.

Zamknięcie całkowite zasuw [hh:mm:ss]	10:01:35	00:38:16	08:22:10	16:01:01	11:01:27	00:18:24	19:02:07	11:39:03	10:39:52
Czas zamykania zasuw [s]	10	11	10	14	12	10	12	13	12
Wzrost ciśnienia $\times 10$ [MPa]	13,853	15,686	14,201	13,855	14,167	13,793	13,933	12,515	14,320
Przyrost ciśnienia $\times 10$ [MPa]	9,200 (198%)	8,570 (120%)	8,740 (160%)	8,826 (176%)	9,153 (183%)	8,666 (169%)	8,684 (165%)	8,721 (230%)	9,212 (180%)
Maksymalny wzrost ciśnienia [hh:mm:ss]	10:04:00	00:40:24	08:25:30	16:03:40	11:03:58	00:20:48	19:04:13	11:41:27	10:42:29
Czas osiągnięcia maksimum [s]	155	139	210	173	163	154	138	157	169
Maksymalne ciśnienie $\times 10$ [MPa]	14,714	16,289	15,404	14,709	14,865	14,629	14,688	13,319	15,002
Maksymalny przyrost ciśnienia $\times 10$ [MPa]	10,061 (216%)	9,173 (129%)	9,943 (182%)	9,680 (192%)	9,851 (196%)	9,502 (185%)	9,439 (180%)	9,525 (251%)	9,894 (194%)

5. Analiza wyników

Ciśnienie początkowe w rurociągu na wejściu do stacji pomp SP2 zmienia się od wartości 0,38 MPa do 0,71 MPa. Jest to spowodowane wybranym przez operatora reżimem pracy systemu rurociągowego, który określa wykorzystywaną moc stacji pomp systemu, zależną między innymi od wymaganego w danym momencie wydatku przepływu ropy naftowej. Istotnym czynnikiem różnicującym ciśnienie są także pojawiające się opory przepływu, uzależnione między innymi od gęstości właściwej ropy naftowej oraz jej lepkości, zależącej w dużym stopniu od temperatury ropy, na którą wpływa temperatura na zewnątrz rurociągu uwarunkowana porą dnia, roku i panującą pogodą.

Czas zamykania zasuw wyznaczony na podstawie zebranych danych zmienia się w zakresie 10–14 s. Określenie czasu rozpoczęcia zamykania zasuw jest stosunkowo proste, ponieważ następuje wtedy gwałtowny wzrost ciśnienia w rurociągu, natomiast uchwycenie momentu całkowitego zamknięcia zasuw jest już kłopotliwe, gdyż punkt zmiany trendu krzywej z gwałtownego wzrostu wartości ciśnienia w dużo wolniejszy jego przyrost nie zawsze jest wyraźnie zaznaczony na wykresie. Okazuje się, że częstotliwość rejestracji ciśnienia jest zbyt mała, co może prowadzić do nieznacznej rozbieżności w określonych wartościach czasu zamykania zasuw dla dziewięciu rozważanych przypadków.

Na skutek całkowitego zamknięcia zasuw ciśnienie wewnątrz rurociągu na wejściu do stacji pomp SP2 wzrasta do wartości 1,25–1,57 MPa. Odpowiada to przyrostowi ciśnienia wewnętrznego w rurociągu o 0,86–0,92 MPa, co oznacza wzrost o 120–230% ciśnienia początkowego. Średni przyrost ciśnienia wynosi około 180% wartości początkowej, czyli obserwuje się prawie trzykrotny wzrost ciśnienia wewnątrz rurociągu. Wartości wzrostu ciśnienia po czasie zamknięcia zasuw różnią się nieco w dziewięciu analizowanych przypadkach, ale generalnie są do siebie zbliżone. Ze wzoru Żukowskiego–Allieiego (1) wynika, że na wartość wzrostu ciśnienia wpływa początkowa liniowa prędkość przepływu ropy naftowej (końcowa jest w każdym przypadku równa zero), gęstość właściwa ropy oraz prędkość przemieszczania się fali uderzeniowej ciśnienia w ropie. We fragmencie rurociągu „Drużba” poddanym badaniom liniowa prędkość przepływu ropy naftowej jest dobierana na poziomie około $v = 1$ m/s w zależności od wymaganego wydatku przepływu. Wynika stąd, że w każdym z rozważanych przypadków prędkość początkowa przepływu ropy mogła być nieco inna. Gęstość właściwa transportowanej rurociągiem ropy naftowej nieznacznie się różniła i wynosiła $\rho = 867\text{--}876$ kg/m³. Podobnie prędkość rozchodzenia się fali uderzeniowej ciśnienia mogła się różnić ze względu na różną gęstość właściwą i temperaturę ropy naftowej. Omówione czynniki tłumaczą różnice w wartościach uzyskanych przyrostów ciśnienia wewnątrz rurociągu.

Czas osiągnięcia maksymalnego przyrostu ciśnienia, czyli pełnego uderzenia hydraulicznego, zmieniał się w zakresie 138–210 s i średnio wyniósł około 150 s. Występujące różnice w wartościach wynikały częściowo z niedokładności odczytu spowodowanych bardzo małymi zmianami ciśnienia w tym okresie oraz nałożeniem się odbitej fali powrotnej ciśnienia.

Osiągnięte maksymalne wartości ciśnienia na wejściu stacji pomp SP2 (czyli wartości uderzenia hydraulicznego) zmieniają się w zakresie 1,33–1,63 MPa, co odpowiada przyrostowi ciśnienia 0,92–1,01 MPa (stanowi to 129–251% wartości początkowej, średnio 200%). Oznacza to, że uderzenie hydrauliczne spowodowane całkowitym zatrzymaniem przepływu ropy naftowej w rurociągu powoduje średnio trzykrotny wzrost ciśnienia na wejściu do stacji pomp SP2. Wyjaśnienie różnic w wartościach wzrostu ciśnienia przy użyciu równania Żukowskiego–Allieiego (1) jest analogiczne jak podane wcześniej w odniesieniu do przyrostu ciśnienia po całkowitym zamknięciu zasuw.

6. Podsumowanie

Magistralowy rurociąg dalekosiężny ropy naftowej „Družba” jest nadal podstawowym kanałem importu tego surowca do części krajów Europy, mimo wprowadzonych ograniczeń w imporcie ropy z Rosji na skutek agresji na Ukrainę. Różne jego odcinki zostały zbudowane 60–70 lat temu, dlatego można mieć obecnie obawy dotyczące stanu technicznego rurociągu i związanego z tym bezpieczeństwa jego pracy. Ropociąg „Družba” pracuje nadal bezawaryjnie w warunkach ustawionych standardowych reżimów pracy, jednak każde odchylenie ciśnienia eksploatacyjnego od wartości typowych może stwarzać zagrożenie dla jego bezpieczeństwa.

W pracy poddano analizie dziewięć zdarzeń z jednego roku polegających na kontrolowanym lub niekontrolowanym wyłączeniu jednej ze stacji pomp tłoczących ropę naftową w wybranym fragmencie rurociągu. Wyłączenie stacji pomp skutkowało automatycznym zamknięciem zasuw na wejściu i wyjściu stacji w celu ich ochrony. Całkowite zablokowanie przepływu rozpędzonej ropy naftowej powodowało gwałtowny wzrost ciśnienia wewnątrz rurociągu na wejściu do stacji pomp, które to zjawisko określa się mianem uderzenia hydraulicznego. Wykonane pomiary ciśnienia czujnikami zamontowanymi wewnątrz rurociągu, prowadzącymi automatyczną jego rejestrację co sekundę, wykazały, że w poszczególnych dziewięciu rozważanych przypadkach w czasie 10–14 s od zatrzymania stacji pomp ciśnienie wewnątrz rurociągu na wejściu stacji pomp wzrosło średnio ponad 2,5 razy, natomiast po osiągnięciu maksymalnej wartości w czasie 2–3 min było średnio ponad trzy razy wyższe od ciśnienia początkowego sprzed zdarzenia. Tak gwałtowny przyrost ciśnienia w bardzo krótkim czasie stanowi duże odchylenie od wartości ustalonych w standardowym reżimie pracy systemu rurociągowego, stwarzając niebezpieczeństwo jego rozszczelnienia.

Literatura

- Kodura A., 2018, *Wpływ długości przewodu za zaworem kulowym na parametry uderzenia hydraulicznego z polietylenu o dużej gęstości*, Ochrona Środowiska, vol. 40(4), s. 15–20.
- Magda W., 2018, *Rurociągi podmorskie: zasady projektowania*, Wydawnictwa Naukowo-Techniczne, Warszawa.
- Mambretti S., 2014, *Water Hammer Simulations*, WIT Press, Southampton.
- Michałowski W.S., Trzop S., 2006, *Rurociągi dalekiego zasięgu*, Wydawnictwo Fundacja Odyseum, Warszawa.
- Niełacny M., 2005, *Uderzenia hydrauliczne w systemach wodociągowych*, Wydawnictwo Politechniki Poznańskiej, Poznań.
- Thorley A.R.D., 2004, *Fluid Transients in Pipeline Systems: A Guide to the Control and Suppressions of Fluid Transients in Liquids in Closed Conduits*, ASME Press, New York.

Aerial SLAM algorithm with global alignment using satellite imagery

Jakub Karbowski¹ , Bartosz Bartoszewski² , Daria Kokot³ ,
Michał Jan Kwiecień⁴ , Tymoteusz Turlej⁴ 

¹ AGH University of Krakow, Faculty of Computer Science, Krakow

² AGH University of Krakow, Faculty of Electrical Engineering, Automatics, Computer Science and Biomedical Engineering, Krakow

³ AGH University of Krakow, Faculty of Physics and Applied Computer Science, Krakow

⁴ AGH University of Krakow, Faculty of Mechanical Engineering and Robotics, Krakow

Abstract: Simultaneous localization and mapping (SLAM) algorithms can be used in aerial vehicles with a down-facing camera for navigation and terrain mapping. However, they suffer from an incremental position error. Each neighbouring image pair match introduces an alignment error which accumulates to a global position error proportional to the map size. By using a reference satellite image map this error can be reduced to a constant factor. This work proposes a novel image alignment algorithm called Embed Match which can be applied to any SLAM algorithm. Embed Match uses a neural network to represent a map region as a field of embedding vectors which are a semantic representation of the area. By comparing the embedding fields of a satellite image and the camera image, an alignment function is defined. Maximising this function leads to image alignment. The proposed algorithm shows a radically different approach from keypoint-based alignment methods which usually struggle with matching drastically different images, such as satellite and real camera images.

Keywords: SLAM, aerial navigation, mapping, neural network

ALGORYTM AERIAL SLAM Z GLOBALNYM USTAWIENIEM PRZY UŻYCIU ZDJĘĆ SATELITARNYCH

Streszczenie: Algorytmy symultanicznej lokalizacji i mapowania (SLAM) mogą być stosowane w pojazdach powietrznych z kamerą skierowaną w dół do nawigacji i mapowania terenu. Jednakże występuje błąd inkrementacyjny pozycji. Każde dopasowanie sąsiedniej pary obrazów wprowadza błąd wyrównania, który wpływa na błąd globalnej pozycji proporcjonalnie do rozmiaru mapy. Korzystając z referencyjnej mapy obrazów satelitarnych, ten błąd można zredukować do stałego współczynnika. W tej pracy proponowany jest nowatorski algorytm wyrównywania obrazów o nazwie Embed Match, który może być stosowany do dowolnego algorytmu SLAM. Embed Match wykorzystuje sieć neuronową do reprezentowania regionu mapy jako pola wektorów osadzania, które są semantyczną reprezentacją obszaru. Porównując pola osadzania obrazu satelitarnego i obrazu kamery, definiuje się funkcję wyrównywania. Maksymalizacja tej funkcji prowadzi do wyrównywania obrazu. Proponowany algorytm stanowi radykalnie inne podejście niż metody wyrównywania oparte na punktach charakterystycznych, które zazwyczaj mają trudności z dopasowywaniem się do drastycznie różnych obrazów, takich jak obrazy satelitarne i rzeczywiste obrazy z kamery.

Słowa kluczowe: SLAM, nawigacja satelitarna, mapowanie, sieć neuronowa

1. Introduction

SLAM algorithms play a vital role in the field of aerial robotics by providing real-time localization and constructing maps of the environment using internal sensor data (Gupta and Fernando 2022). SLAM algorithms not using external reference information such as pre-loaded maps exhibit an incremental mapping error, which accumulates over time and adversely affects the accuracy of the constructed maps. This accumulated error becomes particularly critical in scenarios involving extensive map coverage, long-distance navigation, and tasks reliant on global spatial relationships.

Most advances in SLAM algorithms are related to the reduction of this error. SLAM algorithms can utilise many different sensor types. Of particular interest are visual SLAM algorithms, which can be divided into visual-only, visual-inertial and visual-depth systems (Macario Barros et al. 2022).

SLAM algorithms incrementally build and refine a map of the environment. When new sensor data such as a camera image is collected, it is used to update the map. First, an estimate of the robot's position is used to roughly align new sensor data with the current map. Next, the new data is locally aligned with the nearest map features. Finally, a global alignment procedure takes into account features of the entire map created up to this point in time to perform long distance alignment and loop closing. Without the global alignment step, each local alignment of new sensor data would introduce an incremental error to the map. Loop closing helps reduce this error, as long as the robot visits the same position many times. However, the global position error of each point in the map is proportional to the size of the map. This kind of error cannot be eliminated without external reference data. A SLAM algorithm with loop closing, assuming robust operation, guarantees a relative positioning error between two map points that is proportional to the distance between these points. An ideal system would guarantee a constant positioning error, independent of the relative position of two points. If this error is measured in relation to a global frame of reference, such as the Earth, the system could be used to replace external global positioning systems.

The primary objective of this study is to introduce a novel approach that reduces the incremental positioning error to a constant factor, particularly focusing on top-down aerial SLAM algorithms (Avola et al. 2019). The proposed algorithm is described as a standalone component which can be used to align the map in any SLAM algorithm.

Traditional methods for image alignment often rely on detecting and matching characteristic keypoints, accompanied by complex feature descriptor computations (Bay et al. 2006, Rublee et al. 2011). However, these methods face challenges when

dealing with images from different sources that exhibit visual disparities due to varying lighting conditions and camera characteristics. Moreover, such methods tend to be computationally demanding and may require distinct, visually characteristic points in the image.

The proposed Embed Match algorithm uses image embeddings generated by a neural network. These embeddings encode the semantic content of a given map region and enable alignment across domains, irrespective of lighting conditions and camera characteristics. A similar method is used in the SuperPoint algorithm (Quach et al. 2021) and HFNet-SLAM (Liu et al. 2023). These algorithms focus on detecting high quality keypoints that can be used for matching images from different sources. Embed Match does not detect individual keypoints but instead considers the entire image as a dense field of embedding vectors.

In summary, this paper introduces the Embed Match algorithm as a novel solution to the challenge of global alignment in aerial SLAM algorithms. By using neural network-generated semantic embeddings, Embed Match offers a new approach that holds promise for enhancing the accuracy and applicability of aerial mapping and navigation tasks.

2. Methods

2.1. High level proof of the algorithm

Embed Match works alongside an external SLAM algorithm. The goal of Embed Match is to guarantee a constant relative position error between any two points on the created map. The map is represented by a point cloud where each point is associated with an image taken by the aerial vehicle at that position. To simplify the proof of the algorithm’s correctness, we consider a 1-dimensional SLAM problem. Let x_{real}^i and x_{est}^i be the actual and estimated positions of image point i . We want to guarantee that:

$$\exists e_{\text{em}} \forall i : \left| x_{\text{real}}^i - x_{\text{est}}^i \right| < e_{\text{em}} \quad (1)$$

holds true even when new points are added to the map. We call e_{em} the Embed Match error. We define the *EmbedMatch*(i) procedure which, given an image point i , adjusts x_{est}^i so that:

$$\left| x_{\text{real}}^i - x_{\text{est}}^i \right| < e_{\text{em}} \quad (2)$$

The limitation of $EmbedMatch(i)$ is that it requires the initial condition:

$$\left| x_{\text{real}}^i - x_{\text{est}}^i \right| < e_{\text{max}} \quad (3)$$

where e_{max} is the maximum Embed Match error – the maximum allowed divergence of the image point being aligned. The $EmbedMatch(i)$ procedure shall be described later. Next, we assume the existence of a $LocalAlign(i, j)$ procedure, which aligns neighbouring image points i and j . Such a procedure is provided by the SLAM algorithm used alongside Embed Match. In the case of ORB-SLAM3 (Campos et al. 2021), it is the matching algorithm using the ORB feature descriptor (Rubblee et al. 2011). $LocalAlign(i, j)$ assures that:

$$\left| \left(x_{\text{real}}^i - x_{\text{real}}^j \right) - \left(x_{\text{est}}^i - x_{\text{est}}^j \right) \right| < e_{\text{loc}} \quad (4)$$

for neighbouring points i and j , where e_{loc} is the local alignment error. Using $LocalAlign(i, j)$ for a sequence of neighbouring image points accumulates e_{loc} over the length of the sequence. This accumulation results in a position error that is proportional to the size of the map, and is the main problem solved by Embed Match.

Initially, the SLAM algorithm collects the x_{est}^0 origin point estimate. If the robot's global starting position x_{real}^0 can be estimated with an error less than e_{max} , $EmbedMatch(0)$ can be used to assure that:

$$\left| x_{\text{real}}^0 - x_{\text{est}}^0 \right| < e_{\text{em}} \quad (5)$$

Next, let us consider a generic point i that is initialised to point 0. For this point, Equation (2) is initially satisfied because of Equation (5). The robot now collects the next image point $x_{\text{est}}^{i+1} = x_{\text{est}}^j$. Since this point will be neighbouring with point i , $LocalAlign(i, j)$ can be used to satisfy Equation (4). From Equations (2) and (4), it follows that (omitting exact steps for brevity):

$$\left| x_{\text{real}}^j - x_{\text{est}}^j \right| < e_{\text{em}} + e_{\text{loc}} \quad (6)$$

which is simply the accumulation of the Embed Match error and the local alignment error. Under the condition that:

$$e_{\text{em}} + e_{\text{loc}} < e_{\text{max}} \quad (7)$$

$EmbedMatch(j)$ can be used to satisfy:

$$\left| x_{\text{real}}^j - x_{\text{est}}^j \right| < e_{\text{em}} \quad (8)$$

This method can be used to satisfy Equation (1) by being applied to each image point in the sequence of collected sensor samples.

2.2. Embed Match procedure

The Embed Match procedure is used to align the input image to a global frame of reference. It works by maximising the satellite alignment score of the input image. The satellite alignment score is defined as a function of the estimated latitude and longitude of the input image. Given an input image and a latitude and longitude, the score is the mean patch similarity between the input image and the corresponding satellite image taken around the given latitude and longitude. The real-world size of the area represented by the input image is assumed to be known. It can be computed from the altitude of the camera and its focal length. To compute the mean patch similarity between two images, both images are divided into overlapping patches with a moving window technique. The moving window size and stride are hyper-parameters of Embed Match and depend of the used camera setup. For images with a resolution of 512×512 pixels, a window of size 128 pixels and stride 32 pixels was chosen empirically. After obtaining image patches, the mean embedding alignment (MEA) score is computed between pairs of corresponding patches. The MEA score is used as the satellite alignment score, taking values from -1 to 1 .

2.3. Mean embedding alignment score

The MEA score is the core idea behind Embed Match. It is a measure of alignment between two images. For each patch in the input images, an embedding vector is computed with a neural network, in this case the OpenAI CLIP model (Radford et al. 2021). This process is illustrated in Figure 1. The MEA score is the mean of cosine similarity metrics between pairs of corresponding patch embeddings. If two patches show the same area, the cosine similarity of their embeddings will be close to 1 . As the patches become more distinct, the cosine similarity approaches -1 . The choice of an embedding neural network has significant effects on the quality of MEA. The CLIP model produces embeddings which are spatially-invariant. Spatially-invariant embeddings can only be used with a small patch size because embeddings of larger patches are less sensitive to small translations of the image. Since CLIP is trained with a text-based contrastive learning objective, its embeddings encode a general description of a given patch. With a smaller patch size, two neighbouring patches will have distinct contents which is used to guide image alignment.

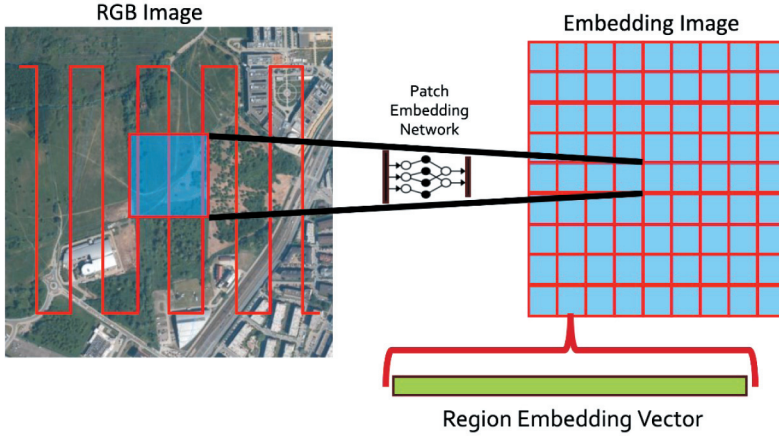


Fig. 1. Representing images with Embed Match

2.4. Maximising mean embedding alignment

Since the MEA is defined as a pure function of two variables, many numerical optimisation algorithms can be used. Embed Match assumes a maximum Embed Match error $e_{\max}$. Therefore, Bayesian optimisation could be an efficient method to optimise MEA as a Gaussian process. With the assumption that the area within $e_{\max}$ of the global maximum of MEA contains only one maximum, a gradient descent algorithm is used instead. The function gradient is approximated with the symmetric difference quotient. The discrete step h is a hyper-parameter of Embed Match defined proportionally to the size of the input image. With a step of 10%, the MEA is sampled at a position translated by 10% of the image size. After numerically computing the gradient with regard to the estimated latitude and longitude, a gradient descent iteration updates the position estimate. When the MEA stops increasing, the learning rate of gradient descent is halved. This iterative process is repeated for a fixed number of iterations. The result of the process is the latitude and longitude with the highest MEA score.

3. Experiments

Measuring the effectiveness of Embed Match when paired with an actual SLAM algorithm is out-of-scope for this work. Such an application is explained by the proof illustrating the intended usage of Embed Match. Instead, we focus exclusively on the image alignment procedure which is the core component of Embed Match. The model used for generating embeddings is the OpenAI CLIP ViT base variant with

a patch size of 32 (not to be confused with the Embed Match sliding window patch size). The experiments were carried out based on custom data collected with a DJI Mavic Zoom quadcopter.

3.1. Visualising patch embeddings

The first experiment is meant to provide insights into the quality of the embeddings used for alignment. Two example images are used for the experiment. One is a satellite map image from the mapy.geoportal.gov.pl service and the other is a photo taken by a DJI Mavic Zoom quadcopter of the same area from an altitude of 100 m. The images are split into patches according to the Embed Match algorithm and patch embeddings are computed. The embeddings are reshaped into a 3-dimensional tensor where the third dimension is the embedding vector of one patch. The embeddings are processed with primary component analysis (PCA) to produce an RGB image. Bicubic interpolation is used to visualise the image. The resulting images can be interpreted as a semantically-compressed version of the source images.

3.2. MEA convergence

The MEA score is intended to be the highest when images are perfectly aligned. This assumption is tested by computing the MEA between a source image and target images translated by values taken from an even grid of points. A grid of 5×5 cells with a cell size of 5 m is created around the true geographic origin of the input image. This grid of MEA scores is visualised as a 3-dimensional surface showing whether the MEA has a correct global maximum.

3.3. Image alignment

Finally, the entire Embed Match procedure is used to align a real drone image with a satellite map reference. The path taken by the gradient descent algorithm is analysed, as well as the observed values of MEA. The final image alignment result is visualised. This experiment shows the intended usage of Embed Match alongside a SLAM algorithm.

4. Results

The quadcopter drone image and the corresponding satellite image used for the experiments are shown in Figure 2.



Fig. 2. Comparison of drone (a) and satellite (b) images of the same area

4.1. Visualising patch embeddings

Figure 3 shows side-by-side the patch embeddings generated with images from Figure 2. Both PCA-reduced images contain visually similar areas. The content of the source images differs greatly since one was taken years before the other. Many objects have been moved, some have disappeared, and some have been added. Even though the images come from completely different sources, their embeddings are similar.

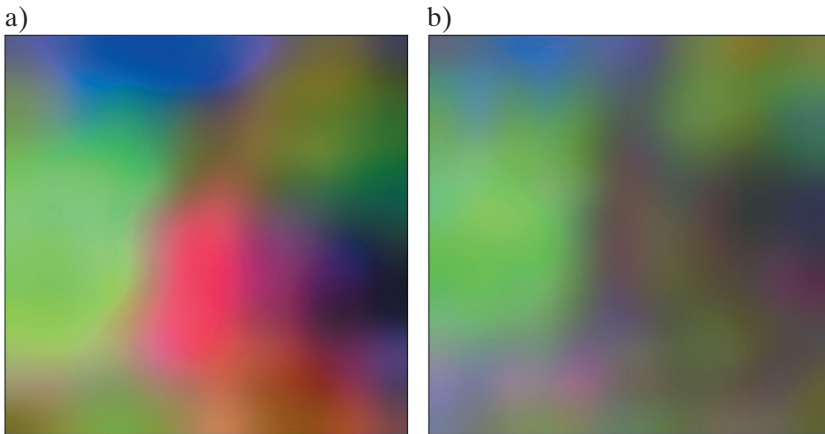


Fig. 3. Comparison of drone (a) and satellite (b) image embeddings

4.2. MEA convergence

Figure 4 visualises the MEA score as a 3-dimensional function of latitude and longitude. The global maximum of MEA is in the middle, corresponding to the actual

global position of the image being aligned. No other local maxima can be seen, and the function has a smooth surface. This validates the use of gradient descent and shows that the MEA correctly assigns the highest value to the place of perfect alignment.

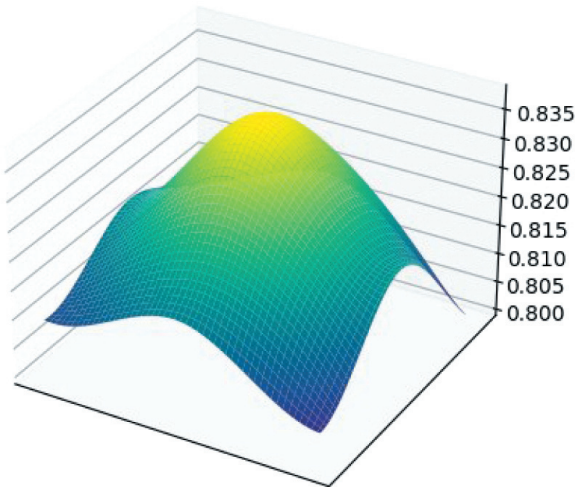


Fig. 4. Visualisation of the MEA score function

4.3. Image alignment

Figure 5 shows the results of Embed Match image alignment. The procedure begins with the actual drone image and an initial guess of the latitude and longitude. After ten iterations of Embed Match, the image is correctly aligned. The iterative progress of Embed Match is depicted in Figure 6, illustrating the path traversed by the algorithm. After a few initial large steps, the algorithm converges on the global maximum of MEA. Figure 7 shows the value of MEA at each iteration of Embed Match.

