Selected problems in mechanical engineering 2022

Editors:

Magdalena Mieloszyk

Gdańsk 2024

Pod redakcją/ Editors:
dr hab. inż. Magdalena Mieloszyk, prof. IMP PAN

Projekt okładki:
Mirosław Sawczak

Skład komputerowy LaTeX, łamanie, adjustacja:
Paweł Kudela, Katarzyna Majewska, Magdalena Mieloszyk

ISBN 978-83-66928-09-1

WYDAWNICTWO INSTYTUTU MASZYN PRZEPŁYWOWYCH PAN
Instytut Maszyn Przepływowych im. Roberta Szewalskiego
Polskiej Akademii Nauk
ul. Fiszera 14, 80-231 Gdańsk
tel. (+48) 58-52-25-141; fax (+48) 58-341-61-44
e-mail: redakcja@imp.gda.pl https://www.imp.gda.pl/wydawnictwo

Recenzenci monografii/ Reviewers of the monograph:
Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie/
AGH University of Science and Technology
dr hab. inż. Paweł Madejski, prof. AGH

Instytut Maszyn Przepływowych Polskiej Akademii Nauk/
Institute of Fluid Flow Machinery, Polish Academy of Sciences
prof. dr hab. inż. Alicja Krella
prof. dr hab. inż. Janusz Badur
dr hab. inż. Mariusz Banaszkiewicz
dr hab. inż. Jan Cebula
dr hab. inż. Paweł Flaszyński, prof. IMP PAN
dr hab. inż. Paweł Kudela, prof. IMP PAN
dr hab. Katarzyna Majewska, prof. IMP PAN
dr hab. Paweł Malinowski, prof. IMP PAN
dr hab. inż. Robert Matysko
dr hab. inż. Polesek-Karczewska, prof. IMP PAN
dr hab. inż. Tomasz Ochrymiuk, prof. IMP PAN
dr hab. inż. Maciej Radzieński, prof. IMP PAN
dr hab. inż. Mirosław Sawczak, prof. IMP PAN
dr hab. inż. Ryszard Szwaba, prof. IMP PAN
dr hab. inż. Tomasz Wandowski, prof. IMP PAN
dr inż. Mohammad Ali Fakih
dr inż. Ksawery Kuligowski
dr inż. Tomasz Przybyliński

Kaunas University of Technology, Lithuania
dr. Rūta Rimasaŭskienė

Politechnika Śląska/ Silesian University of Technology
prof. dr hab. inż. Małgorzata Jastrzębska
dr inż. Andrzej Sachajdak

Contents

Przedmowa

W 2019 roku została utworzona Trójmiejska Szkoła Doktorska Polskiej Akademii Nauk (TSD PAN) prowadzona wspólnie przez:

- Instytut Maszyn Przepływowych PAN (IMP PAN),
- Instytut Budownictwa Wodnego PAN (IBW PAN),
- Instytut Oceanologii PAN (IO PAN).

Szkoła Doktorska oferuje kształcenie w ramach trzech dyscyplin: inżynierii mechanicznej, inżynierii lądowej, geodezji i transportu oraz nauki o Ziemi i środowisku.

Mam ogromną przyjemność przedstawić Państwu monografię będącą podsumowaniem aktywności naukowej kolejnego roku działalności TSD PAN zawierającą prace naukowe zaproponowane przez doktorantów Szkoły Doktorskiej.

Niniejsza monografia zawiera zbiór prac z dyscypliny inżynieria mechaniczna. Ich krótkie podsumowanie doktoranci TSD PAN zaprezentowali w ramach *III Seminarium Naukowego TSD PAN*, które odbyło się w dniach 13-14 czerwca 2022 r. Wydarzenie to jest kontynuacją cyklu corocznych otwartych seminariów naukowych, na których doktoranci Szkoły Doktorskiej prezentują swoje osiągnięcia naukowe.

dr hab. inż. Magdalena Mieloszyk, profesor IMP PAN
Dyrektor TSD PAN

Preface

In 2019 the Tricity Doctoral School of the Polish Academy of Sciences (TSD PAN) was established, which is run jointly by:

- Institute of Fluid-Flow Machinery PAS (IMP PAN),
- Institute of Hydro-Engineering PAS (IBW PAN),
- Institute of Oceanology PAS (IO PAN).