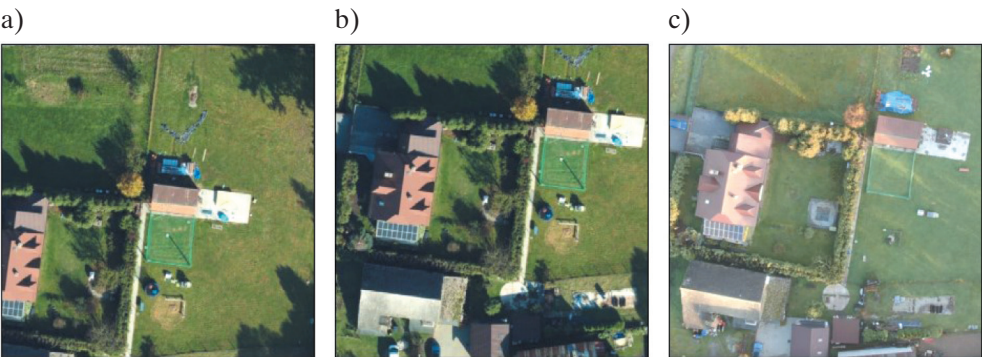


Fig. 5. Aligning images with Embed Match: a) initial guess; b) aligned; c) actual

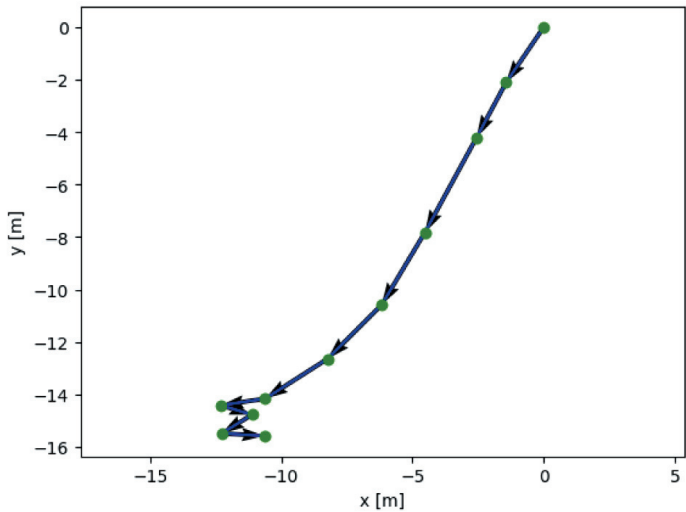


Fig. 6. Path taken by Embed Match

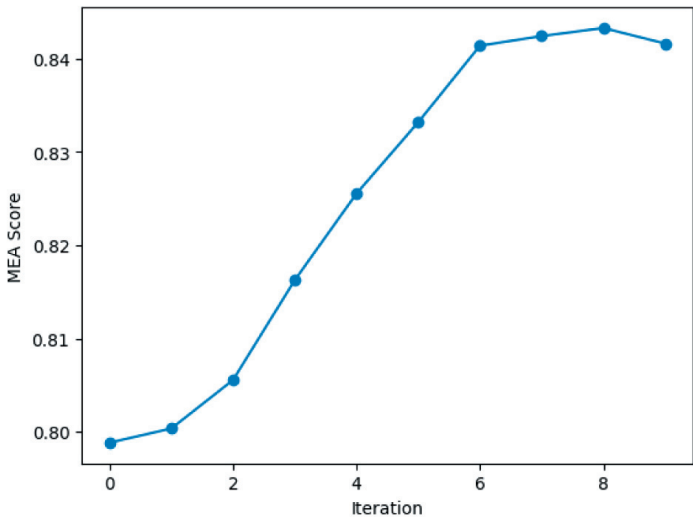


Fig. 7. Optimising MEA with gradient descent

5. Discussion

The results show that Embed Match successfully aligns drone images to a satellite map. The visual similarity of PCA-reduced embeddings can be questioned. Since PCA performs a reduction from 512 embedding features to 3 RGB values, most of the information is lost.

However, analysing the MEA function shows that comparing cosine similarity of patch embeddings is a valid method of aligning images. Only the relative comparison of embeddings is important. Further work could analyse the MEA score of images where one is covered by snow and other anomalies.

The algorithm could be improved by implementing a more advanced MEA optimisation method, such as Adam (Kingma and Ba 2015). Many other optimisation methods can be used, both gradient-based and gradient-free. It is also possible to use MEA to align an image without an initial position guess. In such case an optimisation method that can avoid local maxima has to be used.

The final result obtained with Embed Match demonstrates its effectiveness. The algorithm adjusts the initial position guess by 19 m, which is 30% of the input image width. This shows that Embed Match can be used to align the map created by a SLAM algorithm even with a large initial error.

6. Conclusions

This work explored the application of neural embedding models to the task of cross-domain image alignment. A novel algorithm called Embed Match is introduced. A mathematical proof is used to describe a visual top-down aerial SLAM algorithm using satellite imagery. The proposed solution guarantees a constant upper bound on the global position error of each point on the created map. The Embed Match algorithm introduces a new class of image alignment methods based on whole-image embedding similarity. Experiments are conducted that explore the behaviour of the proposed methods. The results show that the algorithm correctly aligns images. Future research directions for embedding-based image alignment algorithms are discussed.

References

- Avola D., Cinque L., Fagioli A., Foresti G.L., Massaroni C., Pannone D., 2019, *Feature-Based SLAM Algorithm for Small Scale UAV with Nadir View*, [in:] Ricci E., Rota Bulñ S., Snoek C., Lanz O., Messelodi S., Sebe N. (eds.), *Image Analysis and Processing – ICIAP 2019: 20th International Conference, Trento, Italy, September 9–13, 2019, Proceedings, Part II*, Lecture Notes in Computer Science, vol. 11752, Springer, Cham, pp. 457–467. https://doi.org/10.1007/978-3-030-30645-8_42.
- Bay H., Tuytelaars T., Van Gool L., 2006, *SURF: Speeded Up Robust Features*, [in:] Leonardis A., Bischof H., Pinz A. (eds.), *Computer Vision – ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7–13, 2006, Proceedings, Part I*, Lecture Notes in Computer Science, vol. 3951, Springer, Berlin, Heidelberg, pp. 404–417. https://doi.org/10.1007/11744023_32.

- Campos C., Elvira R., Rodríguez J.J.G., Montiel J.M.M., Tardós J.D., 2021, *ORB-SLAM3: An accurate open-source library for visual, visual-inertial, and multimap SLAM*, IEEE Transactions on Robotics, vol. 37(6), pp. 1874–1890. <https://doi.org/10.1109/TRO.2021.3075644>.
- Gupta A., Fernando X., 2022, *Simultaneous localization and mapping (SLAM) and data fusion in unmanned aerial vehicles: Recent advances and challenges*, Drones, vol. 6(4), 85. <https://doi.org/10.3390/drones6040085>.
- Kingma D.P., Ba J.L., 2015, *Adam: A method for stochastic optimization*, [in:] *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7–9, 2015, Conference Track Proceedings*. <https://doi.org/10.48550/arXiv.1412.6980>.
- Liu L., Aitken J.M., 2023, *HFNet-SLAM: An accurate and real-time monocular SLAM system with deep features*, Sensors, vol. 23(4), 2113. <https://doi.org/10.3390/s23042113>.
- Macario Barros A., Michel M., Moline Y., Corre G., Carrel F., 2022, *A comprehensive survey of visual SLAM algorithms*, Robotics, vol. 11(1), 24. <https://doi.org/10.3390/robotics11010024>.
- Quach C.H., Phung M.D., Le H.V., Perry S., 2021, *SupSLAM: A robust visual inertial SLAM system using SuperPoint for unmanned aerial vehicles*, [in:] Bao V.N.Q., Luong M.D. (eds.), *2021 8th NAFOSTED Conference on Information and Computer Science (NICS): NICS 2021: Proceedings: December 21–22, 2021, Hanoi, Vietnam*, IEEE, Piscataway, pp. 507–512. <https://doi.org/10.1109/NICS54270.2021.9701527>.
- Radford A., Kim J.W., Hallacy C., Ramesh A., Goh G., Agarwal S., Sastry G., Askell A., Mishkin P., Clark J., Krueger G., Sutskever I., 2021, *Learning transferable visual models from natural language supervision*, [in:] Meila M., Zhang T. (eds.), *Proceedings of the 38th International Conference on Machine Learning: 18–24 July 2021, Virtual*, Proceedings of Machine Learning Research, vol. 139, pp. 8748–8763. <https://doi.org/10.48550/arXiv.2103.00020>.
- Rublee E., Rabaud V., Konolige K., Bradski G., 2011, *ORB: An efficient alternative to SIFT or SURF*, [in:] Metaxas D.N., Quan L., Sanfeliu A., Van Gool L. (eds.), *IEEE International Conference on Computer Vision, ICCV 2011, Barcelona, Spain, November 6–13, 2011*, IEEE, Piscataway, pp. 2564–2571. <https://doi.org/10.1109/ICCV.2011.6126544>.

Charakterystyka mikrostruktury i własności materiałów stosowanych do wyrobu kling szermierczych

Robert Karpiński , Adam Kokosza 

AGH Akademia Górniczo-Hutnicza, Wydział Inżynierii Metali i Informatyki Przemysłowej, Kraków

Streszczenie: W pracy przedstawiono wyniki analizy mikrostruktury, własności i stanu naprężeń w klingach szermierczych wykorzystywanych podczas treningów nowoczesnej szermierki klasycznej. Badania wykonano na trzech klingach od producenta 1 po różnych etapach ich eksploatacji (nowa, eksploatowana, zniszczona) oraz na jednej klindze od producenta 2, która uległa zniszczeniu. Dzięki zastosowanym metodom badawczym określono, jak pod wpływem ich eksploatacji zmienia się stan naprężeń w klingach oraz w jaki sposób przebiega degradacja ich powierzchni. Wyciągnięto także pewne wnioski na temat ich obróbki cieplnej oraz różnicy pomiędzy procesami produkcyjnymi różnych producentów.

Słowa kluczowe: szpada, szermierka, klinga, stal sprężynowa, mikrostruktura, metoda szumów Barkhausena, SEM, pomiary twardości

CHARACTERIZATION OF MICROSTRUCTURE AND PROPERTIES OF MATERIALS USED IN FENCING BLADES

Abstract: This paper is dedicated to the analysis of the microstructure, properties and stress states of fencing blades used in modern classical fencing trainings. The research has been conducted on three blades from manufacturer 1 after three different stages of exploitation (new, used and destroyed) and on one destroyed blade from manufacturer 2. Thanks to the applied research methods, the change in stress states and the degradation of the surface of the blades has been evaluated. Some conclusions about their heat treatment and differences in the manufacturing methods between different manufacturers have been made as well.

Keywords: smallsword, fencing, fencing blade, spring steel, microstructure, Barkhausen noise analysis, SEM, hardness tests

1. Wprowadzenie

Zarówno podczas treningów, jak i zawodów szermierczych, pod wpływem kontaktu z bronią czy ciałem przeciwnika, klingi często ulegają znacznemu wygięciu. Wygięcia te najczęściej mają charakter sprężysty, ale w skrajnym wypadku mogą okazać się trwałe. Badania na temat takich obciążeń (Kibaroglu i in. 2017) wykazują, że naprężenia działające w warstwie wierzchniej klingi podczas starcia mogą być bardzo wysokie, co przy cyklicznym charakterze wygięć może prowadzić do pojawienia się pęknięć, szczególnie w miejscach takich jak wyszczerbienia lub obszary silnie zdeformowane.

W innych pracach jako przyczynę zniszczenia podaje się energię odkształcenia oraz szczątkowe odkształcenia plastyczne (Vivaldi i in. 2018) lub akumulację mikrouszkodzeń związanych ze zmęczeniem niskocyklowym (Jiang i in. 2022). Przykładem niebezpieczeństw wynikających z niekontrolowanego zniszczenia broni podczas zawodów jest tragiczna śmierć sowieckiego szermierza Władimira Smirnowa w 1982 roku. Szacuje się, że między latami 1980 a 2006 wystąpiło osiem wypadków śmiertelnych powiązanych z szermierką olimpijską na skutek obrażeń spowodowanych ostrą końcówką złamanej broni (Roi i in. 2008).

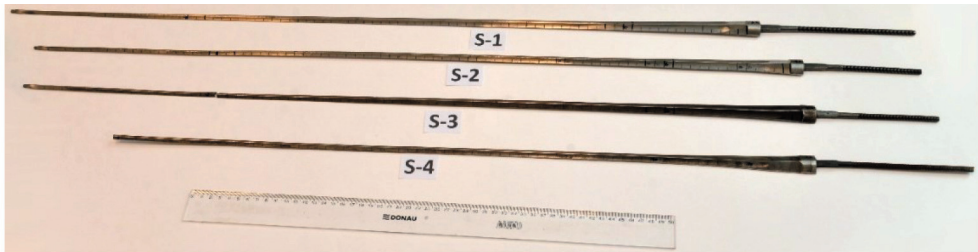
Aby ograniczyć ryzyko niebezpiecznych wypadków, do wyrobu kling stosuje się materiały, które muszą spełnić określone wymogi dotyczące wytrzymałości na rozciąganie, uderności i wytrzymałości zmęczeniowej. Obecnie dwie najczęściej stosowane grupy materiałów na klingi szermiercze to krzemowe stale sprężynowe (do produkcji broni treningowej oraz broni używanej na zawodach na niskim szczeblu) oraz stale typu maraging (do produkcji broni stosowanej na zawodach na najwyższym szczeblu, organizowanych pod patronatem Międzynarodowej Federacji Szermierczej FIE – Fédération Internationale d’Escrime) (Smallman i in. 1999).

O ile klingi ze stali typu maraging zostały dokładnie opisane w regulaminie FIE i wymagania dotyczące ich użytkowania są znane, o tyle odnośnie do kling ze stali sprężynowych nie ma takich warunków poza wewnętrznymi wymaganiami wytwórców.

Zdaniem autorów opisane w niniejszej pracy badania pozwolą na lepsze poznanie zagadnień dotyczących naprężeń, mikrostruktury i potencjalnych przyczyn niszczenia kling wykonywanych z tej grupy materiałów, tj. ze stali sprężynowych.

2. Materiał do badań

Do badań dostarczono cztery klingi (rys. 1) wykorzystywane na treningach nowoczesnej szermierki klasycznej: trzy klingi jednego producenta po różnych etapach eksploatacji (S-1, S-2, S-3) oraz jedną zniszczoną klingę od drugiego producenta (S-4).



Rys. 1. Dostarczone do badań klingi szpadowe

W tabeli 1 zestawiono natomiast wyniki analizy składu chemicznego każdej z czterech kling. Badania te zostały przeprowadzone za pomocą spektrometru iskrowego WAS Foundry-Master.

Tabela 1

Wyniki analizy składu wykonane za pomocą spektrometru iskrowego WAS Foundry-Master

Oznaczenie	Fe [% mas.]	C [% mas.]	Si [% mas.]	Mn [% mas.]	P [% mas.]	S [% mas.]	Cr [% mas.]	Ni [% mas.]
S-1	reszta	0,44	1,63	0,66	0,02	0,02	0,17	0,11
S-2	reszta	0,39	1,65	0,65	0,04	0,04	0,08	0,04
S-3	reszta	0,35	1,60	0,65	0,02	0,02	0,09	0,05
S-4	reszta	0,25	1,72	0,78	0,03	0,01	0,19	0,10

Jak widać, badane klingi cechują się średnią, a nawet niską, zawartością węgla oraz podwyższoną zawartością krzemu. Zwraca uwagę to, że skład chemiczny klingi S-4 różni się od pozostałych. W klindze tej stwierdzono wyraźnie najmniejszą zawartość węgla (0,25%) i jednocześnie największą zawartość krzemu (1,72%), co potwierdza, że pochodziła ona od innego dostawcy. Wyniki analizy pozwalają na zaklasyfikowanie materiału, z którego zostały wykonane klingi, jako prostych stali sprężynowych typu 60S2A (PN-H-84032:1974), ale z obniżoną zawartością węgla – na poziomie 0,25–0,44% mas.

3. Metodyka

Podczas badań wstępnych oceniono naprężenia w warstwie wierzchniej całych kling. Za pomocą metody Barkhausena i urządzenia Stresscan 500C zmierzono wartość parametru magnetosprężystego (MP) na całej ich długości. Parametr taki zależy od mikrostruktury i stanu naprężeń w materiale. W przypadku materiałów twardych lub przy wysokim poziomie naprężeń ściskających wartość tego parametru jest niska i rośnie ze spadkiem twardości i wzrostem naprężeń rozciągających (Stresstech 2023).

Na potrzeby dalszych badań klingi zostały pocięte. Wycięte z nich próbki zostały wykorzystane do szczegółowych badań powierzchni bocznej, a w przypadku kling S-3 i S-4 również do analizy przełomów. Badania te wykonano przy użyciu skaningowego mikroskopu elektronowego PHENOM XL.

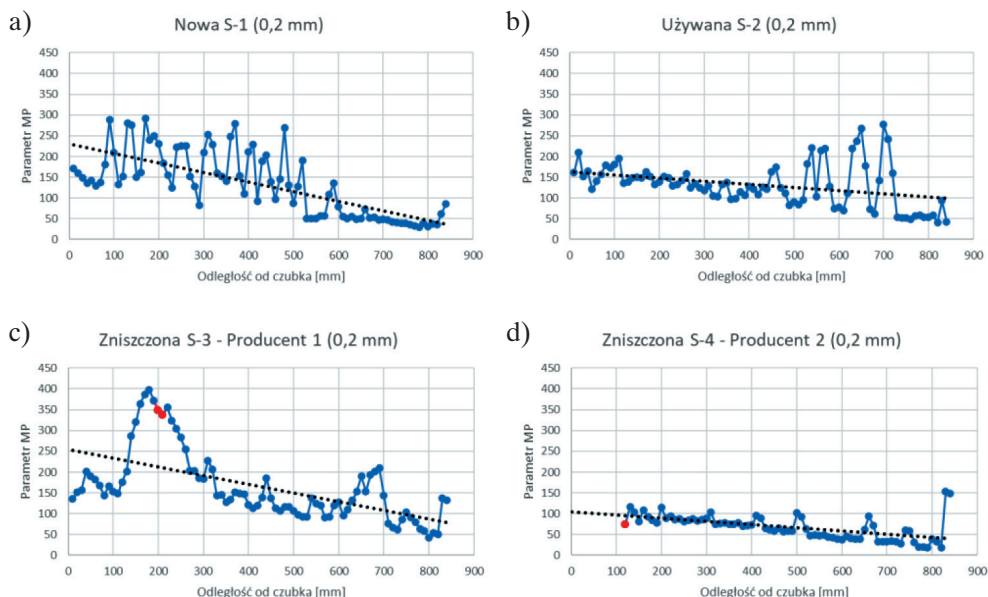
W dalszej kolejności wycięte próbki poddano badaniom metalograficznym. W tym celu próbki zostały zainkludowane, a następnie wyszlifowane i wypolerowane. Jako odczynnik trawiący wykorzystano 4-procentowy Nital. Badania mikrostruktury wykonano za pomocą mikroskopu świetlnego Zeiss AxioVert 200 MAT przy wykorzystaniu obiektywów o powiększeniu $5\times$, $20\times$ oraz $100\times$.

Po badaniach mikrostruktury na wszystkich zglądach poprzecznych próbek wyciętych z kling wykonano pomiary twardości metodą Vickersa. Do badań tych wykorzystano twardościomierz HPO 250 z obciążeniem 20 kg.

4. Wyniki badań i ich dyskusja

4.1. Ocena stanu naprężeń – pomiary parametru MP

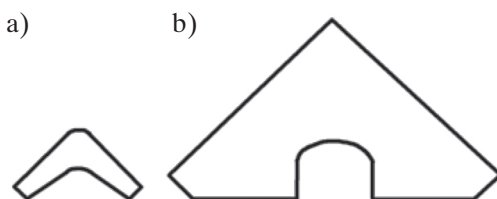
Na rysunku 2 zestawiono wyniki pomiarów parametru MP na długości kling w stanie dostawy (niepociętych) dla głębokości 0,2 mm. Dodatkowo na każdym z wykresów zaznaczono linię trendu parametru MP. Na zniszczonych klingach S-3 i S-4 na czerwono zaznaczono wyniki pomiarów najbliższej przełomu.



Rys. 2. Zestawienie wyników pomiarów parametru MP na głębokości 0,2 mm kling w stanie dostawy

Zauważyć można, że parametr MP w nieużywanej klindze S-1 (rys. 2a) zmienia się od ok. 180 na czubku do ok. 35 przy rękojeści. Ze względu na geometrię klingi i specyfikę technologii jej wykonania, a w szczególności jej obróbkę cieplną (hartowanie), parametr MP powinien wzrastać w kierunku rękojeści, ponieważ niższa twardość rzutuje na zwiększoną wartość parametru MP (Nahak 2017). Tymczasem pokazane na rysunku 4 zmiany mają trend odwrotny. Najbardziej prawdopodobne są trzy przyczyny takiej sytuacji:

- 1) Geometria klingi. Pomimo spodziewanej największej twardości przy czubku, związanej z hartowaniem i wytworzeniem się w tym obszarze martenzytu oraz powstaniem największych naprężeń ściskających (niski parametr MP), kształt klingi zwracający się w stronę czubka sprawia, że w „cienkiej” części trudno jest zachować bardzo duże naprężenia ściskające. Najprawdopodobniej, z powodu niskiego ilorazu objętości materiału i powierzchni, naprężenia takie w tej części klingi ulegają częściowej relaksacji.
- 2) Wycięty rowek na kabel sygnalizacyjny. Ten element jeszcze bardziej obniża iloraz objętości materiału i powierzchni, ponieważ objętości ubywa, a powierzchnia jest w większym stopniu rozbudowana. Sprawia to, że naprężenia ściskające mogą być generowane w jeszcze mniejszej objętości materiału (por. rys. 3), przez co łatwiej są rozpraszane.
- 3) Obróbka wykończeniowa przez szlifowanie. Proces ten powoduje zazwyczaj powstanie w materiale naprężeń rozciągających (Tota i in. 1994).



Rys. 3. Geometria przekroju poprzecznego klingi w jej różnych częściach:
a) przy czubku; b) przy rękojeści

Na rysunku 2a można również zauważyć, że w odległości od ok. 80 mm od czubka aż do ok. 520 mm występują skokowe zmiany parametru MP. Można przypuszczać, że przyczyną takich zmian jest także specyfika procesu produkcji, np. związana z technologią kucia klingi lub sposobem jej szlifowania.

Pokazany na rysunku 2a wykres parametru MP traci swój skokowy charakter w części najbardziej zbliżonej do rękojeści (ponad 600 mm od czubka) i przyjmuje charakter niemal liniowy. Warto przypomnieć, że ta część klingi jest najszerza (por. rys. 1), a rowek najwęższy w stosunku do przekroju. Liniowy charakter wykresu można zatem powiązać z mniejszym wpływem procesu frezowania rowka na naprężenia występujące w tej części klingi.

Na rysunku 2b pokazano zmiany parametru MP na kładzie S-2 od producenta 1 użytkowanej przez ok. 24 miesiące. Zależność ta różni się znacząco od tej pokazanej na rysunku 2a. Przede wszystkim zauważyć można, że zależność nie ma już skokowego charakteru w części klingi pomiędzy czubkiem a punktem oddalonym o 500 mm. Może to być związane z częstym plastycznym wyginaniem klingi podczas treningów (Kibaroglu i in. 2017). Klinga po odkształceniu jest zazwyczaj ręcznie naprostowywana, co wprowadza do niej dodatkowe naprężenia, które prawdopodobnie kompensują naprężenia związane z wytwarzaniem.

Należy też zauważyć lokalny wzrost wartości parametru MP w części zbliżonej do rękojeści oraz skokowy charakter wykresu w tym obszarze. Można przypuszczać, że jest on związany z ze sposobem posługiwania się szpadą i stawiania zasłon, bowiem ta część klingi często używana jest do parowania natarć przeciwnika. Może to powodować liczne drobne odkształcenia i wyszczerbienia, co również wpływa na lokalną relaksację naprężeń względem ich poziomu po procesie produkcji (por. rys. 2a).

Pokazany na rysunku 2c wykres ilustrujący zmiany parametru MP w zniszczonej kładzie S-3 różni się zasadniczo od wykresów na rysunkach 2a i 2b. Najważniejszą różnicą jest silny wzrost parametru MP w bezpośrednim sąsiedztwie przełomu. Najprawdopodobniej jest to efekt całkowitej relaksacji naprężeń ściskających w tym obszarze, który następuje w momencie złamania klingi.

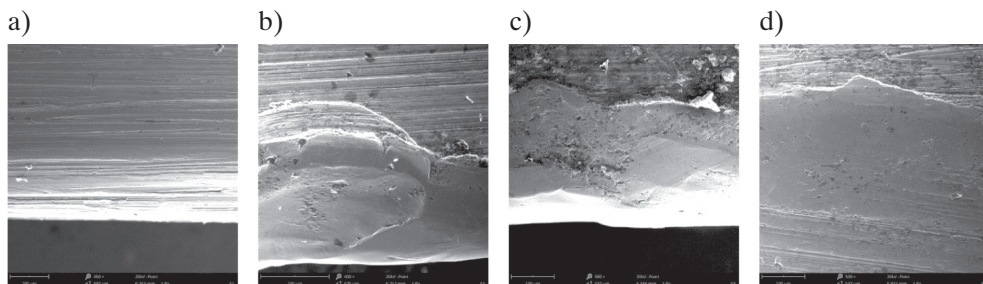
W kładzie S-4 (rys. 2d) wartości parametru MP są znacząco niższe, zmieniając się od ok. 70 przy przełomie do ok. 10 przy rękojeści z nielicznymi lokalnymi skokowymi zmianami. Z jednej strony tak niskie wartości MP wskazywać mogą na większą twardość klingi, z drugiej – na znacznie większy poziom naprężeń ściskających w kładzie S-4 w porównaniu z klingami S-1, S-2 i S-3. Równocześnie zauważalnie inny charakter wykresu sugeruje różnice w procesie produkcyjnym, potwierdzając, że klinga ta pochodzi od innego producenta i, jak wynika z tabeli 1, wykonana jest z innego materiału.

4.2. Analiza powierzchni kling na skanującym mikroskopie elektronowym

Na rysunku 4 pokazano obrazy powierzchni bocznej wszystkich badanych kling, które ujawniono przy użyciu skanującego mikroskopu elektronowego na próbkach wyciętych ok. 200 mm (klingi S-1, S-2 i S-3) lub ok. 120 mm (klinga S-4) od wierzchołka i zarejestrowano przy powiększeniach z zakresu 400–500×.

Obserwacje SEM potwierdzają, że klinga S-1 jest klingą nową (nieużywaną). Na rysunku 4a widać, że na jej powierzchni widoczne są wyłącznie rysy po szlifowaniu.

Na obrazie krawędzi klingi używanej S-2 (rys. 4b) nie widać już rys po szlifowaniu, a jedynie płaskie, silnie odkształcone obszary wskazujące na płynięcie materiału od krawędzi ku górze. Jest to prawdopodobnie efekt wielokrotnych zderzeń z innymi klingami podczas starć szermierczych. Co ważne, wewnątrz odkształconych obszarów pojawiają się zauważalne mikropęknięcia, co przy dalszym użytkowaniu może powodować poważniejsze wyszczerbienia również prowadzące do zniszczenia.

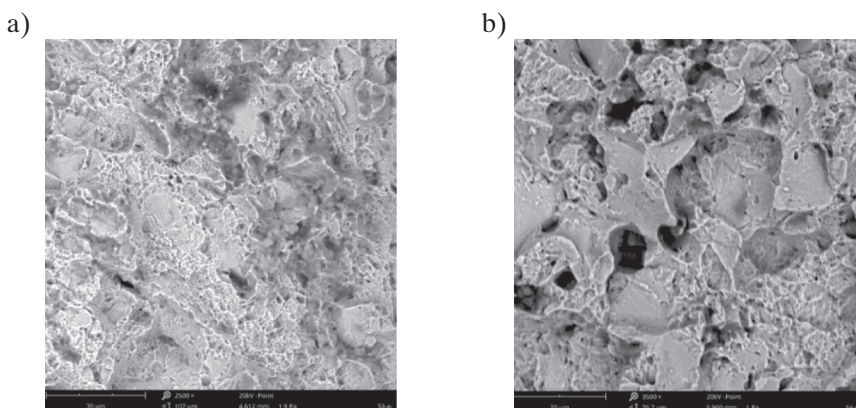


Rys. 4. Fotografie SEM okolic krawędzi kling: a) S-1; b) S-2; c) S-3; d) S-4

Podobnie do poprzedniej wygląda pokazana na rysunku 4c krawędź klingi S-3 (użytkowanej i zniszczonej). W tym przypadku na odształconych obszarach widoczne są także wtórne defekty, takie jak mikropęknięcia i wyszczerbienia w odształconej plastycznie warstwie. Sugeruje to, że klinga ta najprawdopodobniej była dłużej użytkowana.

Krawędź klingi S-4, która również uległa zniszczeniu, ale pochodziła od innego producenta (rys. 4d), także wygląda podobnie do krawędzi klingi S-2. W tym jednak przypadku odształcona plastycznie warstewka obejmuje większy obszar od krawędzi w porównaniu z obiema wcześniej omówionymi klingami. Jest to najprawdopodobniej skutek nieco mniejszej zawartości węgla w materiale tej klingi, co mogło spowodować łatwiejsze odształcanie się jej powierzchni podczas starć.

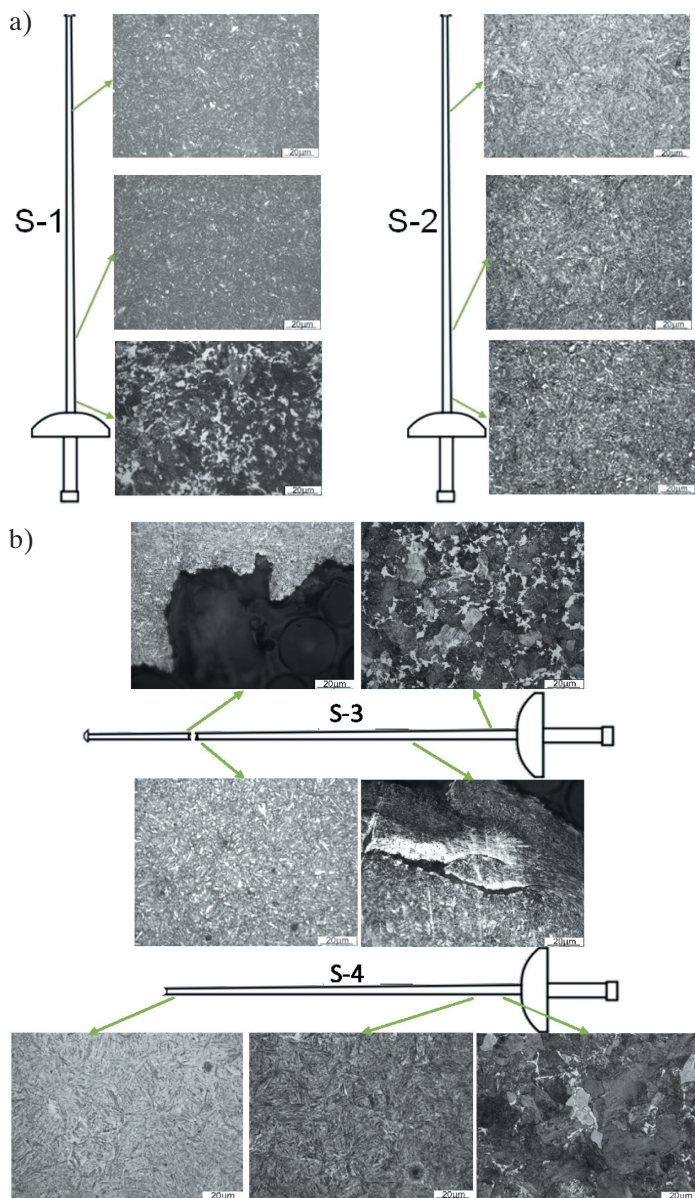
Na rysunku 5 pokazano przełomy, które powstały po złamaniu kling S-3 i S-4. Jak widać na rysunku 5a, w przypadku klingi S-3 przełom ma charakter mieszany, z dominującym jednak udziałem przełomu ciągliwego. Również w przypadku klingi S-4 (rys. 5b) dominuje przełom ciągliwy. Można wprowadzić zauważyć płaskie obszary przełomu, które mogłyby wskazywać na ich kruchy charakter, ale bardziej prawdopodobną przyczyną ich pojawienia się wydaje się „zaklepywanie się” powierzchni nierozwiniętego pęknięcia podczas korzystania z takiej częściowo pękniętej klingi.



Rys. 5. Analiza SEM przełomów złamanych kling S-3 (a) i S-4 (b)

4.3. Analiza mikrostruktury badanych kling

W ramach niniejszej pracy badaniami objęto cztery klingi, z których do dalszych badań wycięto po pięć lub sześć próbek. Na każdej z nich wykonano obserwacje mikrostruktury. W niniejszej pracy ograniczono się do przedstawienia najważniejszych i najciekawszych wyników. Na rysunku 6 pokazano mikrostruktury ujawnione na długości badanych kling.

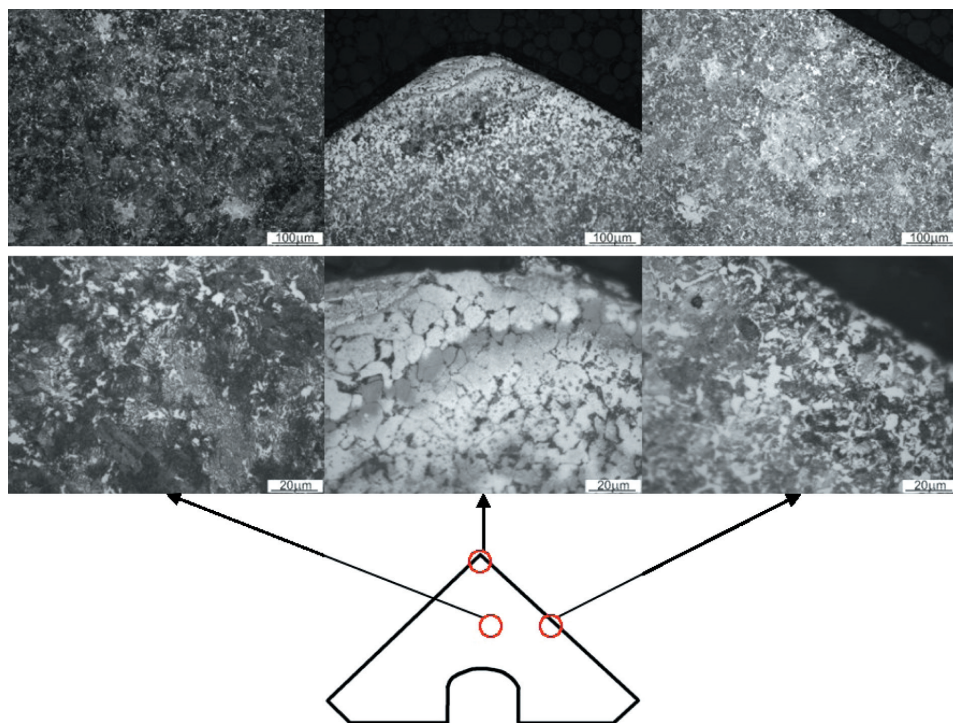


Rys. 6. Mikrostruktura badanych kling: a) niezniszczonych; b) zniszczonych

Jak można zauważyć, w przypadku wszystkich kling od wierzchołka do pewnej odległości od rękojeści występuje jedynie faza o charakterze iglastym, będąca najprawdopodobniej martenzytem odpuszczonym. W przypadku klingi S-4 należy jednak zwrócić uwagę na różnice w morfologii martenzytu względem kling od producenta 1. O ile w przypadku kling S-1, S-2 i S-3 morfologia ujawnionego martenzytu jest zbliżona, o tyle w przypadku klingi S-4 jest ona różna w każdym z miejsc, gdzie ta faza występuje. Wynikać to może z jednej strony z użycia innego materiału (niższa zawartość węgla) (Ji i in. 2011), z drugiej zaś strony może to być efekt innej technologii obróbki cieplnej.

Warto zwrócić też uwagę na zgład wykonany na kładzie S-3 w odległości ok. 600 mm od wierzchołka. Widoczny jest na nich odkształcony plastycznie fragment materiału, który najprawdopodobniej powstał w trakcie użytkowania. Potwierdza to wnioski z obserwacji tej klingi przedstawione na rysunku 2.

W przypadku wszystkich analizowanych kling najciekawsze okazały się obserwacje wykonane na próbkach wyciętych z części położonej najbliżej rękojeści. Najważniejsze wyniki obserwacji próbki wyciętej z klingi S-1 pokazano na rysunku 7.

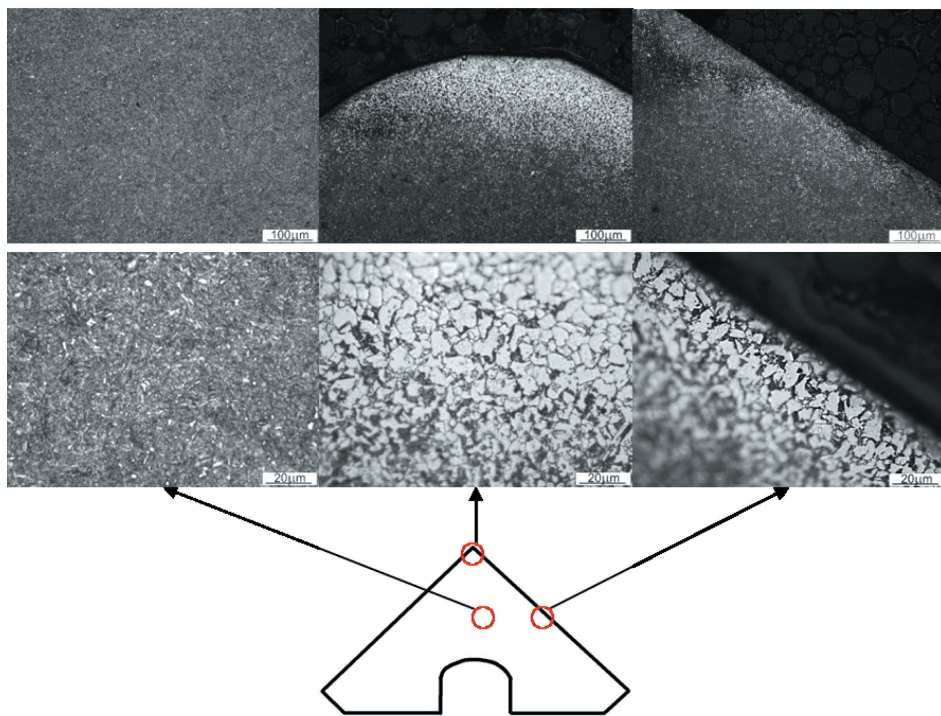


Rys. 7. Mikrostruktury ujawnione na zgładzie poprzecznym w pobliżu rękojeści klingi S-1

Na przekroju tej próbki nie stwierdzono w ogóle struktury martenzytycznej. Dominuje struktura charakterystyczna dla stali podeutektoidalnej, tj. ferrytyczno-perlityczna, z przeważającym udziałem perlitu. Ponadto widać, że mikrostruktura w pobliżu powierzchni na tym fragmencie klingi S-1 jest odwęglona. Szczególnie dobrze widoczne jest to na obszarze w pobliżu krawędzi klingi, gdzie dominuje mikrostruktura ferrytu.

Takie zmiany mikrostruktury sugerują, że – w przeciwieństwie do pozostałej części klingi S-1 (rys. 6) – ten jej fragment nie uległ zahartowaniu podczas obróbki cieplnej. Odwęglenie przy krawędzi wskazuje, że nastąpiło działanie wysokiej temperatury (austenityzowanie), jednak podczas chłodzenia w tym fragmencie klingi na całym przekroju nie osiągnięto prędkości krytycznej, np. na skutek zbyt płytkiego zanurzenia klingi w ośrodku chłodzącym.

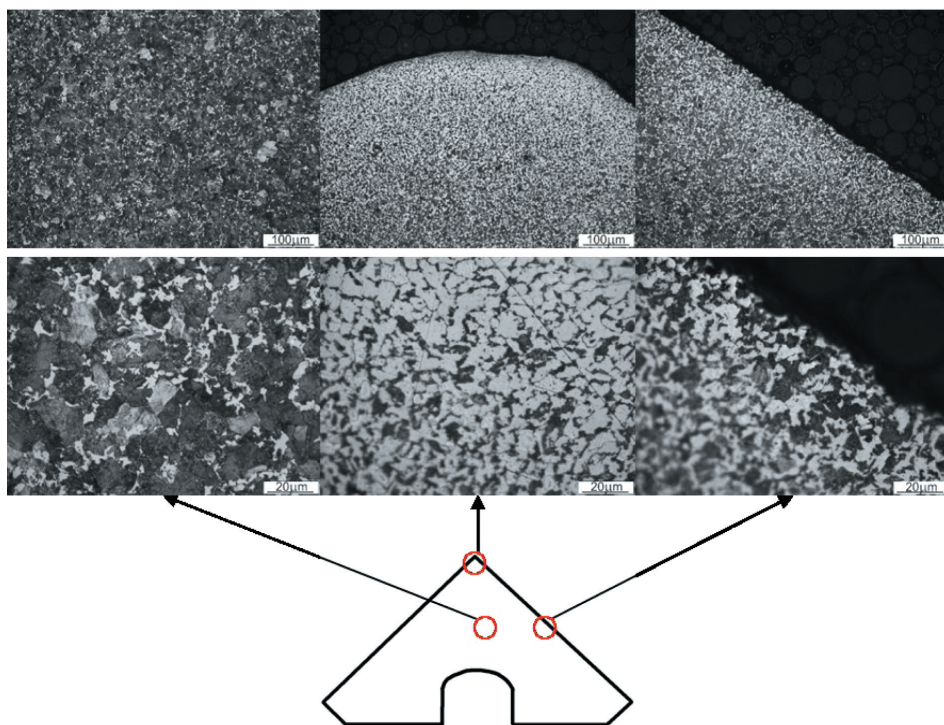
Na rysunku 8 pokazano mikrostrukturę próbki wyciętej w pobliżu rękojeści klingi S-2. Najistotniejszą różnicą w porównaniu z klingą S-1 jest to, że tutaj w rdzeniu klingi wystąpiła mikrostruktura iglasta. Najprawdopodobniej jest to wspomniany już martenzyt, co sugeruje, że podczas obróbki cieplnej w przeciwieństwie do klingi S-1 ta część klingi S-2 uległa zahartowaniu. Wskazuje to na możliwe różnice w wytwarzaniu kling S-1 i S-2, mimo że pochodzą one od tego samego producenta i należą do tej samej linii wyrobów.



Rys. 8. Mikrostruktury ujawnione na zglądzie poprzecznym w pobliżu rękojeści klingi S-2

W tym obszarze klingi S-2 również stwierdzono odwęglenie w pobliżu krawędzi, w którym dominuje mikrostruktura ferrytyczna. Odwęglenie materiału podczas obróbki cieplnej najprawdopodobniej uniemożliwiło zahartowanie tego obszaru klingi, pozostawiając drobne obszary ferrytu przy powierzchni.

Na rysunku 9 przedstawiono zdjęcia mikrostruktur klingi S-3. Można zauważyć, że analogicznie do klingi S-1 mikrostruktura w rdzeniu ma charakter ferrytyczno-perlityczny, a przy powierzchni występuje odwęglenie. Jest ono jednak silniejsze, co może wskazywać na zastosowanie innych parametrów w trakcie obróbki cieplnej tej klingi (Zhang i in. 2013).



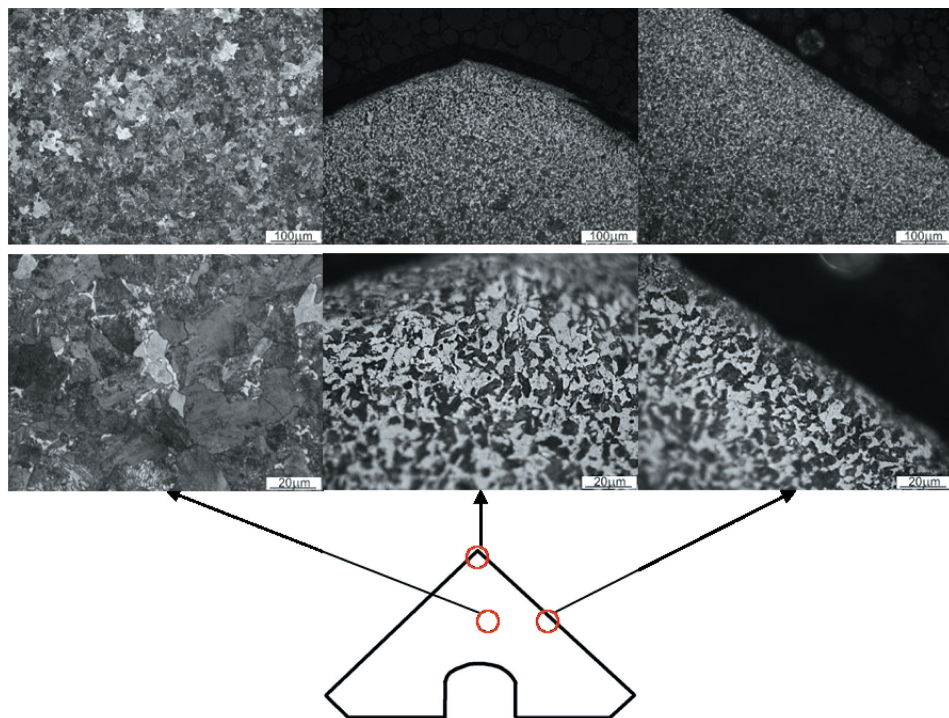
Rys. 9. Mikrostruktury ujawnione na zglądzie poprzecznym w pobliżu rękojeści klingi S-3

Powyższe wnioski pokazują, że proces produkcyjny kling producenta 1 nie jest powtarzalny, bowiem pojawiają się pewne różnice zarówno w sposobie, jak i w parametrach obróbki cieplnej produkowanych kling.

Na rysunku 10 pokazano mikrostruktury w próbce wyciętej najbliżej rękojeści klingi S-4.

Jak widać, tutaj na całym przekroju występuje mikrostruktura ferrytyczno-perlityczna. Brak efektów koagulacji cementytu wskazuje, że ten obszar klingi S-4

(od producenta 2) był poddawany austenitzowaniu, jednak podobnie jak w klingach S-1 i S-3 również nie był on chłodzony na tyle szybko, aby utworzył się martenzyt.



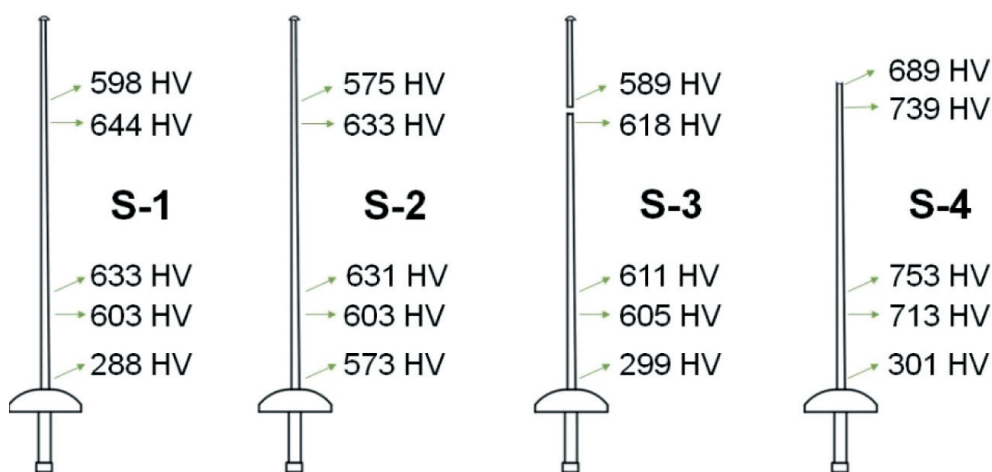
Rys. 10. Mikrostruktury ujawnione na zglądzie poprzecznym w pobliżu rękojeści klingi S-4

Ponadto na pokazanych na rysunku 10 obrazach mikrostruktur widoczny jest znaczny rozrost ziarna w rdzeniu. Kolonie perlitu w tym przypadku są znacznie większe niż w klingach od producenta 1. Jest to zapewne efekt zastosowania wyższej temperatury austenitzowania (Lee i Lee 2008) – co najprawdopodobniej związane było z innym składem chemicznym materiału klingi S-4 i niższą zawartością węgla. W klingie tej wystąpiło również najłabsze odwęglenie obszarów leżących przy powierzchni. Oznacza to, że proces obróbki cieplnej klingi S-4 wyraźnie różni się od procesów obróbki kling S-1, S-2 i S-3, ale – co ważne – w przypadku tej klingi zastosowano skuteczną ochronę przed odwęgleniem.

4.4. Analiza wyników pomiarów twardości

Na rysunku 11 pokazano wyniki pomiarów twardości HV20 na przekrojach poprzecznych wszystkich kling.

Analizując wyniki tych pomiarów dotyczące kling od producenta 1 (klingi S-1, S-2 i S-3), łatwo stwierdzić, że ich twardość pozostaje podobna i spada dopiero w pobliżu rękojeści. W klingach S-1 i S-3 przy rękojeści twardość wynosi ok. 30 HV, co potwierdza ujawnioną w tym miejscu mikrostrukturę ferrytyczno-perlityczną (por. rys. 7 i 9). Pokazany charakter zmian twardości najprawdopodobniej związany jest z technologią obróbki cieplnej kling, tj. hartowania i średniego odpuszczania. Ponieważ klingi mają większe pole przekroju przy rękojeści niż przy wierzchołku, zmienia się również prędkość chłodzenia podczas hartowania, która jest największa przy wierzchołku, co powoduje powstawanie w tym obszarze większej ilości martenzytu.



Rys. 11. Wyniki pomiarów twardości na zglądach poprzecznych badanych kling

W przypadku klingi S-4 również zauważalny jest spadek twardości w kierunku rękojeści, a najniższą twardość ponownie zaobserwowano bezpośrednio przy rękojeści. Co ważne, twardość w tym miejscu na obu klingach jest bardzo zbliżona.

Przed wszystkim stwierdzono wyższą twardość tej klingi, co potwierdza zarówno wyniki wcześniejszych pomiarów parametru MP (rys. 2d), jak i wyniki obserwacji metalograficznych (rys. 6). Dość zaskakujące jest to, że klinga S-4, której analiza składu chemicznego wykazała najniższą zawartość węgla (por. tab. 1), cechuje się najwyższymi twardościami na niemal całej długości. Zapewne jest to efekt obróbki cieplnej, której ta klinga została poddana, tzn. prawdopodobnie została odpuszczona przy niższych temperaturach od kling pochodzących od producenta 1. Spowodowało to z jednej strony osiągnięcie wyższych twardości po obróbce cieplnej, ale mogło to się wiązać ze spadkiem udarności i spowodować jej przedwczesne zniszczenie.

5. Wnioski

- 1) Składy chemiczne badanych kling szermierczych są zbliżone do składu klasycznej stali sprężynowej 60S2A, ale z obniżoną zawartością węgla. Zawartość węgla w stali, z której wykonano klingę S-4 (0,25%) pochodzącą od producenta 2, jest wyraźnie niższa od pozostałych, co wskazuje na różne materiały stosowane do wyrobu kling.
- 2) Porównując zmiany parametru MP na klingach tego samego producenta po różnych etapach eksploatacji, można stwierdzić, że parametr MP zmienia się w trakcie użytkowania klingi. Skoki w części klingi oddalonej od rękojeści zanikają, natomiast bliżej niej pojawiają się nowe. Najprawdopodobniej jest to skutkiem procesów związanych z relaksacją naprężeń i pojawianiem się mikroodkształceń wynikających z uderzania klingi o klingę.
- 3) Na podstawie analizy SEM powierzchni badanych kling można stwierdzić, że przyczyną zużywania kling, zwłaszcza przy ich krawędzi, są mikroodkształcenia powstające na skutek uderzeń o siebie broni. Powodują one powstawanie struktury „łusek” i rozwarstwień w pobliżu krawędzi klingi, które mogą być miejscami зароdkowania pęknięć.
- 4) W mikrostrukturze badanych kling, niemal na całej ich długości, stwierdzono występowanie martenzytu odpuszczonego. Różnice w mikrostrukturze pojawiają się dopiero przy samej rękojeści, gdzie poza klingą S-2 wystąpiła mikrostruktura ferrytyczno-perlityczna. Powyższe obserwacje potwierdzają również pomiary twardości. Oznacza to, że podczas hartowania ten obszar klingi najprawdopodobniej nie był poddawany bezpośredniemu działaniu czynnika chłodzącego. W przypadku klingi S-2 (w odróżnieniu od kling S-1 i S-3) w tym obszarze wystąpiła struktura martenzytyczna, co oznacza, że producent tej klingi stosuje odstępstwa od wyżej opisanej technologii.
- 5) Pomiędzy klingami S-1, S-2 i S-3, pochodzącymi od producenta 1 a klingą S-4 pochodzącą od producenta 2 występują różnice w morfologii martenzytu oraz różnice w jego twardości. Oznacza to, że poza innym materiałem producent 2 zastosował najprawdopodobniej inny wariant obróbki cieplnej, polegający na obniżeniu temperatury odpuszczania.

Literatura

- Ji Y., Liu Z., Ren H., 2011, *Morphology and formation mechanism of martensite in steels with different carbon content*, Advanced Materials Research, vol. 201–203, s. 1612–1618. <https://doi.org/10.4028/www.scientific.net/AMR.201-203.1612>.
- Jiang H., Shen J., Yao X., Van Horne C., Lu X., Xiong Y., Cha L., 2022, *En-garde! A review of fencing blade material development*, Metals (Basel), vol. 12(2), 236. <https://doi.org/10.3390/met12020236>.

- Kibaroglu D., Baydogan M., Cimenoglu H., Bas B., Yagsi C., Aliyeva N., 2017, *Failure analysis of fencing blades*, Journal of Physics: Conference Series, vol. 843, 012009. <http://doi.org/10.1088/1742-6596/843/1/012009>.
- Lee S., Lee Y., 2008, *Prediction of austenite grain growth during austenitization of low alloy steels*, Materials & Design, vol. 29, s. 1840–1844. <https://doi.org/10.1016/j.matdes.2008.03.009>.
- Nahak B., 2017, *Material characterization using Barkhausen noise analysis technique – a review*, Indian Journal of Science and Technology, vol. 10(14), s. 1–10. <https://doi.org/10.17485/ijst/2017/v10i14/109697>.
- PN-H-84032:1974, *Stal sprężynowa (resorowa) – Gatunki*, Polski Komitet Normalizacyjny, Warszawa.
- Roi G.S., Bianchedi D., 2008, *The science of fencing: implications for performance and injury prevention*, Sports Medicine, vol. 38(6), s. 465–481. <https://doi.org/10.2165/00007256-200838060-00003>.
- Smallman R.E., Bishop R.J., 1999, *Modern Physical Metallurgy and Materials Engineering: Science, Process, Applications*, 6th ed., Butterworth-Heinemann. <https://doi.org/10.1016/B978-0-7506-4564-5.X5000-9>.
- Stresstech.com, *Stresstech Bulletin 2: The properities of Barkhausen noise*. <https://www.stresstech.com/stresstech-bulletin-2-the-properities-of-barkhausen-noise/> [dostęp: 26.06.2023].
- Tota J., Kokosza A., Pacyna J., 1994, *Badania porównawcze naprężeń po szlifowaniu aparatem Stresscan 500C*, [w:] *XVII Naukowa Szkoła Obróbki Ściernej: Kraków – Myślenice, 8–10 września 1994: Referaty*, Politechnika Krakowska, Kraków, s. 339–346.
- Vivaldi F., Cortese L., Coppola T., Mazzarano A., Nalli F., 2018, *A study on the dynamic structural behavior of Olympic sabres*, Procedia Structural Integrity, vol. 8, s. 345–353. <https://doi.org/10.1016/j.prostr.2017.12.035>.
- Zhang C., Zhou L., Liu Y., 2013, *Surface decarburization characteristics and relation between decarburized types and heating temperature of spring steel 60Si2MnA*, International Journal of Minerals, Metallurgy, and Materials, vol. 20, s. 720–724. <https://doi.org/10.1007/s12613-013-0789-1>.

Kosmiczne materiały – platery

Karol Kuglarz , Grzegorz Michta 

AGH Akademia Górniczo-Hutnicza, Wydział Inżynierii Metali i Informatyki Przemysłowej, Kraków

Streszczenie: Artykuł opisuje zalety technologii platerowania wybuchowego oraz jej spójność z aktualnie panującym nurtem w rozwoju materiałów. Zgrzewanie wybuchowe pozwala przede wszystkim na łączenie ze sobą materiałów, których nie jesteśmy w stanie połączyć innymi metodami spajania. W pracy zaprezentowano badania mikrostrukturalne dziesięciowarstwowego plateru złożonego z trzech materiałów: żelaza Armco, Ti Gr. 1 i Al 1050. Omówiono zalety każdego z materiałów i uzasadniono jego użycie. Pokazano także metodykę badań materiału wielowarstwowego. W pierwszej kolejności pokazano wpływ zgrzewania wybuchowego na grubość każdej z warstw, a także zaprezentowano wyniki pomiarów podstawowych parametrów złącza platerowanego. Wyznaczono również współczynnik RGP, który pozwolił określić jakość wykonanego połączenia. W dalszej części przedstawiono wyniki badań mikrostruktury, które obrazują charakterystyczne cechy złącza zgrzewanego wybuchowo. Na koniec wskazano na możliwość zastosowania takiego plateru.

Słowa kluczowe: materiały platerowane, zgrzewanie wybuchowe, kompozyt metaliczny, materiał wielowarstwowy, odporność balistyczna

SPACE MATERIALS – CLADDING MATERIALS

Abstract: The paper presents the advantages of explosive cladding technology and its consistency with the current trend in materials development. Explosive cladding allows, first of all, to join materials that we are not able to join with other methods. The paper shows microstructural studies of ten-layer plating composed of three materials: Armco iron, Ti Gr. 1 and Al 1050. Each of the materials is discussed in terms of its advantages and its use is justified. The methodology of multi-layer material testing was also shown. First of all, the impact of explosive welding on the thickness of each layer is reported, as well as the results of measurements of the basic parameters of the clad joint are displayed. The RGP factor was also determined, which allowed to determine the quality of the connection made. The following section presents the results of microstructure tests, which show the characteristics of the explosive welded joint. Finally, a possible application of such plater is showed.

Keywords: cladding materials, explosive cladding, metallic composite, multi-layer material, ballistic resistance

1. Wstęp

Rozwój konstrukcji w wielu dziedzinach przemysłu, m.in. lotniczym, energetycznym i zbrojeniowym, wiąże się z rosnącymi wymaganiami wobec materiałów. Stawia to przed inżynierami wyzwanie projektowania materiałów coraz bardziej zaawansowanych i posiadających nie tylko lepsze parametry wytrzymałościowe, ale także zróżnicowane właściwości, w tym cechy funkcjonalne. Coraz częściej wymaga się, aby materiał spełniał kilka funkcji jednocześnie przy jak najmniejszej masie. Pozwala to uzyskać oszczędności na etapie produkcyjnym dzięki zmniejszonemu zużyciu materiału, a także na etapie eksploatacyjnym dzięki zmniejszeniu masy konstrukcji.

Bardzo dobrze odpowiada tym wymaganiom technologia platerowania wybuchowego pozwalająca na wytworzenie wielowarstwowych materiałów metalicznych. Posiada ona zalety, które idealnie wpasowują się w nurt rozwoju technologii wytwarzania materiałów zarówno tych konwencjonalnych, jak i do zastosowań specjalnych – w tym kosmicznych. Zgrzewanie wybuchowe pozwala na łączenie materiałów, których nie można połączyć inną metodą. Dzięki temu można na przykład uzyskiwać połączenia metali lekkich, jak magnez, aluminium czy tytan, oraz ich stopów. Ponadto możemy łączyć te materiały w szerokim zakresie ich grubości, a także powierzchni. W pewnych przypadkach pozwala to na obniżenie kosztów dzięki zmniejszeniu zużycia drogich metali. Następną ważną zaletą technologii zgrzewania wybuchowego jest możliwość wykonania łączenia w procesie symultanicznym, czyli łączenia ze sobą wielu warstw materiału w jednym akcie strzałowym. Wprowadzenie warstw o zróżnicowanych i jednocześnie jednoznacznie określonych właściwościach pozwala uzyskać nowe oczekiwane cechy funkcjonalne i konstrukcyjne materiału platerowanego, przy zachowaniu właściwości technologicznych, korozyjnych, mechanicznych i prądowych materiałów składowych (Walczak 1989, Dya i in. 2001, Gałka 2015). Otrzymane w ten sposób nowe, unikatowe właściwości platerów umożliwiają ich zastosowanie w konstrukcjach urządzeń, w których do tej pory nie były stosowane.

W pracy przedstawiono badania mikrostrukturalne dziesięciowarstwowego plateru złożonego z trzech materiałów: żelaza Armco, Ti Gr. 1 i Al 1050.

2. Materiał użyty do badań

Przedmiotem badań był dziesięciowarstwowy plater w postaci blachy powstały metodą wybuchową, o końcowej grubości 7,66 mm, składający się z trzech materiałów. Materiał ten został opracowany przez firmę Explomet, a technologia jego wytworzenia została zastrzeżona. Jednak kluczowymi parametrami procesu były: ustawienie odpowiedniej odległości pomiędzy poszczególnymi warstwami materiałów składowych w układzie strzałowym i dobór odpowiedniej mieszanki wybuchowej. Parametry te

dobrano w taki sposób, by zagwarantować odpowiednią prędkość i siłę detonacji, co miało wpływ na pęd i energię zderzenia łączonych blach. Warstwy zewnętrzne stanowiło żelazo Armco, które zaliczane jest do grupy stali niskowęglowych magnetycznie miękkich. Dużą zaletą tego materiału jest jego czystość chemiczna i metalurgiczna (praktycznie brak naturalnych zanieczyszczeń). Dzięki temu materiał cechuje się bardzo dobrymi właściwościami magnetycznymi, podwyższoną odpornością na korozję, bardzo dobrą podatnością na obróbkę plastyczną na zimno oraz bardzo dobrą spawalnością. Wszystkie te cechy są bardzo pożądane w tym metalicznym kompozycie. Skład chemiczny żelaza Armco przedstawiono w tabeli 1 (Zumbilev 2016).

Tabela 1
Skład chemiczny żelaza Armco [% wag.]

Fe	C	Si	Mn	Cr	Ni	S	P	Mo
reszta	0,02	0,01	0,07	0,02	0,03	0,002	0,002	0,02

Kolejnym materiałem użytym do wytworzenia plateru był stop aluminium Al 1050. Jego skład chemiczny został przedstawiony w tabeli 2. Podobnie jak żelazo Armco jest to materiał o wysokim stopniu czystości. Cechuje się dobrą odpornością na korozję, podatnością na obróbkę plastyczną na zimno, wysoką przewodnością cieplną oraz dobrą spawalnością (Esmaeili i in. 2011).

Tabela 2
Skład chemiczny stopu Al 1050 [% wag.]

Al	Ti	Zn	Mn	Fe	Si	Mg	Cu
reszta	0,05	0,07	0,05	0,4	0,25	0,05	0,05

Ostatnim materiałem użytym do wytworzenia tego plateru był Ti Gr. 1. Ten niestopowy tytan jest podatny na obróbkę plastyczną i może być formowany na zimno. Ponadto jest bardzo dobrze spawalny i cechuje się wysoką udurocznością. Doskonale nadaje się do zastosowań w konstrukcjach o obniżonej masie i wysokiej odporności na korozję. Jego niska rozszerzalność cieplna powoduje, że naprężenia termiczne są niższe niż w innych materiałach metalicznych. Skład chemiczny tego materiału został przedstawiony w tabeli 3 (Prazmowski i in. 2017).

Tabela 3
Skład chemiczny Ti Gr. 1 [% wag.]

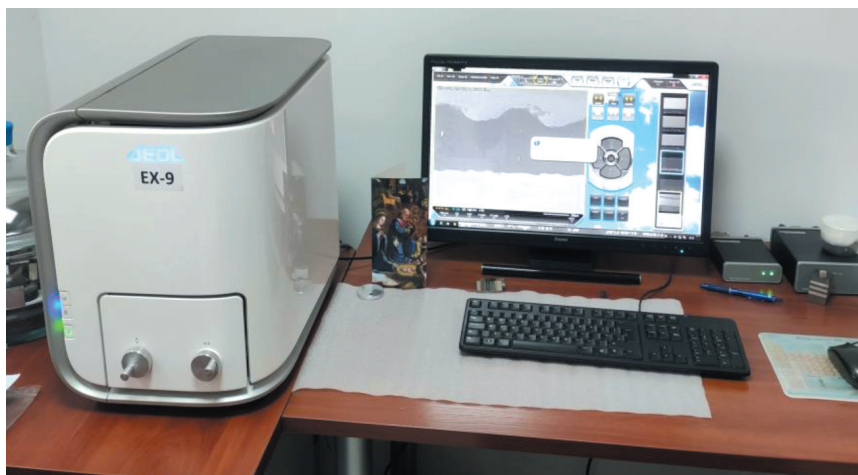
Ti	C	Fe	N	O	H
reszta	0,08	0,02	0,03	0,18	0,015

W trakcie procesu platerowania materiał podstawowy stanowiło żelazo Armco o grubości 1,5 mm. Następnie na niego nałożony został tytan o grubości 0,8 mm a w dalszej kolejności aluminium o grubości 0,5 mm. W sumie naprzemiennych warstw tytan – aluminium było osiem. Ostatnią nakładaną warstwę stanowiło żelazo Armco również o grubości 1,5 mm.

Otrzymany plater w dalszych procesach przetwarzania poddany został walcowaniu do grubości końcowej 2,7 mm oraz w kolejnych krokach celowym obciążeniom cieplnym oraz formowaniu w oryginalnej, współautorskiej technologii kształtującej jego wyjątkowe właściwości i cechy funkcjonalne.

3. Metodyka badań

Z otrzymanej platerowanej blachy przy użyciu wody wycięto próbkę o wymiarach 25 mm × 25 mm. Taki rodzaj cięcia zapobiegł nagrzewaniu i zmianom mikrostruktury w badanym materiale. Następnie próbkę poddano procesom szlifowania oraz polerowania i wstępnym obserwacjom przy użyciu mikroskopu skaningowego JEOL 6000 (rys. 1). Badania te miały na celu wyznaczenie grubości poszczególnych warstw materiału po procesie platerowania wybuchowego, które następnie zostały zestawione z grubościami tych materiałów przed procesem platerowania (tab. 6).



Rys. 1. Mikroskop skaningowy JEOL 6000

Po wykonaniu badań wstępnych próbka została zainkludowana, wyszlifowana i wypolerowana. Dużą trudność stanowiło trawienie wielowarstwowego plateru, gdyż każdy z użytych materiałów musiał być trawiony innym odczynnikiem chemicznym. Skład

chemiczny odczynnik użytego do trawienia tytanu przedstawiono w tabeli 4, a odczynnik do trawienia aluminium – w tabeli 5. Żelazo Armco wytrawiono Nitałem o stężeniu 1%.

Tabela 4

Skład chemiczny odczynnika do trawienia tytanu [ml]

HNO ₃	HF	H ₂ O
2	2	96

Tabela 5

Skład chemiczny odczynnika do trawienia aluminium [ml]

H ₂ SO ₄	HF	H ₂ O
10	5	85

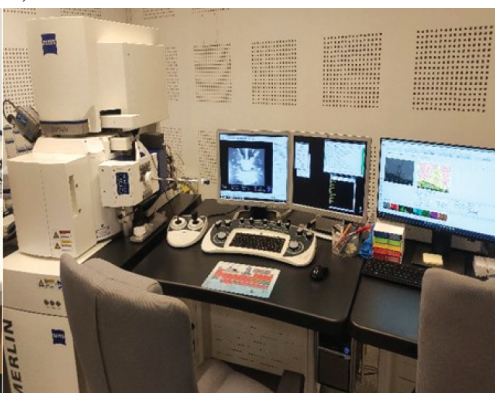
Tak wytrawiona próbka została poddana badaniom mikrostrukturalnym przy użyciu: mikroskopu świetlnego (*light microscope*, LM) oraz skaningowego mikroskopu elektronowego (*scanning electron microscope*, SEM).

W badaniu przy użyciu mikroskopu świetlnego Axio Imager M1m firmy Zeiss (rys. 2a) zastosowano różne techniki obrazowania, takie jak: jasne pole widzenia oraz kontrast Nomarskiego (*differential interference contrast*, DIC). W przypadku skaningowego mikroskopu elektronowego Merlin Gemini II firmy Zeiss (rys. 2b), który ma wbudowany system mikroanalizy EDS (*energy-dispersive X-ray spectroscopy*), również zastosowano różne techniki obrazowania, takie jak: elektrony wtórne (*secondary electron*, SE) i elektrony wstecznie rozproszone (*backscattered electron*, BSE).

a)



b)



Rys. 2. Mikroskopy użyte do badań metalograficznych: a) mikroskop świetlny firmy Zeiss; b) skaningowy mikroskop elektronowy firmy Zeiss

Wykonane podczas obserwacji mikroskopowych zdjęcia posłużyły do pomiaru podstawowych parametrów złącza platerowanego, takich jak: długość fali, wysokość fali oraz pole powierzchni przetopień. Pomiary te wykonano przy użyciu programu ImageJ, a ich wyniki posłużyły także do wyznaczenia współczynnika równoważnej grubości przetopień RGP. Ponadto zdjęcia wykorzystano do określenia średniej wielkości ziarna. W tym celu za pomocą programu ImageJ zmierzono powierzchnię 30 ziaren, a następnie obliczono średnią.

Współczynnik RGP jest to parametr opracowany przez firmę Explomet i stosowany do oceny poprawności zastosowanych parametrów zgrzewania wybuchowego. Oblicza się go, korzystając ze wzoru:

$$RGP = \frac{\sum A}{L} [\mu\text{m}] \quad (1)$$

gdzie:

A – sumaryczne pole powierzchni przetopień [mm^2],

L – całkowita długość złącza [mm].

Przyjmuje się, że w poprawnie wykonanym złączu platerowanym wartość RGP nie powinna przekraczać 12 μm .

Należy wspomnieć, że żelazo Armco i tytan wytrawiły się w zadowalający sposób, natomiast aluminium najprawdopodobniej uległo pasywacji i jego wytrawienie nie było zadowalające. Trzeba przeprowadzić ponowną preparatykę i dopracować trawienie stopu aluminium – trawić świeżo po polerowaniu lub trawić przed użyciem odczynników trawiących żelazo Armco i tytan.

4. Badania złącza platerowanego

W pierwszej kolejności zmierzono grubość poszczególnych warstw po procesie platerowania. Nominalnie żelazo powinno mieć 1,5 mm, tytan 0,8 mm, a aluminium 0,5 mm. Jednak proces platerowania spowodował niewielkie zmniejszenie grubości blach. Wyniki pomiarów grubości przedstawiono w tabeli 6 (każdy wynik jest średnią z sześciu punktów pomiarowych).

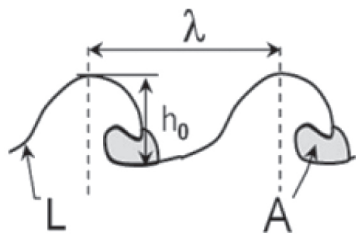
Z przeprowadzonych pomiarów wynika, że docisk generowany w procesie platerowania wybuchowego spowodował zmniejszenie grubości plateru o 0,54 mm $\pm$ 0,09 mm.

Następnie wykonano pomiary złącza platerowanego zgodnie ze schematem przedstawionym na rysunku 3.

Najpierw zmierzono wysokość i długość fali. Wyniki przedstawiono w tabeli 7 (każdy wynik jest średnią z pięciu punktów pomiarowych).

Tabela 6
Grubość poszczególnych warstw po procesie platerowania

Warstwa		Grubość po platerowaniu [mm]	Grubość nominalna [mm]
Materiał nakładany	Fe Armco	1,410	1,500
	Al 1050	0,448	0,500
	Ti Gr. 1	0,765	0,800
	Al 1050	0,434	0,500
	Ti Gr. 1	0,754	0,800
	Al 1050	0,466	0,500
	Ti Gr. 1	0,778	0,800
	Al 1050	0,464	0,500
	Ti Gr. 1	0,715	0,800
Materiał podstawowy	Fe Armco	1,427	1,500
	Suma [mm]		
		7,66	8,2



Rys. 3. Podstawowe parametry złącza platerowanego: λ – długość fali, h_0 – wysokość fali, L – długość linii złącza, A – pole powierzchni przetopień
Źródło: Kowalski (2014)

Tabela 7
Wyniki pomiarów wysokości i długości fali

Złącze	h_0 [μm]	λ_0 [μm]
Fe-Al	18	100
Al-Ti	141	655
Ti-Al	27	214
Al-Ti	82	590
Ti-Al	21	143
Al-Ti	79	446
Ti-Al	13	103
Al-Ti	68	326
Ti-Fe	182	833

Na podstawie otrzymanych wyników można stwierdzić, że fala w złączu Fe-Al jest najniższa i zarazem najkrótsza. Następnie nieznacznie wyższa i dłuższa jest fala w połączeniu Ti-Al. Dalej w kolejności są wymiary złącza Al-Ti, które ma dłuższą i wyższą falę niż złącze Fe-Al i Ti-Al. Natomiast najdłuższą i zarazem najwyższą falę uzyskano w połączeniu tytanu i żelaza. Ponadto możemy zauważyć, że w przypadku tych samych złączy ich wymiary maleją wraz z oddaleniem się od materiału nakładanego. Może być to związane z malejącą siłą uderzenia materiału napędzanego (nakładanego).

Następnie wyznaczono wartość współczynnika RGP. Wyniki pomiarów badanego złącza wybuchowego zostały przedstawione w tabeli 8.

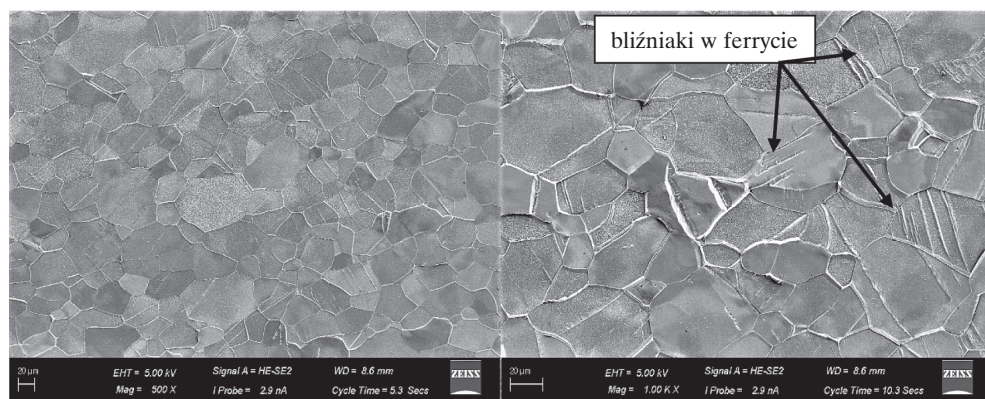
Tabela 8
Wartość współczynnika RGP dla poszczególnych złączy

Złącze	Średnie RGP [μm]
Fe-Al	2,34
Ti-Al	5,97
Fe-Ti	5,12

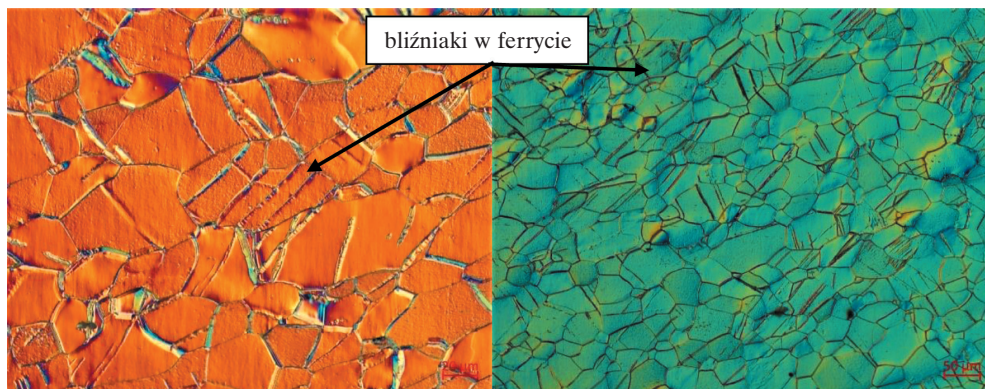
Otrzymane wyniki wynoszą znacznie poniżej $12 \mu\text{m}$. Dlatego można stwierdzić, że proces platerowania został poprawnie przeprowadzony.

5. Badania mikrostruktury

W celu oceny wpływu procesu platerowania, a także sprawdzenia jakości złącza wybuchowego na próbce przeprowadzono obserwację mikrostruktury. Badania zaczęto od obserwacji poszczególnych warstw materiałów składowych plateru. W pierwszej kolejności zbadano żelazo Armco. Przykładowe zdjęcia mikrostruktury żelaza zostały przedstawione na rysunkach 4 i 5.



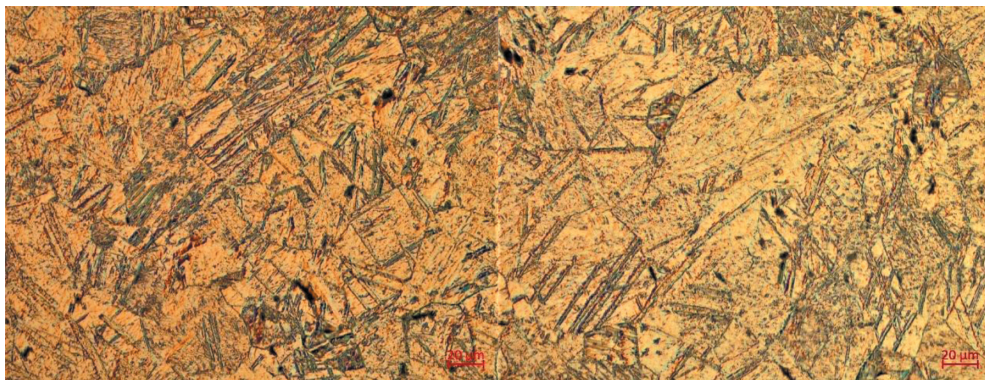
Rys. 4. Przykładowa mikrostruktura żelaza Armco, zgląd trawiony, SEM



Rys. 5. Przykładowa mikrostruktura żelaza Armco, zgład trawiony, LM, kontrast Nomarskiego

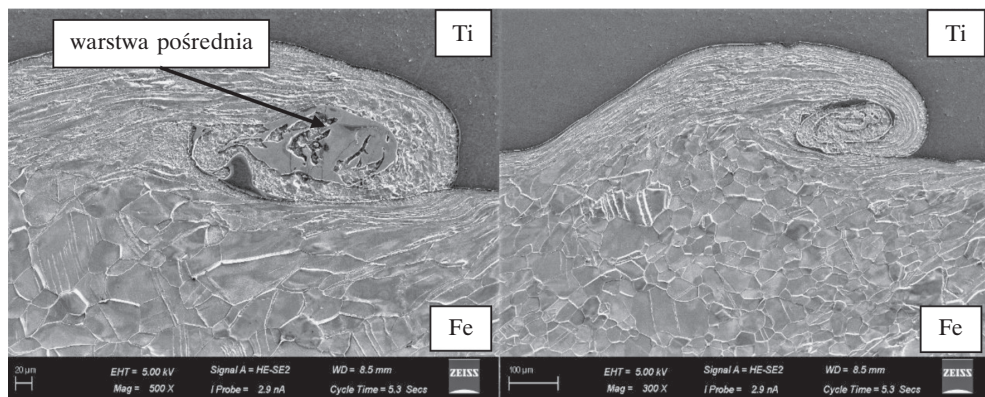
W trakcie badań mikrostruktury przy użyciu programu ImageJ zmierzono również średnią powierzchnię ziarna ferrytu, która wyniosła $1360 \mu\text{m}^2$. Ciekawostką jest wystąpienie w ferrycie wyraźnie widocznych bliźniaków. Są to struktury powstające w ściśle określonych warunkach procesu: duże ziarno, duży gradient temperatury i duża szybkość odkształcenia (Osuch 2010, Fronczek i in. 2016, Pocica i Tuz 2016). Wszystkie te czynniki charakteryzują proces zgrzewania wybuchowego, podczas którego mamy do czynienia z bardzo szybkim odkształceniem i dużym gradientem temperatur. Ponadto pomiary powierzchni ziarna pokazały, że warunek dużego ziarna również został spełniony.

Następnie wykonano badania mikrostruktury tytanu. Wyniki przedstawiono na rysunku 6, na którym dobrze widoczne są ziarna tytanu. Ponadto w mikrostrukturze można zaobserwować czarne obszary, które powstały w wyniku podtrawienia próbki.

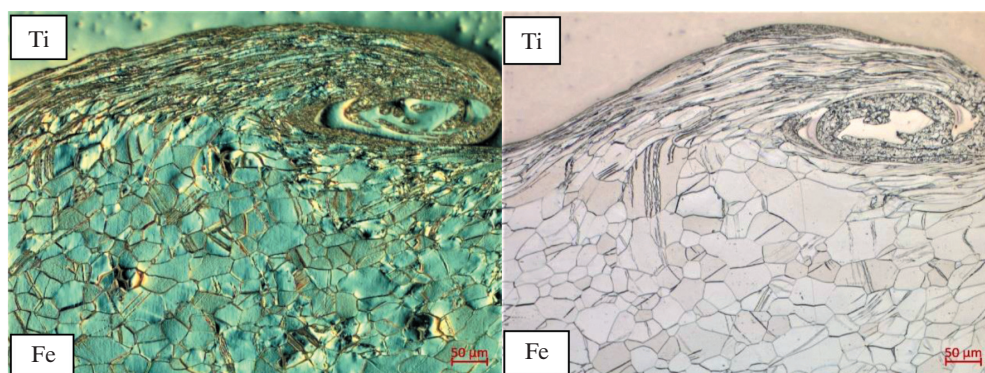


Rys. 6. Przykładowa mikrostruktura tytanu, zgład trawiony, LM, kontrast Nomarskiego

W następnym kroku obserwowano połączenia pomiędzy poszczególnymi materiałami składowymi badanego plateru. Na początku poddano obserwacji mikrostrukturę połączenia wybuchowego tytan-żelazo. Wyniki przedstawiono na rysunkach 7 i 8.



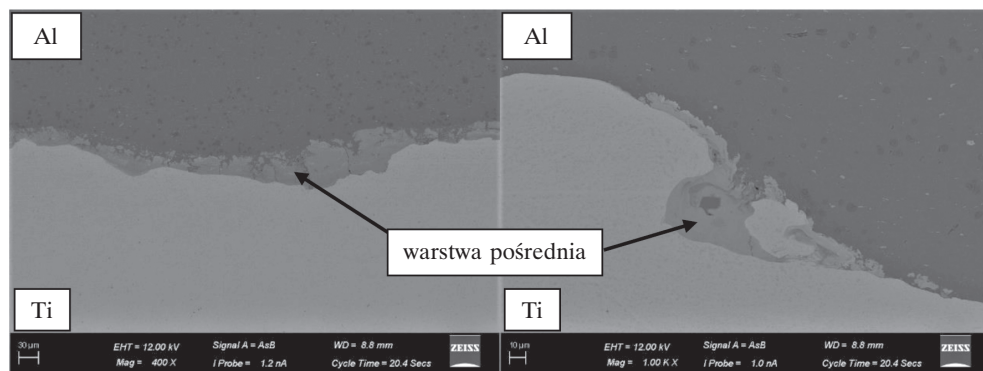
Rys. 7. Przykładowa mikrostruktura złącza Fe-Ti, zgląd trawiony, SEM



Rys. 8. Przykładowa mikrostruktura złącza Fe-Ti, zgląd trawiony, LM, kontrast Nomarskiego

Wyraźnie widoczny jest charakterystyczny dla procesu platerowania wybuchowego falisty charakter połączenia z warstwą pośrednią, której skład chemiczny zależy od materiałów, pomiędzy którymi występuje. Powstaje ona w wyniku przejścia przez stan ciekły materiałów łączonych. Warstwa pośrednia może powstawać na dwa sposoby: jako nowa faza, która będzie cechowała się znaczną twardością i kruchością, lub jako mieszanina drobnodyspersyjna, powstająca w wyniku nadtopienia warstw wierzchnich platerowanych metali (Kuglarz 2023).

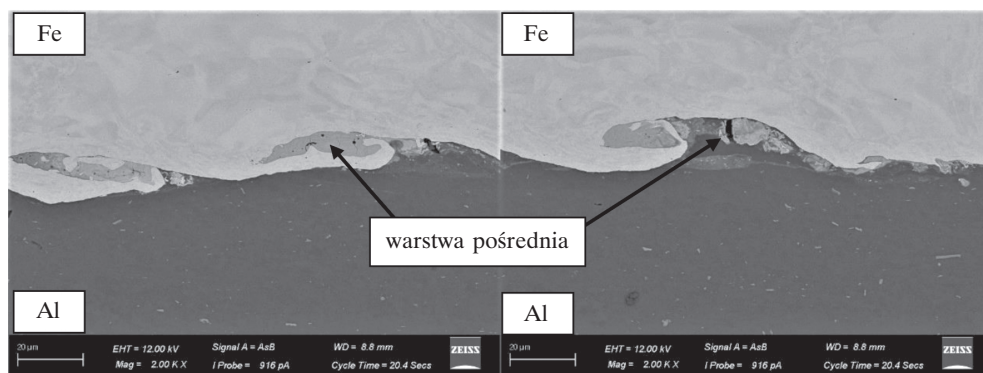
Następnie w tej samej próbce, lecz pomiędzy innymi warstwami, przeprowadzono obserwację złącza wybuchowego Al-Ti. Wyniki tego badania przedstawiono na rysunku 9.



Rys. 9. Przykładowa mikrostruktura złącza Al-Ti, zgląd trawiony, SEM

Podobnie jak w przypadku złącza Fe-Ti tu również jest dobrze widoczny falisty charakter połączenia oraz powstała warstwa pośrednia.

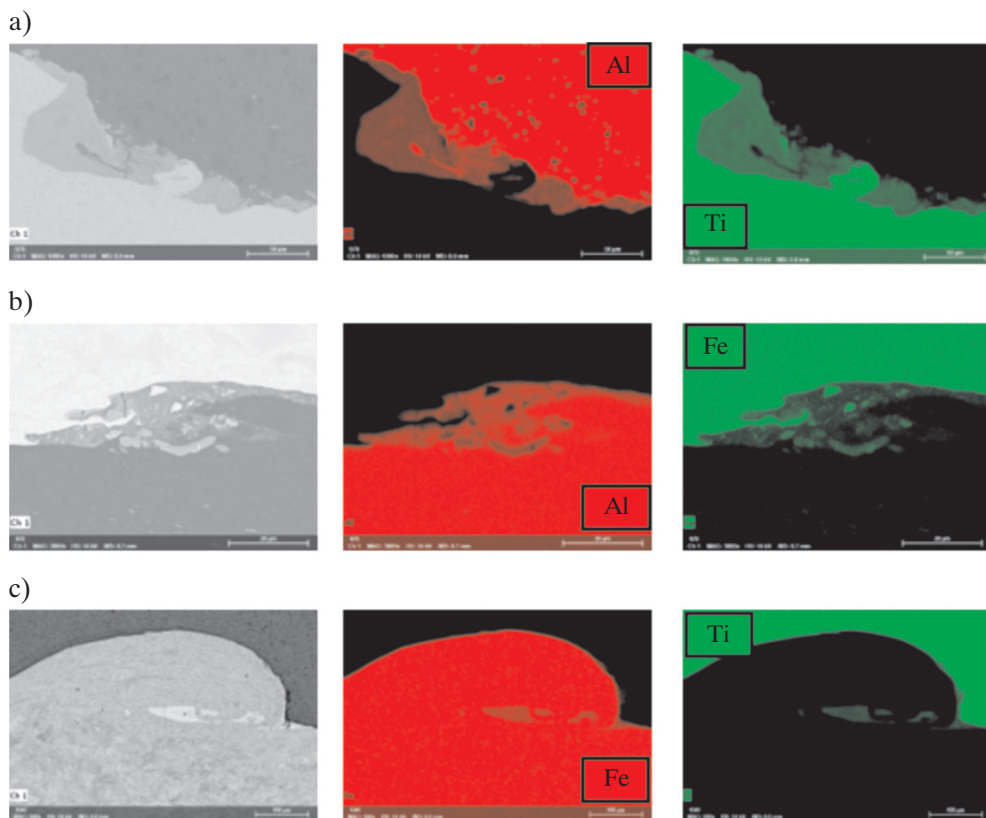
Na ostatnim etapie, kontynuując badania tej samej próbki, przebadano trzeci rodzaj połączenia, czyli złącze Fe-Al. Rezultaty obserwacji przedstawiono na rysunku 10.



Rys. 10. Mikrostruktura złącza Fe-Al, zgląd trawiony, SEM

W złączy Fe-Al tak jak w poprzednich złączach widoczny jest falisty charakter połączenia oraz powstała warstwa pośrednia.

Za pomocą rentgenowskiej spektroskopii promieniowania charakterystycznego z dyspersją energetyczną EDX przeprowadzono także analizy złącz zgrzewanych wybuchowo. Badanie to miało na celu analizę składu chemicznego pod kątem jakościowym poszczególnych warstw plateru. Na rysunku 11 przedstawiono wynik analizy EDX w postaci map jakościowych pierwiastków występujących w badanych obszarach.



Rys. 11. Rozkład powierzchniowy pierwiastków:
a) złącze Al-Ti; b) złącze Fe-Al; c) złącze Fe-Ti

Pierwiastki występujące na uzyskanych mapach jakościowych analizowanych obszarów to odpowiednio: żelazo, aluminium i tytan. W linii łączenia poszczególnych blach występują miejsca o mniej intensywnych kolorach przypisanych do poszczególnych pierwiastków, co świadczy o obecności warstwy pośredniej o składzie odpowiadającym związkom chemicznym powstałym z poszczególnych pierwiastków. Analizując otrzymane wyniki, można stwierdzić, że spośród dwóch materiałów tworzących złącze większemu odkształceniu ulega ten o mniejszej twardości. Ponadto odnosząc się do innych dostępnych badań (Paul i in. 2011), można powiedzieć, że stopień wymieszania pierwiastków w złączu platerowanym może być różny. Zależy to m.in. od grubości warstwy międzymetalicznej w danym miejscu. Warto zwrócić uwagę, że różny stopień wymieszania pierwiastków może powodować powstanie różnych związków międzymetalicznych, a te z kolei mogą w odmienny sposób wpływać na właściwości całego plateru.

6. Zastosowanie wytworzonego plateru

Wytworzony plater jest pierwszym tego typu materiałem na świecie, dlatego też jego unikatowe właściwości będą jeszcze badane zarówno w formie przed obróbką cieplną, jak i po specjalistycznej obróbce cieplno-plastycznej. Oczekuje się, że po tych zabiegach materiał będzie się cechował bardzo dobrą wytrzymałością mechaniczną i – co ważne – odpornością balistyczną. Dzięki temu będzie mógł znaleźć zastosowanie szczególnie w aeronautyce jako poszycie rakiet, sond kosmicznych, czy łazików stosowanych do badań planetarnych.

7. Podsumowanie

W pracy zaprezentowano wyniki badań mikrostrukturalnych innowacyjnego dieściowarstwowego plateru składającego się z żelaza Armco, Ti Gr. 1 i Al 1050. Przeprowadzone obserwacje i pomiary pozwoliły jednoznacznie stwierdzić, że plater został wykonany poprawnie i nadaje się do dalszych etapów obróbki cieplnej oraz cieplno-plastycznej. Ponadto zdjęcia mikrostruktury ujawniły cechy charakterystyczne dla procesu zgrzewania wybuchowego, takie jak falisty charakter połączenia oraz zgmiot ziarna w pobliżu strefy złącza platerowanego. Badania mikrostrukturalne ujawniły również bliźniaki występujące w ferrycie, które powstały na skutek charakterystycznych dla tego procesu zjawisk. Przeprowadzone na koniec badania EDX umożliwiły obserwację wymieszania pierwiastków i powstawanie warstw pośrednich. Wstępne wyniki zaprezentowanych badań świadczą o tym, że perspektywy dalszego rozwoju tego materiału są obiecujące, a po osiągnięciu założonych parametrów ma on dużą szansę na zastosowanie w różnych dziedzinach przemysłu, w tym w aeronautyce czy przemyśle kosmicznym.

Praca zrealizowana w ramach współpracy AGH z firmą EXPLOMET Gałka, Szulc spółka komandytowa i w ramach prac statutowych AGH nr 16.16.110.663. Wykonana dzięki wsparciu badawczemu Międzynarodowego Centrum Mikroskopii Elektronowej dla Inżynierii Materiałowej IC-EM AGH.

Literatura

- Dyja H., Maranda A., Trębiński R., 2001, *Technologie wybuchowe w inżynierii materiałowej*, Wydawnictwo Wydziału Metalurgii i Inżynierii Materiałowej Politechniki Częstochowskiej, Częstochowa.
- Esmaeili A., Zareie Rajni H.R., Sharbati M., Besharati Givi M.K., Shamanian M., 2011, *The role of rotation speed on intermetallic compounds formation and mechanical behavior of friction stir welded brass/aluminum 1050 couple*, Intermetallics, vol. 19(11), s. 1711–1719. <https://doi.org/10.1016/j.intermet.2011.07.006>.

- Fronczek D.M., Wojewoda-Budka J., Chulist R., Sypien A., Korneva A., Szulc Z., Schnell N., Zieba P., 2016, *Structural properties of Ti/Al clad manufactured by explosive welding and annealing*, Materials & Design, vol. 91, s. 80–89. <https://doi.org/10.1016/j.matdes.2015.11.087>.
- Gałka A., 2015, *Zastosowanie technologii wybuchowego platerowania metali do wytwarzania nowych zaawansowanych materiałów warstwowych na przykładzie połączenia tytan Ti6Al4V – aluminium AA2519*, Materiały Wysokoenergetyczne, t. 7, s. 73–79.
- Kowalski M., 2014, *Zjawiska degradacyjne w bimetalu stal-tytan przy obciążeniach cyklicznych*, Politechnika Opolska. Wydział Mechaniczny, Opole [rozprawa doktorska].
- Kuglarz K., 2023, *Właściwości złącza spawanego trójwarstwowego plateru 410S – Cu-DHP – 316L*, WIMiP AGH, Kraków [projekt inżynierski].
- Osuch W., 2010, *Bliźniaki przemiany w stalach niskowęglowych*, Wydawnictwa AGH, Kraków.
- Paul H., Faryna M., Prazmowski M., Bański R., 2011, *Changes in the bonding zone of explosively welded sheets*, Archives of Metallurgy and Materials, vol. 56(2), s. 463–474.
- Pocica N., Tuz L., 2016, *Ocena mikrostruktury i wybranych własności mechanicznych złączy zgrzewanych wybuchowo po obróbce cieplnej*, Przegląd Spawalnictwa, r. 88(4), s. 35–38.
- Prazmowski M., Najwer M., Paul H., Andrzejewski D., 2017, Influence of explosive welding parameters on properties of bimetal Ti-carbon steel, MATEC Web of Conferences, vol. 94, 02012. <https://doi.org/10.1051/mateconf/20179402012>.
- Walczak W., 1989, *Zgrzewanie wybuchowe metali i jego zastosowania*, Wydawnictwa Naukowo-Techniczne, Warszawa.
- Zumbilev I.A., 2016, *About the influence of ion carbonitriding on the chemical composition of the carbonitride zone*, Journal of Food and Packaging Science, Technique and Technologies, no. 8, s. 10–13.

On graph models in knowledge engineering

Antoni Ligeża¹ , Weronika T. Adrian¹ , Marek Adrian¹ ,
Krzysztof Kluza¹ , Krystian Jobczyk¹ , Piotr Wiśniewski¹ ,
Mateusz Ślażyński¹ , Paweł Jemioło¹ , Dominik Sepioło¹ ,
Bernadetta Stachura-Terlecka¹ , Mateusz Zaremba¹ ,
Mateusz Szymkowski¹ , Anna Suchenia² , Tomasz Potempa³ 

¹ AGH University of Krakow, Faculty of Electrical Engineering, Automatics, Computer Science and Biomedical Engineering, Krakow

² Cracow University of Technology, Faculty of Electrical and Computer Engineering, Krakow

³ University of Applied Sciences, Polytechnic Faculty in Tarnow, Tarnów

Abstract: The paper presents selected applications of graph models in knowledge engineering. Starting from the basic definition of a graph as a set of nodes connected by edges, the article presents possible extensions of this concept aimed at increasing the power of expression and the ability to process knowledge. In particular, the work focuses on selected applications of graph models in research areas explored by the members of the KRaKE research team.

Keywords: graph, graph models, multigraphs, hypergraphs, causal graphs, knowledge graphs, BPMN, DMN, constraint representation graphs

RZECZ O MODELACH GRAFOWYCH W INŻYNIERII WIEDZY

Streszczenie: W pracy przedstawiono wybrane zastosowania grafowych modeli w inżynierii wiedzy. Wychodząc od bazowej definicji grafu jako zbioru węzłów połączonych krawędziami, zaprezentowano możliwe rozszerzenia tego pojęcia ukierunkowane na zwiększenie siły ekspresji i możliwości przetwarzania wiedzy. W szczególności praca koncentruje się na wybranych zastosowaniach modeli grafowych w obszarach eksplorowanych przez członków grupy badawczej KRaKE.

Słowa kluczowe: graf, modele grafowe, multigrafy, hipergrafy, grafy przyczynowo-skutkowe, grafy wiedzy, BPMN, DMN, grafy reprezentacji ograniczeń

1. Introduction

Over the past several decades, graph models have significantly gained in importance and popularity in technical and scientific applications. Such models, apart from numerical, algebraic and logical models, have a wide range of information representation possibilities and significantly facilitate both the visual and effective presentation of data, information and knowledge, as well as human-computer communication. Moreover, current computer systems enable effective processing of knowledge represented by graph models.

2. A note on evolution of graph definition

The first definition of ‘graph’ was introduced in 1878 by James Joseph Sylvester. He used the word to describe the similarity of the relation between mathematics and chemical structure. Some definitions of graphs that are currently used in different areas include: a simple graph which is a pair $G = (V, E)$, where V is a set of vertices and E is a set of paired vertices (edges). A directed graph is a type of graph in which the edges are directed, a multigraph extends the above definition by allowing multiple edges to have the same pair of endpoints, while a hypergraph allows the edges to join any number of vertices. A conceptual graph is a relatively new definition introduced in 1976 by John F. Sowa to describe schemas used in databases (Sowa 1976). Later, he described various applications of his definition in different fields of science: cognitive science, computer science, and artificial intelligence. The concept has recently been adapted for the knowledge reasoning and representation domain as knowledge graph. It serves as a way to integrate data using a data model based on graph structure. It is used widely in various search engines or linked open data projects. Causal graph, in turn, represents another subtype of directed graphs that are used in statistics and modeling to represent assumptions about the data-generating process.

3. Graphs in knowledge representation and reasoning

Representing knowledge about the world in terms of objects and relations between them is intuitive for humans, so multiple models, of various levels of formality, have been proposed, including informal “mind maps” and various semantic networks, taxonomies (with a defined semantics for “is-a” relationship), thesauri (with relations such as synonyms, antonyms, etc.), up to the most expressive and universal models that are called ontologies. Ontologies, defined in a selected language (ideally based on frag-

ments of formal logic), can express *any* relationships among single objects (instances, individuals) or classes (categories, concepts). In the context of the so-called Semantic Web, which aims to bring formal knowledge representation and reasoning to the Web, Web Ontology Language (OWL) has been adopted as a *de facto* standard for expressing ontologies. Carefully defined variants of OWL allow users to formally represent and process knowledge using description logics. However, the adoption of formal KRR methods is not trivial in the wider community, so a new wave of graph-oriented databases that are simpler, yet easier to use, has gained popularity in recent years. Among the modern implementations of NoSQL databases, Neo4j is probably the most popular and widely-used database system based on graphs. It features intuitive modeling and visualization interface (see Fig. 1) and provides a wide selection of libraries and extensions for data science and graph operations. No wonder it is currently used in numerous business and academic projects.

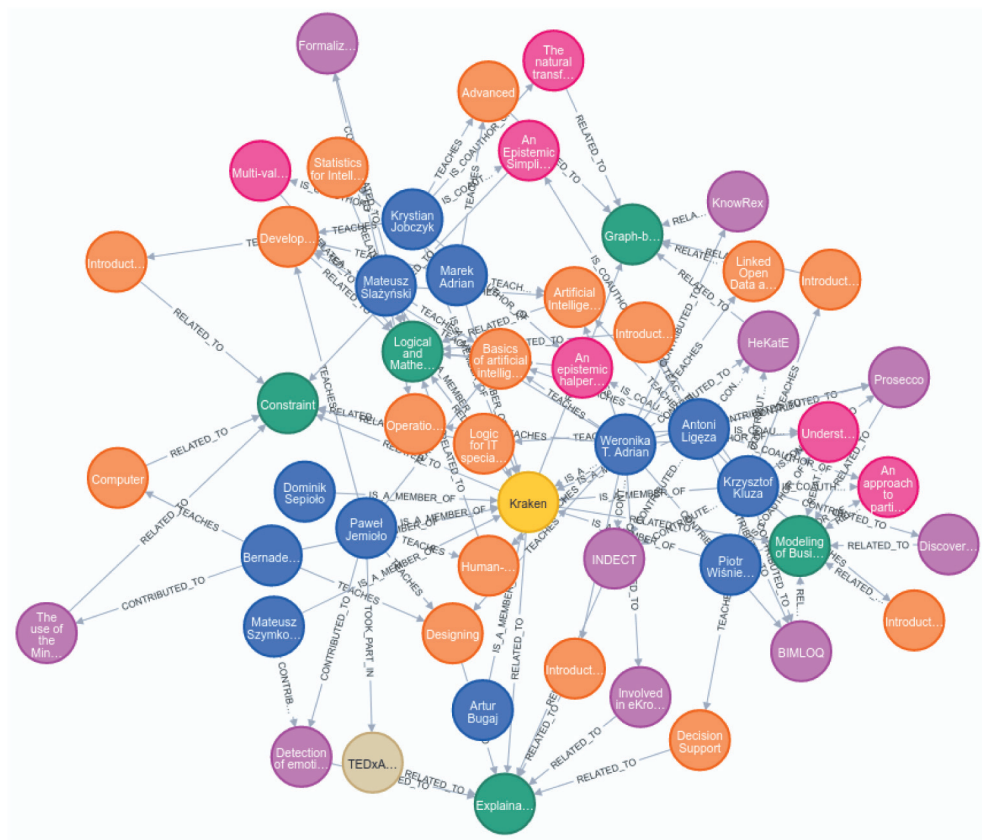


Fig. 1. Neo4j is a modern graph database with multiple extensions for visualization, data science, and knowledge processing. It represents knowledge with labeled property graphs
Source: Pekala (2022)

Graphs in knowledge representation allow for smooth navigation and facilitate knowledge discovery (see Fig. 2). Moreover, a flexible data model allows to pose intuitive queries based on graph patterns and discover paths that represent connections between the entities of the universe of discourse. In order to perform more advanced reasoning tasks, we proposed an extended knowledge graphs model (Adrian et al. 2020), in which the nodes can be of three kinds: agents, objects, and transformations, and within which several “levels” of operations, from simple CRUD tasks up to deductive and inductive reasoning can be performed.

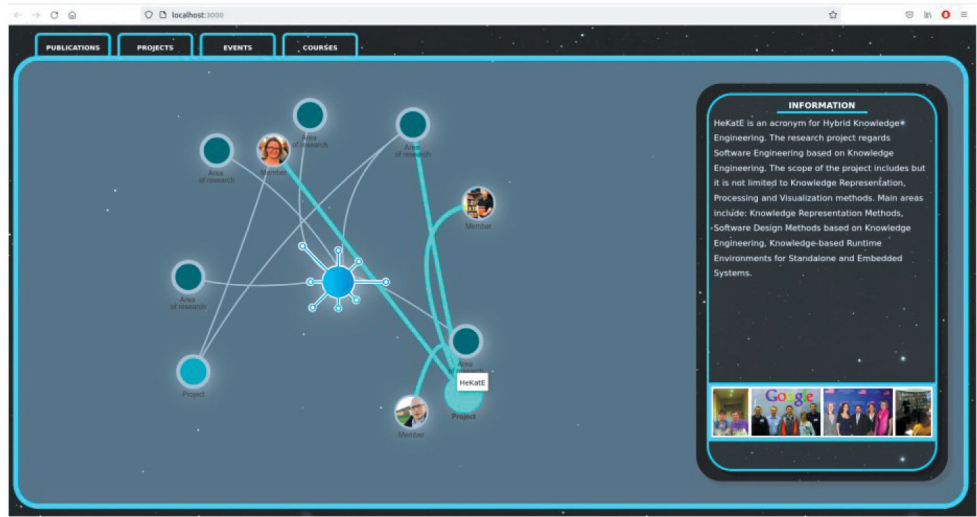


Fig. 2. Intuitive knowledge browsing and discovery of connections between real-world entities.
A case of a research group knowledge representation
Source: Pękala (2022)

4. Causal graphs and Bayes nets

Bayesian networks are a well-established tool for reasoning over uncertain information that can have known and/or latent dependencies between each other. To this end, it uses concepts from two very well-established mathematical domains, namely probability and graph theory. From the first, it takes the notions of independent events and the Bayes theorem. The first concept allows us to directly calculate whether two kinds of observations (for example smoking and developing lung cancer) have any sort of dependency between them. The Bayes theorem, on the other hand, gives us a formula to find how knowledge of one event occurring affects our beliefs about another event, given the inverse dependency between them is known (as an example, we can calculate the probability of developing lung cancer if we are smoking from the information about how many lung cancer patients have been smokers).

Having the necessary tools to calculate possible dependencies between events, we also need to put some structure on them, and for that end, we use graph theory. We construct a graph by assigning a vertex to each event considered in our model. We draw an edge between two vertices if, and only if, we believe to be a direct dependency between those events (as an example, abusing alcohol would be connected by an edge to developing heart disease, but would not be connected to the inevitable heart surgery that comes afterward). This edge would be directed from cause to effect to reflect the casual nature of events in real life. Due to this cause-effect mindset, it is also obvious why no cycles would appear in such representation. Thus, any form of dependencies between events can be represented by a directed acyclic graph.

The Bayesian network model has been studied for a long time by now and many of its strengths, but also many weaknesses have been identified. One of the more significant drawbacks is that, in general, running inference on a Bayesian network is an NP-complete problem, and moreover, any known methods of approximating the results of inference can give results far from useful. Thus, if the model becomes large enough, it is unfeasible to use this particular model of reasoning.

But with some control over the size of the model, we can use Bayesian networks to draw significant conclusions in practical domains, like medicine, as can be seen in the work of Fuster-Parra et al. (2015), where a Bayesian network was implemented to discover relationships among thirteen epidemiological features of the heart age and to analyze their connection to a new concept called cardiovascular lost years. The relationship between smoking, gender, blood pressure and cholesterol to the faster aging of the heart was established (especially for the relation to smoking), while other factors have been found to be independent.

Furthermore, different heuristics can be tried to improve the results given by Bayesian networks. One such attempt was made by Staniec (2020) in his master thesis, where he inquired if clustering the data before running it through a Bayesian network could improve the results of inference. Unfortunately, no improvement was noted, and even degradation in the quality of the inference was observed. This could be the result of the clustered data being too similar, and further inquiries into ways to improve the inference are yet to be tested.

5. Graphs in business processes modeling

Models of business processes constitute graph knowledge representation that help organizations visualize and optimize their processes, what allows for achieving their business goals more effectively. There are various notations for representing these concepts (Kluza et al. 2017, Suchenia et al. 2019). Such models might be modeled manually or generated automatically (Wiśniewski et al. 2019), e.g., from natural

language description (Honkisz et al. 2018) or tasks specification in other form (Wiśniewski et al. 2018).

Since the manual creation of process models usually occurs in a dedicated editor, the use of graph-based models is related to the emerging area of automated modeling. Two main graph representations of a business process are activity graphs and business process graphs.

5.1. Activity graph

An activity graph is a directed graph where each vertex represents a process activity, and each edge corresponds to an admissible precedence relation between two activities (Yan et al. 2019). Activity graphs may also contain nodes that determine the start and end of the workflow. What remains undetermined in such graphs are split and merge flows that represent activities executed in parallel or alternatively. In process models, gateways are used to control such constructs. Although activity graphs provide a general overview of tasks in a process and their sequence, they do not explicitly represent gateways that are an essential part of every process model. As a consequence, additional information about relations between activities is needed to transform an activity graph into a process model.

5.2. Identification of logical gateways in an activity graph

One of the possibilities to determine alternative and parallel gateways in an activity graph is the identification of gateway structures. Having the activity graph as input and knowing which process activities can be executed concurrently, it is possible to define a set of constraints that need to be satisfied by the resulting structure. To obtain a correct graph-based representation of a business process, the formulated constraint satisfaction problem must be resolved for every activity graph node whose degree is greater than one.

5.3. Business process graph

Business process graphs (BP graphs) extend the concept of activity graphs by including nodes that correspond to process flow objects other than activities and start or end events (Heinrich et al. 2020). Assuming that there are no limitations on cyclicity, such graphs may serve as exact copies of process models under the condition that vertices of the graph are annotated with the activity properties, event state descriptions or gateway conditions. Figure 3 shows an example activity graph and the result of its transformation into a BP graph.

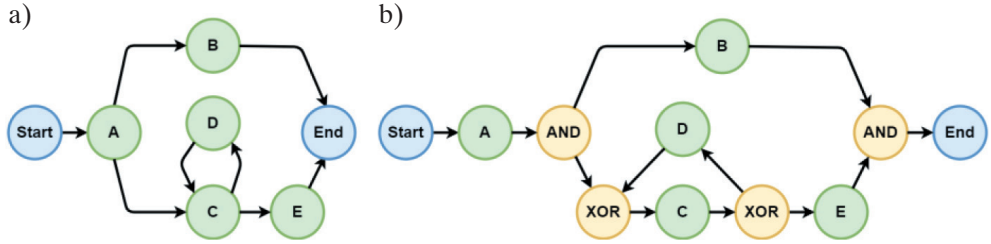


Fig. 3. An activity graph (a) and its corresponding business process graph (b)

Besides the use of process graphs in automated process modeling, such representations are present in process discovery, which consists in extracting the process model from existing event logs. Dymora et al. (2019) used the concept of a BPMN graph as a representation of a business process graph with a graphical distinction of BPMN flow objects.

6. Graphs in logic

The so-called Kripke frame is one of the most fundamental graph structures – exploited in non-classical logic. A single Kripke frame forms a semantic ‘surrogate’ for interpreting the modal-type logic systems, such as temporal, epistemic, deontic, or alethic logic systems.

A single Kripke frame is a tuple $F = \langle U, R \rangle$, where U is a non-empty set of the so-called states, and R is a relation between states. The Kripke frame F is said to be a Kripke model if it is equipped with a standard valuation function, say V , which interprets a given modal logic language as subsets of U .

In order to reconcile a syntax of a given propositional modal language, say L , with its semantic interpretations in terms of Kripke frame-based models, a variety of completeness theorems is formulated.

Example: *Modal logic **K** (a propositional calculus in a modal language with Kripke axiom) is complete with respect to all Kripke models with an arbitrary relation R in them. Modal logic **T** (system **K** with axiom $\phi \rightarrow \Box \phi$) is complete with respect to Kripke models with reflexive relation R in them.*

Putting aside the details of semantic interpretations of modal languages in their Kripke frame-based models, we focus our attention on elucidating Kripke frames (in a more general n -dimension depiction) as a graph-type structure by its similarity to such typical graph structures as trees or irreflexive trees. The construction refers to ideas from (Gabbay et al. 2003).

Having defined a rooted (Kripke) n -frame F with root w_0 , we are in a position to construct another n -frame $U = (W, R_1, \dots, R_n)$ by taking W as the set consisting of $\langle w_0 \rangle$ and all the tuples $\langle w_0, R_1, w_1, \dots, R_k, w_k \rangle$, for $k > 0$ of points of W , and each accessibility relation R_j from the set $\{R_1, \dots, R_n\}$, for $j < k$, and such that:

- $w_j R_j w_{j+1}$, and,
- for any two points $\langle w_1, \dots, w_k \rangle$ and x , the following relation S_j holds:

$$\langle w_0, \dots, w_k \rangle S_j x \Leftrightarrow \exists w \in W (x = \langle w_0, \dots, w_k \rangle R_j w) \quad (1)$$

Informally speaking, S_j creates new states from the previous ones by extending them at the end. The frame U , just introduced, is said to be the *unravelling* of F . The visual effect of this algebraic process is depicted in Figure 4. The left part of it delivers an exemplary Kripke frame, and the right side – its unravelling (Jobczyk 2019).

Let us assume now that R is the smallest reflexive and transitive relation containing the sum relation:

$$R = 1 \leq i \leq n R_i \quad (2)$$

If a rooted n -frame $F = (W, R_1, \dots, R_n)$ of pairwise disjoint relations R_i ($1 \leq j \leq n$) R_i is such that the set $\{y: yRx\}$, for every $x \in W$, is finite and linearly ordered by R restricted to this set, then the frame F is said to be a *tree*. A tree, which forms an *intransitive* n -frame is simply called an *intransitive tree*.

This algebraic reasoning elucidates the graph nature of general Kripke frames and illustrates the method of an intermediate transforming the ordinary (even n -dimensional) Kripke frame to the tree structures (see Fig. 4).

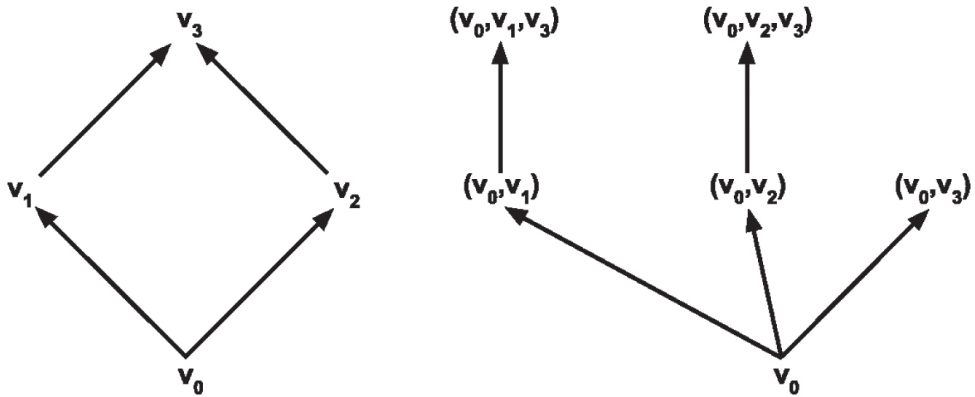


Fig. 4. The graph nature of general Kripke frames

7. Constraint problem solving: hypergraphs for constraint modeling and execution planning

A constraint satisfaction problem (CSP) is one of the classical tasks of artificial intelligence and knowledge engineering. The generic problem statement is relatively simple yet universal; in fact, solving CSP found numerous practical applications. A simple statement of CSP can be formulated as follows: given a *finite set of discrete variables* with *finite domains*, find an *assignment of domain values to all these variables*, so that a *set of predefined constraints is satisfied*. Such problems are solved with advanced constraint programming tools (CP solvers). The core idea of the search is to apply the depth-first search (DFS) with backtracking and extensive constraint propagation. Constraint networks (Lecoutre 2009), hypergraph-based representations are extensively used for modeling and analysis of CSP with the main focus on constraint propagation. However, exploring the problem structure can also be very useful in specific cases. Below, we shall illustrate the use of hypergraphs for *constraint modeling and planning the execution of the search*; the discussion is based on Ligęza (2011, 2012), where further details can be found.

Consider the following, well-known cryptarithmic puzzle:

```

SEND
+ MORE
-----
MONEY

```

The eight variables – S, E, N, D, M, O, R, Y – are to be assigned digits from the domain 0, 1, 2, ..., 9, so that the above constraint is satisfied. Different variables are to be assigned different digits. Leading digits (in our example S and M) must be different from 0.

A naive approach may consist in examining as many as 8^{10} possible assignments; in fact, such problems suffer from a very strong combinatorial explosion. Using contemporary CP methods and tools allows us to easily model the problem and find the only admissible solution in milliseconds. However, exploring the problem structure with a dose of planning leads to the unexpectedly efficient execution of a solution plan.

Consider any of the five columns of the addition – in fact, it is a constraint of the form:

$$X + Y + C_{i-1} = Z + 10 \cdot C_i \quad (3)$$

where X and Y are the respective digits summed up in column i , C_{i-1} (if exists) is the carry digit (0 or 1) from the previous column to the left, and C_i is the carry digit to be passed to the next column.

In the presented example we have four such constraints plus one extra of the form $M = C_4$. The constraints are modeled as a hypergraph (Fig. 5) with the hyperlinks: M, SMO, EON, NRE, and DEY (represented with ovals in the middle line), and vertices: C4, C3, C2, C1, M, S, O, N, E, R, D, Y (represented with small circles).

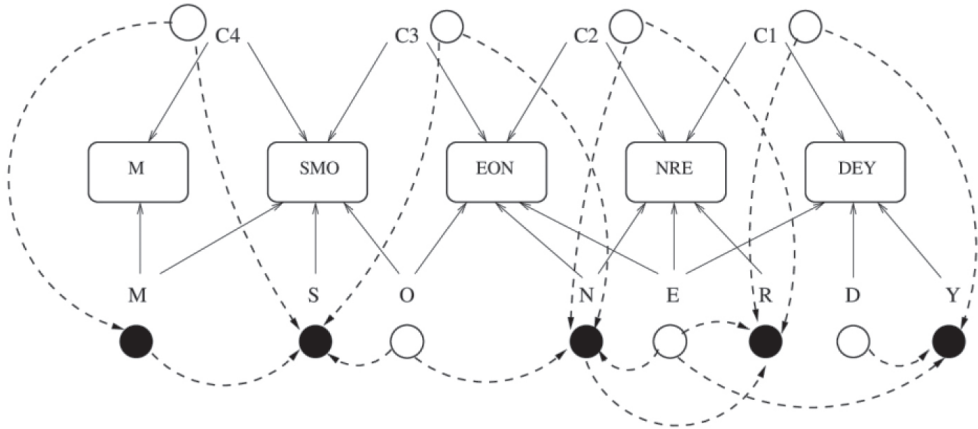


Fig. 5. A hypergraph model of constraints and variables exploration plan

Here, a very efficient plan, exemplified with the dashed line, is as follows: start with C4/2, M, C3/2, O/9, S, C2/2, E/8, N, C1/2, R, D/7, Y. The notation X/N says that we explore variable X with N values in the current domain; X alone means that at the specific point the variable value is calculated in a unique way. Variables, where selection and backtracking are needed, are represented with circles, while variables that can be calculated are represented with black-filled circles.

As for the results, a simple, pure SWI-Prolog program working according to the predefined plan finds the solution in less than 300 steps (logical inferences) and in the time less than 0.000 s CPU on a standard PC/Linux, comparing to around 13,500 steps and 0.005 s CPU when using the clp(fd) library, and around 0.075 s CPU in MiniZinc. Although planning for efficient solving of CSP remains an open issue, this section puts forward some promising ideas of the exploration of hypergraph constraint modeling.

8. Graphs in metaheuristic algorithms

As shown in the previous section, graphs (i.e., constraint hypergraphs) are a very versatile representation language for combinatorial problems. Such formally specified structure may be used to infer an efficient solving strategy tailored to the problem. While the CP solvers, in general, construct the solution by consecutively assigning val-

ues to the variables, many optimization techniques are search based – instead of building a solution, they explore all the possible solutions in an intelligent and often stochastic manner. This family of algorithms is known as metaheuristics and includes such methods as: hill climbing, simulated annealing, and tabu search.

In our work (Ślażyński et al. 2019b), we propose a hyper-heuristic method to impose a structure on the problem search space. Formally, the task involves finding a neighborhood relation: $N : S \rightarrow S$, defined over all the solutions in the search space (S in the formula). The neighborhood relation defines what solutions are considered *accessible* from the given solution. It is used in the search-based optimization to explore the search space, e.g., starting from the solution s_1 algorithm can only move to its neighbors: $s \in N(s_1)$. Given the neighborhood, the search space itself can be viewed as a graph with nodes corresponding to the solutions and edges representing the neighborhood relation itself. The choice of the neighborhood is independent of the problem and may have a crucial impact on the algorithm performance. When the neighborhood is too much connected, the cost of the exploration is prohibitively high. Also, the neighborhood should connect similar solutions, otherwise the search would be effectively random. Finally, the neighborhood may impose intelligent restrictions on the search process by excluding not interesting parts of the search space.

The proposed method exploits a constraint programming model to find an efficient neighborhood for the given combinatorial problem. There are two types of graphs involved in the process. First, there is a typed constraint network – a constraint hypergraph enriched with annotations designed to preserve structural features of a constraint programming model. Secondly, there is the neighborhood itself which (due to the size) is not representable extensionally (i.e., by enumerating all the edges). To bypass this issue, we have defined a formal language called Neighborhood Definition Language (NDL) designed to define the neighborhood intentionally, i.e., it provides a set of neighbors for the given solution. The scope of our work involved the well-specified task: creating an NDL definition of the neighborhood given the TCN representation of the combinatorial problem. Our experiments (Ślażyński et al. 2019a) have shown that such inference is possible, and the neighborhood graph may be found using evolutionary algorithms using the TCN as the source of knowledge and the basis for evaluation of various neighborhood structures.

9. Graphical models in XAI

Graphical models are an efficient form for expressing explanations as they can capture complex dependencies between facts whilst avoiding redundancy (Saha et al. 2021). Graphical models can perform both classification and regression tasks.

Moreover, the structure of tree-based models gives a promise of making a step towards efficient generation of explanations (Sepiło and Ligęza 2022). Several commonly used machine learning models are based on the graph structure. Some of them are compact and interpretable by humans (e.g., decision trees or Bayesian networks). Every prediction generated by those models can be fully explained thanks to their simulatability (the ability of a model to be simulated or thought about strictly by a human (Barredo-Arrieta et al. 2020)), as users can follow and understand the graphical representation of the model’s structure.

However, human capacities limit the complexity of those transparent graphical models, which is usually connected with their weaker predictive power. More complex, opaque models are being implemented. Random forests (RF) are among the most prominent ones. RF is a collection of explainable-by-itself decision (Fig. 6).

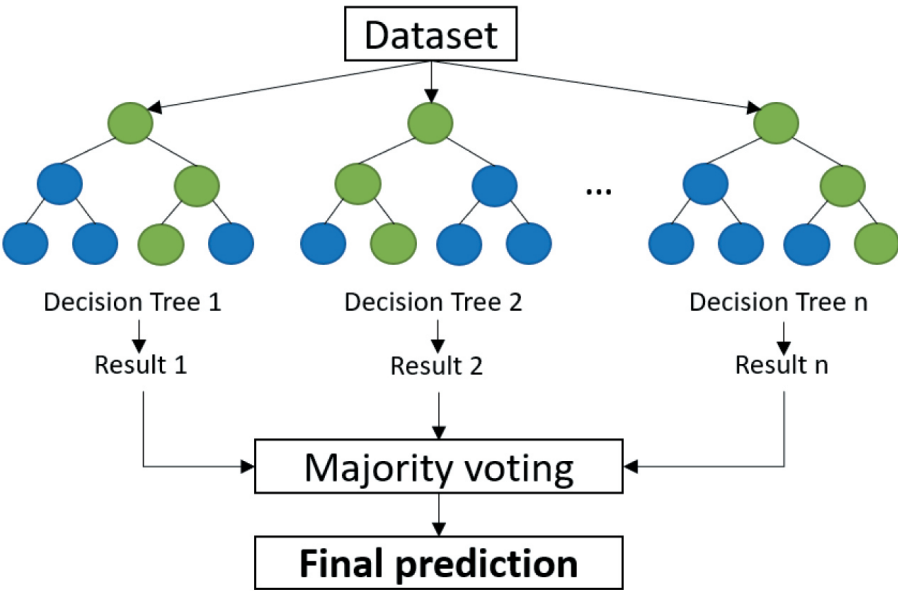


Fig. 6. A random forest operating scheme

The predictions are generated independently by each tree in the collection and then combined into the final decision, usually through majority voting. RF provides better prediction accuracy and is generally more resistant to overfitting than a single decision tree. However, complex ensemble models cannot be easily interpreted. This is the reason why various model-specific (methods that utilize assumptions regarding model structure) and model-agnostic (methods that can be applied to any class of machine learning models) explainability techniques are being developed as part of explainable artificial intelligence (XAI) toolkit.

Nowadays, more and more systems incorporate deep learning techniques. Taking into account their performance, it is not surprising. However, the accuracy is often connected with high complexity. Following this, trust in systems involving deep learning remains low. Therefore, researchers came up with explainable artificial intelligence, which goal is to provide insight into how deep learning models operate (Gunning et al. 2019). Such insight should be comprehensible for users, who often are not experts in a specific field.

In our work (Mróz et al. 2020), we focused on incorporating graphs into the process of explanations generation. More specifically, we used knowledge graphs with the significance of two features presented as the level of thickness of edges and the features encoded as nodes. We believe that presenting the results in the form of graphs may increase the readability and comprehensibility of the model structure (Adrian et al. 2022).

10. Summary

In this overview paper, we provided an outline of basic definitions of various types of graphs and summarized possible extensions of graph concepts and applications of graph-based technologies. Specifically, we presented various applications of graph models in the knowledge engineering area explored by the members of the KRaKE research group, such as knowledge representation, knowledge discovery, business process modeling, logic reasoning, constraint modeling and execution planning, metaheuristic algorithms, as well as explainable artificial intelligence methods.

References

- Adrian W.T., Adrian M., Kluza K., Stachura-Terlecka B., Ligęza A., 2020, *Extended knowledge graphs: A conceptual study*, [in:] Aveiro D., Dietz J., Filipe J. (eds.), *IC3K 2020: Proceedings of the 12th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management: November 2–4, 2020. Vol. 2, KEOD: 12th International Conference on Knowledge Engineering and Ontology Development*, SciTePress – Science and Technology Publications, pp. 173–180. <https://doi.org/10.5220/0010111601730180>.
- Adrian W.T., Kluza K., Zaremba M., Jemioło P., Adrian M., Pękala A., Florek K., Ligęza A., 2022, *Selected applications of graph-based knowledge representation*, [in:] *PP-RAI'2022: Proceedings of the 3rd Polish Conference on Artificial Intelligence: April 25–27, 2022, Gdynia, Poland*, Gdynia Maritime University, Gdynia, pp. 172–175.

- Barredo Arrieta A., Díaz-Rodríguez N., Del Ser J., Bennetot A., Tabik S., Barbado A., Garcia S., Gil-Lopez S., Molina D., Benjamins R., Chatila R., Herrera F., 2020, *Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI*, Information Fusion, vol. 58, pp. 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>.
- Dymora P., Koryl M., Mazurek M., 2019, *Process discovery in business process management optimization*, Information, vol. 10(9), 270. <https://doi.org/10.3390/info-10090270>.
- Fuster-Parra P., Tauler P., Bennasar-Veny M., Ligęza A., Lopez-Gonzalez A.A., Aguiló A., 2016, *Bayesian network modeling: A case study of an epidemiologic system analysis of cardiovascular risk*, Computer Methods and Programs in Biomedicine, vol. 126, pp. 128–142. <https://doi.org/10.1016/j.cmpb.2015.12.010>.
- Gabbay D., Kurucz A., Wolter F., Zakharyashev M. (eds.), 2003, *Many-Dimensional Modal Logic: Theory and Applications*, Studies in Logic and the Foundations of Mathematics, 148, Elsevier, Amsterdam.
- Gunning D., Stefik M., Choi J., Miller T., Stumpf S., Yang G.Z., 2019, *XAI – Explainable artificial intelligence*, Science Robotics, vol. 4(37), eaay7120. <https://doi.org/10.1126/scirobotics.aay7120>.
- Heinrich B., Schiller A., Schön D., Szubartowicz M., 2020, *Adapting process models via an automated planning approach*, Journal of Decision Systems, vol. 29(4), pp. 223–259. <https://doi.org/10.1080/12460125.2020.1800976>.
- Honkisz K., Kluza K., Wiśniewski P., 2018, *A concept for generating business process models from natural language description*, [in:] Liu W., Giunchiglia F., Yang B. (eds.), *Knowledge Science, Engineering and Management: 11th International Conference, KSEM 2018: Changchun, China, August 17–19, 2018: Proceedings, Pt. 1*, Lecture Notes in Computer Science, vol. 11061, Springer, Cham, pp. 91–103. https://doi.org/10.1007/978-3-319-99365-2_8.
- Jobczyk K., 2019, *Multi-valued deontic Halpern–Shoham logic for fuzzy deontic-temporal expressions*, Journal of Intelligent and Fuzzy Systems, vol. 36(5), pp. 5091–5103. <https://doi.org/10.3233/JIFS-179054>.
- Kluza K., Wiśniewski P., Jobczyk K., Ligęza A., Suchenia A., 2017, *Comparison of selected modeling notations for process, decision and system modeling*, [in:] Ganzha M., Maciaszek L., Paprzycki M. (eds.), *FedCSIS: Proceedings of the 2017 Federated Conference on Computer Science and Information Systems: September, 3–6, 2017, Prague, Czech Republic*, IEEE, Piscataway, pp. 1111–1114. <https://doi.org/10.15439/2017F454>.
- Lecoutre Ch., 2009, *Constraint Networks: Targeting Simplicity for Techniques and Algorithms*, ISTE, Wiley, London, UK, Hoboken, USA.

- Ligęza A., 2011, *Models and tools for improving efficiency in constraint logic programming*, Decision Making in Manufacturing and Services, vol. 5(1–2), pp. 68–78.
- Ligęza A., 2012, *Improving Efficiency in Constraint Logic Programming Through Constraint Modeling with Rules and Hypergraphs*, [in:] Ganzha M., Maciaszek L., Paprzycki M. (eds.), *FedCSIS: Proceedings of the Federated Conference on Computer Science and Information Systems 2012: September 9–12, 2012 Wrocław, Poland*, Polskie Towarzystwo Informatyczne, Warsaw; IEEE Computer Society Press, Los Alamitos, pp.101–107.
- Mróz P., Quemy A., Ślażyński M., Kluza K., Jemioło P., 2020, *GBEx – towards graph-based explanations*. [in:] Alamaniotis M., Pan S. (eds.), *ICTAI 2020: IEEE 32nd International Conference on Tools with Artificial Intelligence: 9–11 November 2020: Proceedings*, IEEE, Piscataway, pp. 112–117. <https://doi.org/10.1109/ICTAI50040.2020.00028>.
- Pękała A., 2022, *Strona internetowa zespołu badawczego zaprojektowana w oparciu o graf wiedzy*, AGH University of Science and Technology [engineering thesis, unpublished].
- Saha S., Yadav P., Bauer L., Bansal M., 2021, *ExplaGraphs: An Explanation Graph Generation Task for Structured Commonsense Reasoning*, [in:] Moens M.-F., Huang X., Specia L., Wen-tau Yih S. (eds.), *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing: EMNLP 2021 Virtual Event: Punta Cana, Dominican Republic, 7–11 November, 2021*, Association for Computational Linguistics, pp. 7716–7740. <https://doi.org/10.18653/v1/2021.emnlp-main.609>.
- Sepioło D., Ligęza A., 2022, *Towards explainability of tree-based ensemble models: A critical overview*, [in:] Wojciech Zamojski et al. (eds.), *New Advances in Dependability of Networks and Systems: Proceedings of the Seventeenth International Conference on Dependability of Computer Systems DepCoS-RELCOMEX: June 27–July 1, 2022, Wrocław, Poland*, Lecture Notes in Networks and Systems, vol. 484, Springer, Cham, pp. 287–296. https://doi.org/10.1007/978-3-031-06746-4_28.
- Sowa J.F, 1976, *Conceptual graphs for a data base interface*, IBM Journal of Research and Development, vol. 20(4), pp. 336–357. <https://doi.org/10.1147/rd.204.0336>.
- Staniec A., 2020, *Wybrane problemy odkrywania modeli przyczynowo-skutkowych z użyciem sieci bayesowskich*, Akademia Górniczo-Hutnicza, Kraków [M.Sc. thesis].
- Suchenia A., Łopata P., Wiśniewski P., Stachura-Terlecka B., 2019, *Towards UML representation for BPMN and DMN models*, MATEC Web of Conferences, vol. 252, 02007. <https://doi.org/10.1051/matecconf/201925202007>.
- Ślażyński M., Abreau S., Nalepa G., 2019a, *Generating local search neighborhood with synthesized logic programs*, Electronic Proceedings in Theoretical Computer Science, vol. 306, pp. 168–181. <https://doi.org/10.4204/EPTCS.306.22>.

- Ślażyński M., Abreau S., Nalepa G., 2019b, *Towards a formal specification of local search neighborhoods from a constraint satisfaction problem structure*, [in:] *GECCO 2019: The Genetic and Evolutionary Computation Conference: A Recombination of the 28th International Conference on Genetic Algorithms (ICGA) and the 24rd Annual Genetic Programming Conference (GP): July 13th–17th 2019, Prague, Czech Republic*, ACM, USA, pp. 137–138. <https://doi.org/10.1145/3319619.3321968>.
- Ślażyński M., Abreau S., Nalepa G., 2019c, *Generating local search neighborhood with synthesized logic programs*, *Electronic Proceedings in Theoretical Computer Science* [35th International Conference on Logic Programming (ICLP'19): Las Cruces, USA, September 20–25, 2019], vol. 306, pp. 168–181. <http://doi.org/10.4204/EPTCS.306.22>.
- Wiśniewski P., Kluza K., Ligęza A., 2018, *An approach to participatory business process modeling: BPMN model generation using constraint programming and graph composition*, *Applied Sciences*, vol. 8(9), 1428. <https://doi.org/10.3390/app8091428>.
- Wiśniewski P., Kluza K., Jobczyk K., Stachura-Terlecka B., Ligęza A., 2019, *Overview of generation methods for business process models*, [in:] Douligieris Ch., Karagiannis D., Apostolou D. (eds.), *Knowledge Science, Engineering and Management: 12th International Conference, KSEM 2019, Athens, Greece, August 28–30, 2019: Proceedings, Pt. 2*, *Lecture Notes in Computer Science*, vol. 11776, Springer, Cham, pp. 55–60. https://doi.org/10.1007/978-3-030-29563-9_6.
- Yan Z., Sun B., Chen Y., Wen L., Hu L., Wang J., Yang M., Wang L., 2019, *Decomposed and parallel process discovery: A framework and application*, *Future Generation Computer Systems*, vol. 98, pp. 392–405. <https://doi.org/10.1016/j.future.2019.03.048>.

Symulacja numeryczna obróbki cieplnej odlewu ze staliwa i jej weryfikacja doświadczalna

Szymon Żołyńia¹, Sebastian Sobuła² 

¹ AGH Akademia Górniczo-Hutnicza, Koło Naukowe ZGAREK, Kraków

² AGH Akademia Górniczo-Hutnicza, Wydział Odlewnictwa, Kraków

Streszczenie: Zagadnienie technologii obróbki cieplnej staliwa z wykorzystaniem nowoczesnych narzędzi symulacyjnych ma istotne znaczenie dla współczesnego inżyniera. Zasadniczym celem pracy była weryfikacja doświadczalna wyników symulacji obróbki cieplnej staliwa węglowego L500-II. Zgodnie z założeniem przeprowadzono symulację komputerową procesu hartowania z wykorzystaniem oprogramowania Visual-Weld (Sysweld). Równolegle wykonano badania doświadczalne, w czasie których zarejestrowano krzywe chłodzenia. Po hartowaniu przeprowadzonym w wodzie badano rozkład twardości na przekroju próbki, a także mikrostrukturę w wybranych obszarach z wykorzystaniem mikroskopu świetlnego oraz oprogramowania ImageJ. Weryfikację wyników symulacji komputerowej wykonano, zestawiając ze sobą dane, takie jak rozkład temperatury w próbce podczas chłodzenia, rozkład twardości oraz udział składników fazowych. Wykazano, że wyniki z przeprowadzonej symulacji numerycznej odbiegają od uzyskanych w doświadczeniu, szczególnie w obszarze blisko powierzchni stygnięcia.

Słowa kluczowe: obróbka cieplna, symulacja numeryczna, staliwo, Visual-Weld, Sysweld

NUMERICAL SIMULATION OF HEAT TREATMENT OF STEEL CASTING AND ITS EXPERIMENTAL VALIDATION

Abstract: The problem of the heat treatment technology of cast steel with use of modern simulation tools is important to the today's engineer. The main purpose of thesis was experimental verification of the numerical simulation results of L500-II carbon cast steel heat treatment. In accordance to the assumption, a computer simulation of the heat treatment process was carried out using the Visual-Weld (Sysweld) software. Experimental tests, during which the cooling curves were recorded, were carried out simultaneously. After conducted quenching in water, the hardness distribution on the cross-section of the sample was examined, as well as the microstructure in selected areas, using a light microscope and ImageJ software. It was shown, that the results of the numerical simulation are different from those obtained in the experiment.

Keywords: heat treatment, numerical simulation, cast steel, Visual-Weld, Sysweld

https://doi.org/10.7494/978-83-67427-74-6_9

1. Wprowadzenie

W celu zapewnienia wyrobom stalowym wytrzymałości oraz odporności na czynniki środowiskowe niezbędne jest poddanie ich obróbce cieplnej. Pozwala ona zmaksymalizować korzystne właściwości materiału, co umożliwia m.in. redukcję gabarytów odlewu przy zachowaniu dobrych parametrów technicznych, ponieważ odlewy ze staliwa charakteryzują się bardzo niekorzystną, gruboziarnistą strukturą oraz naprężeniami odlewniczymi w stanie lanym (Głownia 2002). Obróbka cieplna jest definiowana jako odpowiednio dobrane zabiegi cieplne prowadzące do zmiany własności stopu przez zmiany struktury wywołane przemianami fazowymi zachodzącymi w stanie stałym (Adrian 2011). Rosnące wymagania odbiorców wyrobów przy jednoczesnym dążeniu do redukcji kosztów skłaniają technologów do poszukiwania metod optymalizacji procesu produkcyjnego. Nieocenioną pomocą w tym zakresie są programy inżynierskie pozwalające wykonać symulację, która umożliwia ocenę oraz doskonalenie wdrażanych rozwiązań z pominięciem czasochłonnych i kosztowanych badań empirycznych. Zagadnienie symulacji numerycznej procesów przemysłowych jest poruszane w wielu publikacjach naukowych. W branży przetwórstwa metali popularne są programy symulujące procesy odlewania, krystalizacji i stygnięcia metalu w formach (Magma, Procast, Flow-3D, NovaFlow&Solid, QuickCAST, AnyCasting), programy symulujące procesy spajania i obróbki cieplnej (Sysweld, Simufact Welding) i inne, np. Ansys, Abaqus, które również mogą symulować procesy łączenia, ale na podstawie modułu cieplno-mechanicznego, nieuwzględniającego naprężenia powstającego podczas przemian fazowych w stanie stałym. W niniejszym opracowaniu do przeprowadzenia obliczeń numerycznych procesu obróbki cieplnej wykorzystano kompleksowe środowisko symulacyjne ESI Visual-Weld (dawniej Sysweld).

2. Charakterystyka środowiska Visual-Weld

Dziejąca się na naszych oczach czwarta rewolucja przemysłowa jeszcze mocniej integruje procesy produkcyjne z oprogramowaniem inżynierskim, które już teraz odgrywa kluczową rolę we wszystkich obszarach działalności produkcyjnej, konstrukcyjnej i technologicznej. Algorytmy komputerowe w znaczący sposób ułatwiają wszelkie procesy, pozwalają zaoszczędzić czas i zasoby. W oprogramowaniu inżynierskim, w zależności od przeznaczenia, można wyróżnić trzy główne grupy (Stanisławski 2008); są to programy do komputerowego:

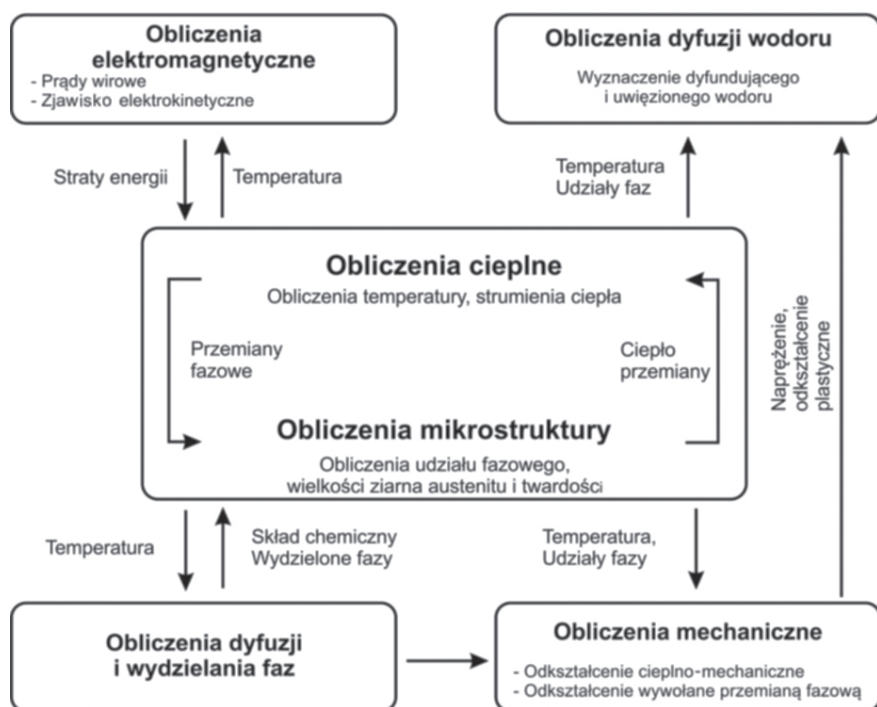
- projektowania (*computer-aided design*, CAD),
- wytwarzania (*computer-aided manufacturing*, CAM),
- analizowania i obliczeń inżynierskich (*computer-aided engineering*, CAE).

Jednym z takich programów jest Sysweld będący częścią większego pakietu narzędzi inżynierskich Visual-Weld grupy Engineering Systems International (ESI). Oprogramo-

wanie zostało stworzone jako specyficzne, zaawansowane narzędzie służące do analizy numerycznej procesów spawania oraz obróbki cieplnej (Kik i in. 2020).

Program ma budowę wielopoziomową, tzn. wyniki obliczeń przekazywane są do kolejnych modułów i służą jako dane wejściowe do dalszych obliczeń. Na rysunku 1 przedstawiono generalną architekturę i schemat działania programu Sysweld (Pont i Guichard 1995, ESI Group 2017). Symulacja jest prowadzona w następujących krokach:

1. obliczenia cieplne i przemian w stanie stałym (mikrostruktury), w których identyfikuje się występujące przemiany na podstawie zmiany temperatury i wyznacza udział fazowy w odniesieniu do czasu i przestrzeni, a jeśli występuje proces nagrzewania, np. indukcyjnego, dane zmian temperatury przekazywane są do modułu obliczeń elektromagnetycznych i po przetworzeniu odbierane z powrotem;
2. obliczenia mechaniczne, w których wyznaczane jest naprężenie i odkształcenie, powstałe w wyniku różnicy temperatury, a także przemian fazowych;
3. obliczenia dyfuzji pierwiastków i wydzielania faz, w których wyznaczane są składniki mikrostruktury powstające w założonym przedziale temperatury i czasu;
4. obliczenia dyfuzji wodoru (spawanie), które łączą wpływ temperatury, naprężenia, a także defektów i przewidują miejsca, w których gromadzi się wodor.



Rys. 1. Struktura i schemat procedury obliczeń w programie Sysweld

Źródło: opracowanie własne na podstawie Pont i Guichard (1995), ESI Group (2017)

Wyznaczenie przemian fazowych w warunkach nieustalonego przepływu ciepła, który występuje w procesach spawania i obróbki cieplnej, jest zagadnieniem złożonym i niezbędnym do prawidłowej oceny właściwości obrabianych przedmiotów. Właściwości cieplne i fizyczne poszczególnych faz w stopach oraz ich wzajemne przemiany istotnie wpływają na stan naprężenia, dlatego główny moduł programu, obliczeń cieplnych i mikrostrukturalnych, realizuje trzy typy zadań:

- 1) wyznacza przemiany zgodnie z kierunkiem zmian temperatury w czasie,
- 2) w zależności od występujących przemian modyfikuje pole temperatury, uwzględniając wpływ ciepła przemiany,
- 3) wyznacza właściwości cieplne i fizyczne (np. współczynnik przewodzenia ciepła) na podstawie aktualnych udziałów objętościowych faz.

Obliczenia cieplne bazują na rozwiązaniu równania przewodzenia ciepła Fouriera (1) dla nieustalonego przepływu ciepła (Lee 2018), za pomocą metody elementów skończonych (MES):

$$\rho \cdot c \cdot \frac{\partial(T)}{\partial t} = \frac{\partial}{\partial x} \left(\lambda_x \frac{\partial T}{\partial x} \right) + \frac{\partial}{\partial y} \left(\lambda_y \frac{\partial T}{\partial y} \right) + \frac{\partial}{\partial z} \left(\lambda_z \frac{\partial T}{\partial z} \right) + Q_V \quad (1)$$

gdzie:

- $\rho \cdot c$ – objętościowe ciepło właściwe [$\text{J}/(\text{m}^3 \cdot \text{K})$],
- T – temperatura [K],
- t – czas [s],
- λ – współczynnik przewodzenia ciepła [$\text{W}/(\text{m} \cdot \text{K})$],
- Q_V – objętościowa szybkość wydzielania ciepła [W/m^3].

Należy pamiętać, że w równaniu (1) objętościowe ciepło właściwe zmienia się z temperaturą, a Q_V ujemne ciepło wydzielające się podczas przemian fazowych w czasie chłodzenia.

Do obliczenia właściwości termofizycznych w zakresie wielofazowym autorzy programu zastosowali regułę mieszanin. Oznacza to, że dana właściwość mieszaniny faz jest wyznaczana jako suma tej właściwości w poszczególnych fazach. W ten sposób są wyznaczone gęstość, ciepło właściwe i inne niezbędne parametry termofizyczne. Udziały (proporcje) faz powstających podczas przemian w stanie stałym wyznaczone są metodą zaproponowaną przez Leblonda i Devaux (1984).

$$\frac{\partial X}{\partial t} = c_L (X_{\text{eq}}(T) - X) \quad (2)$$

gdzie:

- X_{eq} – równowagowy udział fazy w funkcji temperatury,
- c_L – stała empiryczna przyjmująca wartości dodatnie, wyznaczona na podstawie czasu trwałości fazy w danej temperaturze,
- X – udział fazy.

Kinetyka wydzielania faz w procesach dyfuzyjnych wyznaczana jest na podstawie układów czas – temperatura – przemiana (CTT) z uwzględnieniem modelu JMAK (Johson–Mehl–Avrami–Kołmogorow).

$$X_F = 1 - e^{-(K \cdot t)^n} \quad (3)$$

gdzie:

X_F – udział fazy,

K – stała zależna od temperatury,

n – stała materiałowa (zwykle bliska jedności).

Ponadto program wyznacza udziały martenzytu na podstawie równania podanego przez Koistinena i Marburgera (1959).

$$X_M(T) = 1 - e^{-b(M_s - T)} \quad (4)$$

gdzie:

b – parametr materiałowy (bliski 0,011) (Babu 2014, Bhadeshia 2014),

X_M – udział martenzytu,

M_s – temperatura martenzyt start [K].

Obliczenia mechaniczne również są realizowane za pomocą MES, a dane wejściowe są pobierane z modułu obliczeń cieplnych i mikrostrukturalnych. Aby uprościć, w module tym przyjęto, że naprężenia nie wpływają na temperaturę i udziały składników fazowych. Przemiany w stanie stałym mają znaczący wpływ na generowane naprężenia ze względu na istotne zmiany w objętości poszczególnych faz, a także na fakt, że w jednostce czasu może istnieć wiele faz (Pont 1995).

Model dyfuzji bazuje na drugim prawie Ficka. Ponieważ jest wykorzystywany głównie w procesach obróbki cieplno-chemicznej, jego działanie nie zostanie szerzej opisane.

Program Visual-Weld, jako nowoczesne narzędzie służące do analizy numerycznej procesów spawania i obróbki cieplnej, pomaga rozwiązywać problemy dotyczące analizy nieliniowej przewodzenia ciepła w każdej przestrzeni, geometrii nieliniowej odkształceń, izotropowego i kinematycznego umocnienia materiału oraz przemian fazowych itp. (Kik i in. 2020). Wielopoziomowa struktura obejmuje następujące podprogramy (Swacha i in. 2021):

- Visual-Mesh – preprocesor pozwalający na tworzenie elementów (1D, 2D, 3D) oraz generowanie siatki,
- Visual-Weld – program wykorzystywany do modelowania procesów spawania,
- Visual-Heat Treatment – program wykorzystywany do modelowania procesów obróbki cieplnej,
- Sysweld Solvers – program zawierający zespół funkcji numerycznych realizujących właściwe obliczenia,
- Visual-Viewer – postprocesor pozwalający na analizę wyników.

Dla uzyskania wyników jak najbardziej odzwierciedlających rzeczywiste procesy kluczowe jest odpowiednie przygotowanie danych wejściowych, w tym parametrów materiałowych, którymi dysponuje program. Najczęściej właściwości mechaniczne są obliczane z wykorzystaniem algorytmów uwzględniających udziały poszczególnych faz oraz temperatury. Oprócz danych cieplno-mechanicznych, wśród których należałoby wyróżnić m.in. współczynnik przewodzenia ciepła, gęstość, moduł Younga, równie istotne są dane metalurgiczne. Uwzględnienie przemian fazowych oraz ich kinetyki ma znaczący wpływ na informacje zwrotne otrzymywane na etapie analizy cieplnej i strukturalnej oraz wynikającej z niej analizy mechanicznej (Kik i in. 2020).

Sysweld umożliwia symulację różnych zabiegów obróbki cieplnej, m.in. hartowania, odpuszczania, wyżarzania, nawęglania i azotowania (Swacha i in. 2021). W myśl zasady, że jakość wyników może być co najwyżej tak dobra jak jakość zadanych parametrów wejściowych, można stwierdzić, że decydujący wpływ ma poprawne określenie warunków początkowych, w tym utworzenie odpowiednio dopasowanej siatki powierzchniowej 2D oraz objętościowej 3D. Visual-Weld pozwala generować siatkę w układzie *quad* (sześciennym), *tria* (tetragonalnym) oraz *quad-tria* (mieszanym), co w połączeniu z funkcją lokalnego zagęszczania znacząco usprawnia obliczenia, a także korzystnie wpływa na dokładność wyników. Niewątpliwym atutem programu firmy ESI są rozbudowane bazy materiałowe, które są ciągle wzbogacane o kolejne pozycje, ale też mają możliwość samodzielnego aktualizowania. Wymienione atuty aplikacji Sysweld czynią z niej czołowe rozwiązanie w zakresie symulacji obróbki cieplnej oraz procesów spawania (Swacha i in. 2021).

3. Modelowanie procesów spawania i obróbki cieplnej z wykorzystaniem środowiska Visual-Weld

Analiza źródeł literaturowych wskazuje na popularność narzędzi do symulacji procesów spawania i obróbki cieplnej. Wielu autorów wykorzystuje własne programy, ale często bazują one na komercyjnych aplikacjach (środowiskach) wymienionych we wprowadzeniu. Środowisko Visual-Weld znajduje zastosowanie do symulacji pola temperatury, stanu naprężenia, a także obróbki cieplnej – przewidywania udziału faz i twardości obrabianych przedmiotów.

Tak na przykład oprogramowanie inżynierskie wykorzystane zostało do weryfikacji wyników symulacji spoiny wykonanej metodą GTA (spawania łukowego elektrodą wolframową) na stali P91 przez Zubairuddina i in. (2023). W pracy porównano ze sobą wyniki symulacji z programów Sysweld oraz FlexPDE, przy szybkości spawania wynoszącej 1,6 mm/s, sprawności 75% oraz całkowitym dopływie ciepła 424 J/s. Autorzy wykazali, że analiza spoiny przy użyciu modelu podwójnej elipsoidy wykazuje dobrą zgodność

z profilem eksperymentalnym, ponieważ błąd wyników symulacji wyniósł kolejno 7% dla Sysweld oraz 14% dla FlexPDE.

Oprogramowanie może być wykorzystywane do projektowania nowych technik wytwarzania. Tak na przykład Lu i in. (2014) zaproponowali wykorzystanie programu Sysweld do symulacji pola naprężeń powstających podczas napawania powierzchni wielkogabarytowej matrycy do kucia, wykonanej ze staliwa niskostopowego.

Zagadnienie modelowania obróbki cieplnej ma istotne znaczenie w pracy współczesnego inżyniera. Ze względu na moc obliczeniową obecnych komputerów proces ten jest relatywnie szybki i może być stosowany już na etapie planowania produkcji.

W pracach Ryzhkova i in. (2015, 2018) przedstawiono wyniki symulacji hartowania na powietrzu i w oleju stali Mn-Si-Ni-Mo wykonanej z zastosowaniem programu Sysweld, a także wyniki uzyskane na komercyjnej linii produkcyjnej. Autorzy wykazali dobrą zgodność wyników doświadczalnych z symulacją dzięki aktualizacji danych termofizycznych, które pozyskali, przeprowadzając dodatkowe badania.

W pracy Kika i in. (2015) przedstawiono symulację procesu obróbki cieplnej stali 10GN2MFA z wykorzystaniem pakietu Visual-Weld. Badania hartowania i normalizowania prowadzono na prostopadłościanach o wymiarach 50 mm × 40 mm × 20 mm. Wykazano dobrą zbieżność wyników symulacji twardości na przekroju z wynikami hartowania w oleju i normalizowania. Wyniki doświadczalnego badania twardości próbek hartowanych w wodzie charakteryzowały się dużą rozbieżnością w porównaniu z wynikami symulacji.

Oprócz klasycznej obróbki cieplnej rozwijane są także niekonwencjonalne sposoby prowadzenia tego procesu. Nieszablonowe podejście do obróbki cieplnej stali węglowej, polegające na wykorzystaniu wiązki lasera, zaprezentował Evdokimov i in. (2022). W swej pracy badacze przeprowadzili symulację procesu spawania stali C45 z zastosowaniem autorskiego algorytmu numerycznego źródła ciepła, opartego na metodzie elementów skończonych. Zaproponowany algorytm pozwolił wyznaczyć natężenie promieniowania świetlnego na powierzchni przedmiotu obrabianego w funkcji parametrów wiązki lasera. Wykazano, że oprogramowanie *open source* FEniCSx może być używane do laserowej obróbki cieplnej, jeśli w obliczeniach wykorzysta się odpowiednie równanie nieustalonego przepływu ciepła w funkcji czasu.

Symulacja procesów obróbki cieplnej pozwala przewidzieć mikrostrukturę i właściwości obrabianych przedmiotów stalowych. W przypadku odlewów należy pamiętać, że w zależności od grubości ścianki mają inną mikrostrukturę (wielkość ziarna, udział składników fazowych), dlatego symulacja może nie odzwierciedlać ich rzeczywistych właściwości. Zgodnie z aktualnym stanem wiedzy skuteczność omawianego oprogramowania została wykazana głównie w kontekście spawania, a także w znacznie bardziej ograniczonym zakresie w odniesieniu do modułu obróbki cieplnej stali. Badania symulacji obróbki cieplnej odlewanych stopów żelaza (staliwa) nie są publikowane. Z tego powodu oraz ze względu na fakt, że symulacja obróbki cieplnej może mieć realny wpływ na ekonomikę

w przemyśle odlewniczym, postanowiono rozwinąć obszar wiedzy związany z obróbką cieplną staliwa. Przeprowadzono doświadczalną weryfikację wyników symulacji procesu na próbce wyciętej z odlewu staliwnego (gatunek L500II / 270-480), która została poddana hartowaniu zwykłemu w wodzie. Porównano ze sobą dane uzyskane doświadczalnie z otrzymanymi w symulacji.

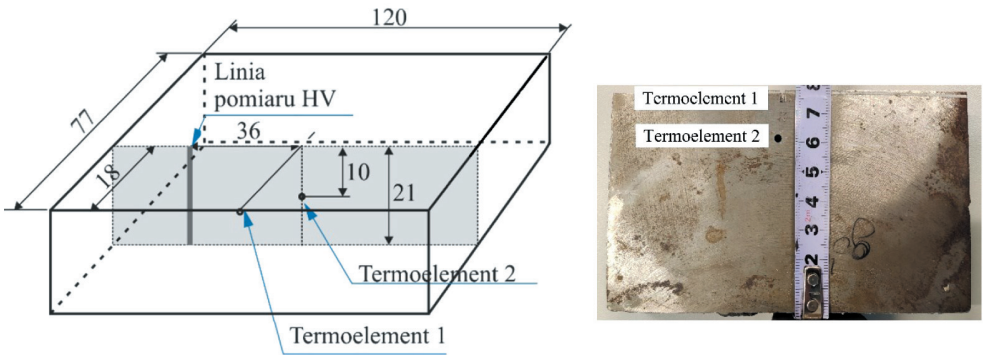
4. Metodyka

Do badań własnych wykorzystano próbkę o kształcie płyty wyciętej z wlewka ze staliwa L500-II. Symulację numeryczną przeprowadzono, przyjmując geometrię identyczną jak geometria próbki poddanej badaniom doświadczalnym. Ze względu na brak w bazie danych programu staliwa przewidzianego do badań do symulacji wybrano stal CF35 o zbliżonym składzie chemicznym (tab. 1).

Tabela 1
Skład chemiczny staliwa L500-II i stali CF35 [% mas.]

Stop	C	Si	Mn	P	S	Ti	Fe
L500-II	0,32	0,48	0,85	0,015	0,011	0,065	reszta
CF35	0,33–0,39	0,15–0,35	0,50–0,80	<0,025	<0,035	–	reszta

Za materiał do badań doświadczalnych posłużyła próbka o wymiarach 120 mm × 77 mm × 21 mm (rys. 2), wycięta z prostokątnego wlewka przylanego o wymiarach 300 mm × 300 mm × 120 mm. W próbce zamontowano dwa termoelementy typu K, tj. NiCr-NiAl – pierwszy z nich na powierzchni próbki, drugi w odległości 18 mm od krawędzi płytki, na głębokości 10 mm. W odległości 36 mm od termoelementu 2 znajdowała się linia pomiaru twardości (rys. 2).

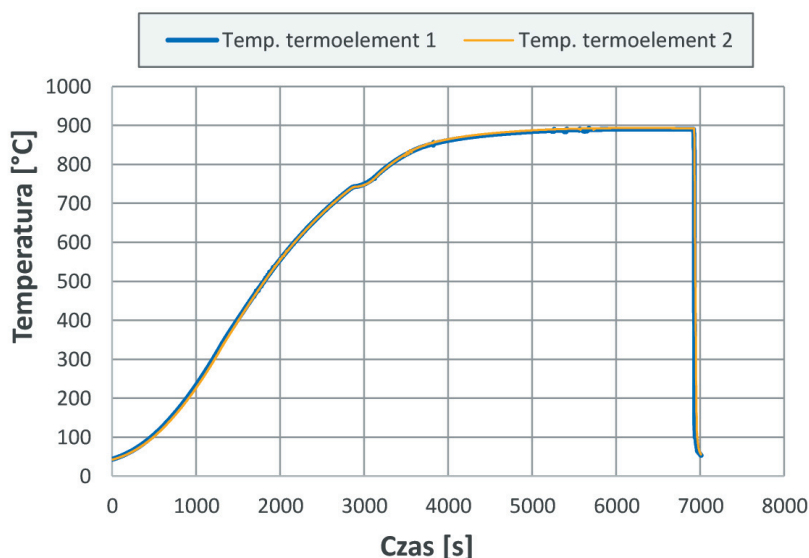


Rys. 2. Model przestrzenny próbki z zaznaczonymi kluczowymi wymiarami oraz widok próbki wykorzystanej do badań rzeczywistych

Próbkę poddano obróbce cieplnej o parametrach tożsamych z zadanymi w symulacji, a więc hartowaniu zwykłemu. Procedura składała się z następujących kroków:

1. nagrzanie próbki do temperatury 890°C,
2. wygrzewanie próbki w zadanej temperaturze przez 30 minut,
3. chłodzenie w wodzie (pojemnik o objętości 8 dm³).

Obróbkę cieplną przeprowadzono w elektrycznym piecu oporowym, o objętości komory roboczej 7 dm³, typu FCF7. Zarejestrowane za pomocą termoelementów zmiany temperatury podczas obróbki cieplnej zostały zobrazowane na wykresie (rys. 3).



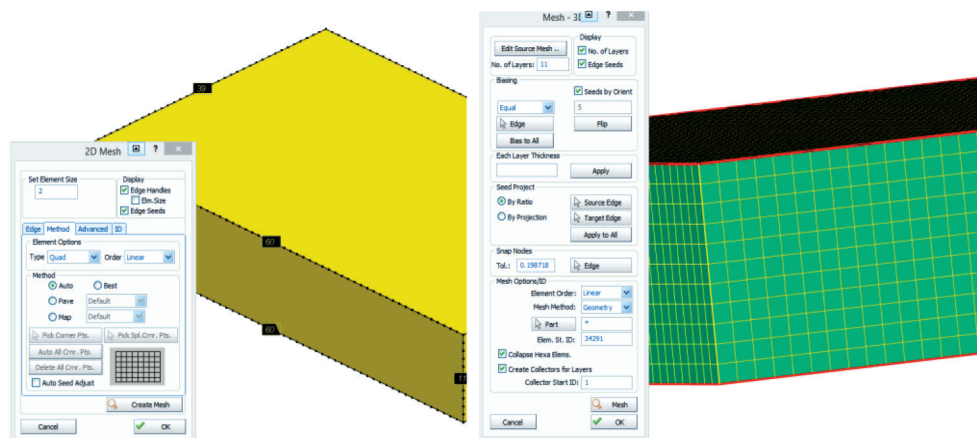
Rys. 3. Zmiany temperatury w trakcie obróbki cieplnej

Następnie przygotowano próbkę do badań metaloznawczych. Płytkę została przecięta w płaszczyźnie prostopadłej do powierzchni, która przechodziła przez punkt mocowania termoelementu 2, zgodnie ze schematem pokazanym na rysunku 2. Powierzchnię próbki wyrównano i wykonano dwie serie pomiarów twardości w punktach znajdujących się w jednakowych odstępach od siebie. Z wyników wyznaczono średnią. Badania wykonywano pod obciążeniem 30 kg. Po zakończonych badaniach twardości wykonano zgląd metalograficzny, który został wytrawiony nitałem. Następnie z zastosowaniem mikroskopu świetlnego LEICA poddano analizie mikrostrukturę tak przygotowanego materiału. Wykonano po pięć zdjęć w dwóch powiększeniach, tj. 100× i 500×, dla każdego z charakterystycznych punktów: na krawędzi próbki (na powierzchni zahartowanej), 5 mm od krawędzi oraz 10 mm od krawędzi (w środku przekroju). Analizę obrazu wykonano za pomocą aplikacji ImageJ.

W celu przeprowadzenia symulacji utworzono projekt w programie Visual-Weld. Kolejno wykonano następujące operacje:

– W module Mesh:

Przygotowano model wirtualny odzwierciedlający geometrię badanej próbki. Najpierw wygenerowano siatkę powierzchniową 2D, a na jej podstawie siatkę objętościową 3D. Rozmiar elementów siatki powierzchniowej ustalono na 2 mm, zaś typ jako sześcienny (*quad*), o stałym zagęszczeniu w każdym kierunku, tj. liniowym. Utworzono kolektory VOLUME (objętość bryły reprezentowana przez model 3D) oraz QUENCHED (powierzchnia, na którą działa medium robocze, tworzona przez elementy siatki 2D). Zdefiniowano klamry, będące pojedynczymi węzłami siatki, służące do unieruchamiania próbki podczas chłodzenia. Na końcu sprawdzono poprawność siatek, upewniono się, że projekt nie zawiera wolnych węzłów, niepożądanych zbieżności oraz nieciągłości siatki, i oceniono jakość elementów (sprawdzano m.in. wartość wyznacznika macierzy Jacobiego siatki 2D). Uzyskany model przedstawiono na rysunku 4.

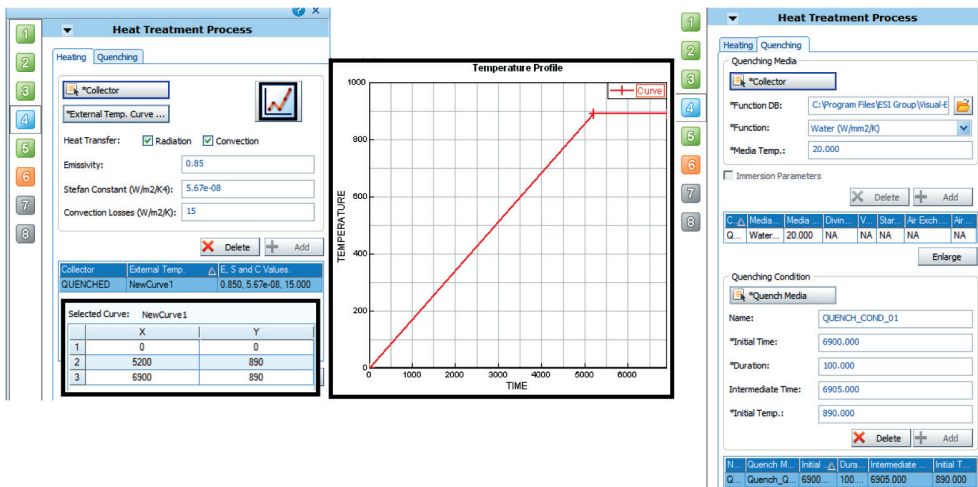


Rys. 4. Model oraz parametry siatek 2D i 3D

– W module Heat-Treatment (H-T):

Uruchomiono narzędzie Advisor, zaznaczono zabiegi obróbki cieplnej, które program ma uwzględnić w obliczeniach. Celem niniejszej pracy było porównanie wyników hartowania, zatem zadano nagrzewanie (*heating*) oraz chłodzenie (*quenching*). Z bazy materiałowej programu wybrano stop, z którego wykonana jest próbka (stal CF35). Krzywą temperaturową nagrzewania oraz parametry chłodzenia zadano na podstawie zmian temperatur zarejestrowanych podczas obróbki cieplnej (rys. 3). Nagrzewanie, trwające łącznie 6900 sekund, odbywało się dwuetapowo: przez 5200 sekund próbka była nagrzewana do temperatury 890°C, następnie wytrzymywana w takich warunkach przez kolejne 1700 sekund. Do chłodzenia zastosowano wodę o początkowej temperaturze wynoszącej 20°C. Czas zanurzenia w wodzie ustalono na

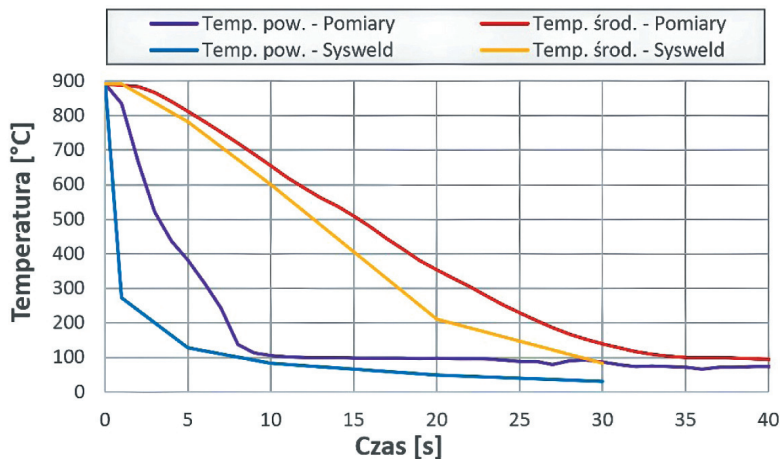
100 sekund, zaś maksymalną różnicę czasową między wyjęciem próbki z pieca a zanurzeniem w medium chłodzącym na 5 sekund. Za pomocą przygotowanych klamer zablokowano stopnie swobody próbki. W panelu funkcyjnym zaznaczono opcję obliczeniową twardości, po czym uruchomiono symulację. Ustalone parametry procesu przedstawiono na rysunku 5.



Rys. 5. Parametry procesu: nagrzewania (po lewej) i chłodzenia (po prawej)

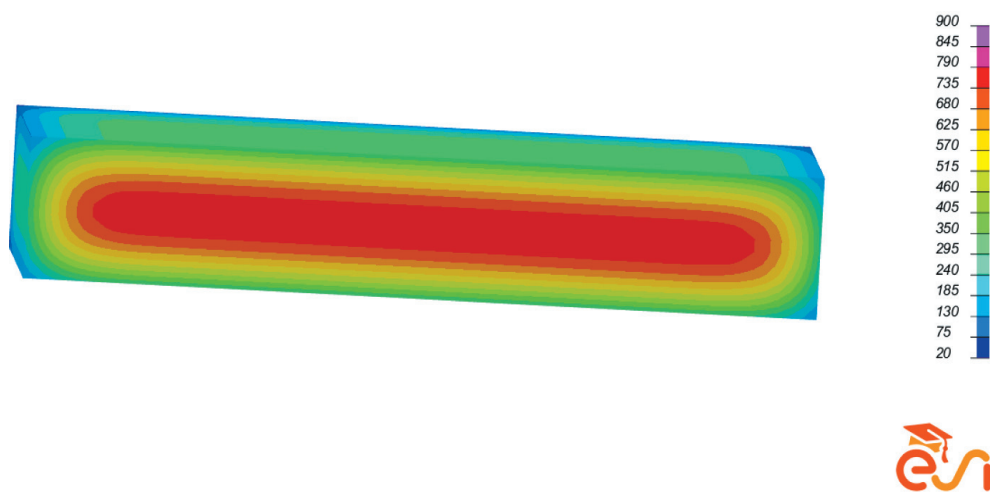
5. Wyniki badań

Na rysunku 6 przedstawiono krzywe stygnięcia zmierzone doświadczalnie i uzyskane na podstawie obliczeń numerycznych w kilku wybranych krokach czasowych (1 s, 5 s, 10 s, 20 s, 30 s).



Rys. 6. Zestawienie krzywych stygnięcia z symulacji oraz pomiarów rzeczywistych

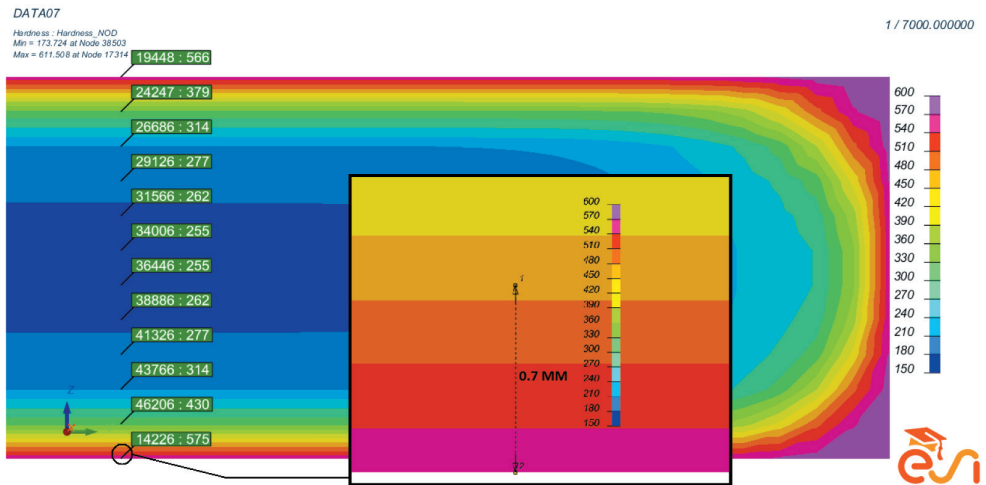
Odnotowuje się różnicę w szybkościach stygnięcia w przypadku danych z programu Sysweld oraz danych empirycznych. Obie krzywe z symulacji wykazują znacznie większą dynamikę chłodzenia (5–10 s) w porównaniu z krzywą opartą na danych pomiarowych. W przypadku termoelementu 1 umieszczonego na krawędzi próbki stwierdzony eksperymentalnie spadek temperatury w pierwszej sekundzie chłodzenia wynosi niewiele ponad 55°C, zaś w symulacji aż 618°C. Inaczej jest w środku przekroju, gdzie różnice są znacznie mniejsze. Po 5 sekundach odnotowano w symulacji 783°C, zaś w pomiarach rzeczywistych 812°C. Szczególny wpływ na twardość powierzchni zahartowanej mają pierwsze sekundy procesu chłodzenia. Niestety dyferencja wyników jest największa w tym zakresie pomiarowym i ma kluczowy wpływ na ilość utworzonego martenzytu, a zatem i na twardość elementu po obróbce cieplnej, co znajduje swoje potwierdzenie w uzyskanych rozkładach twardości. W obu przypadkach odprowadzenie ciepła z powierzchni próbki było znacznie szybsze niż z wewnętrznej jej części. Czas potrzebny do osiągnięcia temperatury 100°C wyniósł 8 sekund w przypadku symulacji oraz 12 sekund w doświadczeniu rzeczywistym. Rozbieżności w szybkości chłodzenia istotnie wpłynęły na niezgodność wyników dotyczących składu fazowego oraz twardości na przekroju próbki w symulacji i w eksperymencie. Na rysunku 7 przedstawiono uzyskany w wyniku symulacji rozkład temperatury na przekroju próbki po 5 sekundach od rozpoczęcia chłodzenia.



Rys. 7. Rozkład temperatury na przekroju próbki po 5 sekundach chłodzenia

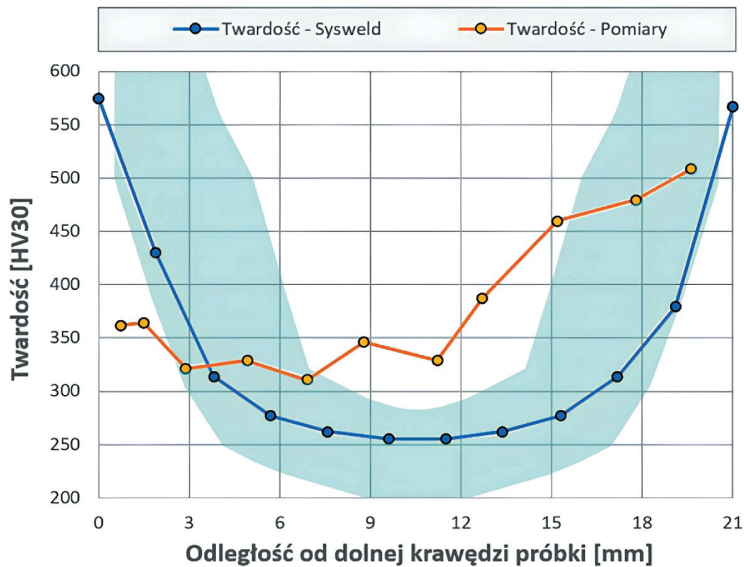
Powierzchnia próbki po tym czasie osiągnęła temperaturę niższą od M_s (400°C dla stali CF35) i w tym obszarze rozpoczęła się przemiana martenzytyczna. W wewnętrznej części próbki temperatura przekracza 700°C, co umożliwia rozpad dyfuzyjny austenitu i rozpoczęcie się przemiany perlitycznej.

Na rysunku 8 przedstawiono zmiany twardości na przekroju próbki uzyskane w wyniku symulacji numerycznej. Wyszczególniono obszar pomiarowy w odległości 0,7 mm od powierzchni hartowanej, gdzie zaobserwowano wyniki twardości z zakresu 480–510 HV30.

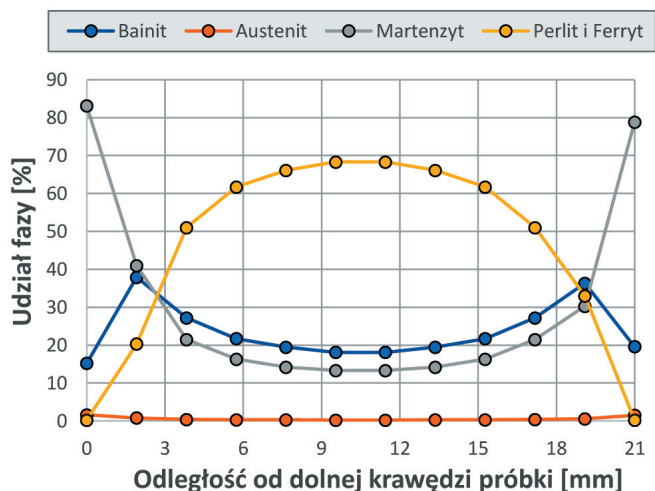


Rys. 8. Rozkład twardości na przekroju z zaznaczonym obszarem na głębokości 0,7 mm

Na rysunku 9 zestawiono wyniki symulacji twardości z danymi uzyskanymi podczas doświadczenia. Dodatkowo na wykresie zacieniowano kolorem turkusowym zakres twardości podawany dla stali CF35 w normie (PN-EN ISO 683-1:2018-09).



Rys. 9. Zestawienie wyników symulacji twardości z wynikami jej pomiarów rzeczywistych

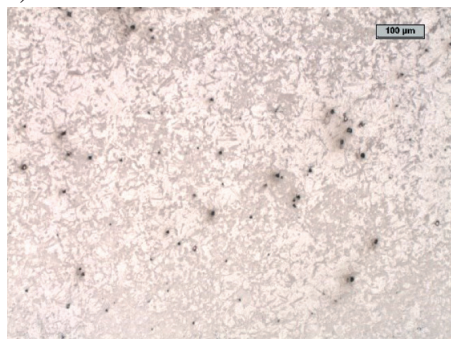


Rys. 11. Porównanie udziału poszczególnych faz na przekroju próbki – symulacja Sysweld

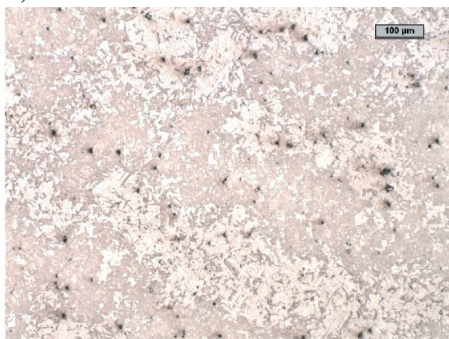
Wybrane wyniki badań metalograficznych zestawiono na rysunku 12. Zdjęcia wykonano w powiększeniach 100 $\times$. Odległość badanych obszarów od powierzchni stygnięcia wynosiła <1 mm, 5 mm i 10 mm. Analiza rysunku 12 wskazuje, że przy powierzchni chłodzenia udział martenzytu jest wyraźnie większy niż w osi symetrii próbki. Cyfrowe badania obrazu wykonane za pomocą programu ImageJ potwierdziły te obserwacje. Ich wyniki przedstawiono na rysunku 13.

Obserwuje się rozbieżne wartości udziału martenzytu w każdym z badanych punktów. Różnica między wynikiem symulacji a wynikami doświadczalnymi udziału martenzytu na krawędzi próbki, tj. w odległości <1 mm od powierzchni hartowanej, wynosi ponad 10%, a w odległości 5 mm od powierzchni – 14%. Trend utrzymuje się także w środku przekroju, gdzie odnotowuje się różnicę wynoszącą ok. 13%.

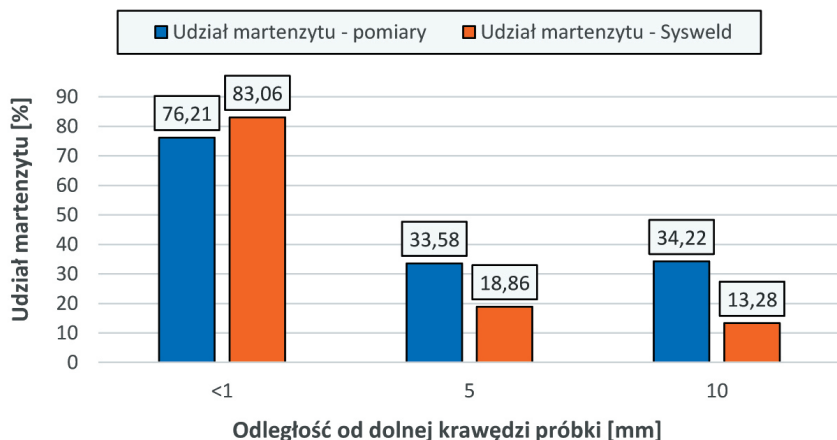
a)



b)



Rys. 12. Mikrostruktura badanych próbek w odległości do 1 mm (a) i 10 mm (b) od powierzchni chłodzenia



Rys. 13. Zestawienie wyników symulacji rozkładu martenzytu z wynikami doświadczalnymi

Pomimo porównywalnej hartowności staliwa L500-II i stali CF35 uzyskane twardości przy powierzchni są mniejsze. Zjawisko to można tłumaczyć mniejszą gęstością ziaren w staliwie w porównaniu ze stałą, a zgodnie z zależnością Halla–Petcha im większe ziarna, tym mniej wytrzymały i twardy jest materiał. Inną przyczyną mniejszej twardości staliwa w porównaniu ze stałą może być większa zawartość wtrąceń niemetalicznych (tlenków, siarczków, azotków), które są preferowanymi miejscami zarodkowania perlitu. Wydaje się, że próbka poddana badaniom empirycznym powinna także zostać poddana procesowi homogenizacji, jednak w warunkach produkcyjnych zabieg ten jest niezwykle rzadko przeprowadzany względem odlewów ze staliwa węglowego, toteż nie przeprowadzono go w opisywanych badaniach.

Należy zauważyć, że przyczyną rozbieżności może być również model matematyczny wykorzystywany w oprogramowaniu Visual-Weld. Z powodów praktycznych tego typu oprogramowanie zwykle zawiera pewne uproszczenia niezbędne do przyspieszenia obliczeń. W badaniach prezentowanych w niniejszej pracy potwierdziły się obserwacje Kika i in. (2015), którzy wskazali na rozbieżność pomiędzy wynikami symulacji w środowisku Visual-Weld oraz wynikami rzeczywistymi uzyskanymi podczas hartowania w wodzie. Otrzymane dane wskazują, że program ma trudności z prawidłowym wyznaczeniem kinetyki powstawania martenzytu, za co może odpowiadać przyjęty model Koistinen i Marburgera (1959). Obecnie w literaturze dostępny jest model Li i in. (1998), który dodatkowo uwzględnia wielkość ziarna austenitu w obliczeniach szybkości powstawania martenzytu. Wydaje się, że na potrzeby symulacji procesów przemian fazowych w odlanych stopach żelaza należałoby dodatkowo wprowadzić parametr opisujący liczbę wtrąceń niemetalicznych, które także mają niebagatelny wpływ na hartowność staliwa.

6. Wnioski

Wyniki symulacji obróbki cieplnej w programie Visual-Weld są rozbieżne z uzyskanymi w doświadczeniu rzeczywistym, natomiast odpowiadają literaturowym, co skłania do wniosku, że zastosowanie oprogramowania grupy ESI w przemyśle odlewnictwa staliwa jest zasadne. Za przyczynę znacznych rozbieżności pomiędzy wynikami uzyskanymi doświadczalnie i w symulacji, zwłaszcza w obszarze bliskim powierzchni zahartowanej, można uznać następujące czynniki:

- staliwo w przeciwieństwie do stali charakteryzuje się niejednorodną i gruboziarnistą mikrostrukturą, która powstaje łatwo szczególnie w odlewach grubościennych; owa niejednorodność w połączeniu z segregacją składników stopowych mogła mieć wpływ na otrzymane wyniki doświadczalne;
- zaimplementowany w oprogramowaniu Visual-Weld model nie uwzględnia niektórych czynników wpływających na hartowność, należą do nich: wielkość ziarna, stopień jednorodności składu chemicznego austenitu, obecność obcych faz takich jak tlenki i siarczki, których udział objętościowy w staliwie zwykle przekracza 0,1%, a które ułatwiają rozpad dyfuzyjny austenitu;
- niższa twardość próbek przy powierzchni chłodzenia jest wynikiem istotnie mniejszej szybkości chłodzenia uzyskanej w doświadczeniu, ponadto wykorzystany zbiornik wody, z powodu stosunkowo niedużej objętości, nie zapewnił stabilnej temperatury podczas chłodzenia.

Literatura

- Adrian H., 2011, *Numeryczne modelowanie procesów obróbki cieplnej*, Wydawnictwa AGH, Kraków.
- Babu K., Prasanna Kumar T.S., 2014, *Comparison of austenite decomposition models during finite element simulation of water quenching and air cooling of AISI 4140 steel*, Metallurgical and Materials Transactions B, vol. 45, s. 1530–1544. <https://doi.org/10.1007/s11663-014-0069-0>.
- Bhadeshia H.K.D.H., 2014, *Physical Metallurgy of Steels*, [w:] Laughlin D.E., Hono K. (eds.), *Physical Metallurgy*, 5th ed., Elsevier, s. 2157–2214. <https://doi.org/10.1016/B978-0-444-53770-6.00021-6>.
- ESI Group, 2017. *ESI SYSWELD: Welding and Heat Treatment Simulation Solution*. <http://www.cef3d.fr/COSS/wp-content/uploads/2018/04/Heat-Treatment-ESI-solution.pdf> [dostęp: 4.08.2023].

- Evdokimov A., Jasiewicz F., Doynov N., Ossenbrink R., Michailov V., 2022, *Simulation of surface heat treatment with inclined laser beam*, Journal of Manufacturing Processes, vol. 81, s. 107–114. <https://doi.org/10.1016/j.jmapro.2022.06.051>.
- Głównia J., 2002, *Odlewy ze stali stopowej*, Fotobit, Kraków.
- Kik T., Slovacek M., Moravec J., Vanek M., 2015, *Numerical simulations of heat treatment processes*, [w:] *Applied Mechanics and Materials*, vols. 809–810, Trans Tech Publications, s. 799–804. <https://doi.org/10.4028/www.scientific.net/amm.809-810.799>.
- Kik T., Moravec J., Nováková I., 2020, *Wsparcie numeryczne badań procesów obróbki cieplnej*, [w:] Górka J. (red.), *Symposium Katedr i Zakładów Spawalnictwa „Nowoczesne zastosowania technologii spawalniczych”: Brenna 15–17 czerwca 2020 r.*, Archives of Foundry Engineering, Katowice – Gliwice, s. 25–36.
- Koistinen D., Marburger R., 1959, *A general equation prescribing the extent of the austenite-martensite transformation in pure iron-carbon alloys and plain carbon steels*, Acta Metallurgica, vol. 7, s. 59–60. [https://doi.org/10.1016/0001-6160\(59\)90170-1](https://doi.org/10.1016/0001-6160(59)90170-1).
- Leblond J.B., Devaux J., 1984, *A new kinetic model for anisothermal metallurgical transformations in steels including effect of austenite grain size*, Acta Metallurgica, vol. 32(1), s. 137–146. [https://doi.org/10.1016/0001-6160\(84\)90211-6](https://doi.org/10.1016/0001-6160(84)90211-6).
- Lee S.H., Kim E.S., Park J.Y., Choi J., 2018, *Numerical analysis of thermal deformation and residual stress in automotive muffler by MIG welding*, Journal of Computational Design and Engineering, vol. 5(4), s. 382–390. <https://doi.org/10.1016/j.jcde.2018.05.001>.
- Li M.V., Niebuhr D.V., Meekisho L.L., Atteridge D.G., 1998, *A computational model for the prediction of steel hardenability*, Metallurgical and Materials Transactions B, vol. 29, s. 661–672. <https://doi.org/10.1007/s11663-998-0101-3>.
- Lu S., Zhou J., Zhang J., 2015, *Optimization of welding thickness on casting-steel surface for production of forging die*, International Journal of Advanced Manufacturing Technology, vol. 76, s. 1411–1419. <https://doi.org/10.1007/s00170-014-6371-9>.
- PN-EN ISO 683-1:2018-09, *Stale do obróbki cieplnej, stale stopowe i stale automatowe – Część 1: Stale niestopowe do hartowania i odpuszczania*.
- Pont D., Guichard T., 1995, Sysweld®: *Welding and Heat Treatment Modelling Tools*, [w:] Atluri S.N., Yagawa G., Cruse T. (eds.), *Computational Mechanics '95*, Springer, Berlin, Heidelberg, s. 248–253. https://doi.org/10.1007/978-3-642-79654-8_41.
- Ryzhkov M.A., Maisuradze M.V., Yudin Y.V., Khuppeev A.V., Ajeet Babu K.P., 2015, *Experience in improving silicon steel component heat treatment quality*, Metallurgist, vol. 59, s. 401–405. <https://doi.org/10.1007/s11015-015-0117-2>.
- Ryzhkov M., Maisuradze M., Kaletin A., 2018, *Heat Treatment Technology Adjustment Using Experimental and Simulation Methods*, [w:] Syngellakis S., Connor J. (eds.), *Advanced Methods and Technologies in Metallurgy in Russia*, Innovation and Discovery in Russian Science and Engineering, Springer, Cham, s. 69–75. https://doi.org/10.1007/978-3-319-66354-8_9.

- Stanisławski M., 2008, *Oprogramowanie CAD, CAM, CAE dla inżynierów projektantów branży mechanicznej*, Projektowanie i Konstrukcje Inżynierskie, nr 7/8 (10/11), s. 9–10.
- Swacha P., Kotyk M., Stachowiak R., Ziółkowski W., 2021, *Analizy numeryczne procesów spawania i obróbki cieplnej z wykorzystaniem oprogramowania SYSWELD*, Developments in Mechanical Engineering, vol. 17(9), s. 47–60. <https://doi.org/10.37660/dme.2021.17.9.4>.
- Zubairuddin M., Albert S.K., Reddy M.S., Baharin Ali, Varaprasad B., Mishra A., Das P.K., Elumalai P.V., 2023, *Thermal analysis of thin P91 steel using FlexPDE and SYSWELD*, Materials Today: Proceedings, vol. 72(part 3), 2023, s. 1550–1555. <https://doi.org/10.1016/j.matpr.2022.09.385>.