The Doctoral School offers education in three disciplines: mechanical engineering, civil engineering, geodesy and transport as well as Earth and environmental science.

It is my great pleasure to present you a monograph which is a summary of the next year of TSD PAN scientific activity, including scientific papers proposed by PhD students of the Doctoral School.

This monograph contains a collection of papers from the discipline of mechanical engineering. Their short summaries were presented by PhD students during the *3rd Scientific Seminar of TSD PAN*, which took place on June 13^{th}-14^{th}, 2022. This event is a continuation of a series of annual open scientific seminars where PhD students of the Doctoral School present their scientific achievements.

dr hab. inż. Magdalena Mieloszyk, profesor IMP PAN

Director of TSD PAN

Chapter 1

The study of Guided Wave mode sensing and separation based on FBG and PZT sensors for structural health monitoring

Sultan Ahamad

Institute of Fluid Flow Machinery, Polish Academy of Sciences,
Fiszera 14, 80-231 Gdansk, Poland

1.1 Introduction

Structural health monitoring (SHM) is an approach for evaluating and monitoring the structures like aeronautical, mechanical, and civil infrastructures. It is the process of monitoring changes in the properties of engineering structures (such as bridges and buildings) over time using periodically sampled response measurements. A well-designed SHM system would evaluate a structure's serviceability, reliability, and durability, allowing owners or decision-makers to deploy resources efficiently. The adoption of the SHM allows for frequent in-service inspections of structures, allowing for the identification of structures that would require maintenance, repair, refit, or replacement regularly. The outcome of this procedure for long-term SHM has frequently updated information on the structure's capacity to execute its intended function in light of the inevitable age and degradation caused by operational settings. SHM is very important for quick condition screening after severe events like earthquakes or blast loads to provide trustworthy information on the structure's integrity in near real-time [1]. It is necessary to carry out real-time health monitoring and evaluation of engineering structures to put an end to the potential hazards and improve the safety performance of civil facilities. It is required to find features in the gathered data that allow one to distinguish between the undamaged (referential or pristine) and damaged structure in order to monitor the condition of a system in real-time. Correlating observed system response parameters, such as vibration amplitude or frequency, with observations of the deteriorated system is one of the most used approaches. SHM systems can help decrease the number of unplanned repair and service operations due to structural deterioration of the utilized structure. The guided wave (GW) propagation approach is one of the numerous diagnostic procedures that can be used in SHM [2].

SHM process involves the sensor/actuator network (different types, numbers, and locations), data acquisition/storage/transmittal hardware, and data processing system/software (presentation of identification results) [3]. Measurements and their processing can be made to either immediately identify any degradation or damage to

a system or indirectly forecast the system's status by assessing the size and frequency of loads experienced. In structural health monitoring (SHM), guided wave-based inspection has been used extensively. Guided-wave propagation is affected by structural damage. Since the detected signal differs from the reference signal of the undamaged structure, the presence of damage can be determined. Inspection of structures can be performed using the piezoelectric sensors but these have a serious drawback of electromagnetic interference [4]. FBG sensors are also used for guided-based SHM and these are immune to electromagnetic interference but relatively very costly as compared to PZT sensors [5]. FBG sensors have a number of potential advantages over traditional piezoelectric sensors, including their small size, immunity to electromagnetic interference, excellent potential for embedding, and multiplexing abilities [6].

1.2 Techniques used in SHM

SHM techniques have been developed for a long time [2]. SHM techniques have the goals of establishing the presence of structural problems (detection), locating (localization) and assessing the degree of structural damage (determination of size/shape/type), evaluating structure safety, and forecasting the structure's remaining service life (prediction), and, if practicable, deciding on a maintenance strategy. Several SHM techniques have been proposed in the literature, including vibration-based, strain-based, guided waves (GW)-based, electromechanical impedance-based, etc. These use various types of sensors and signal processing techniques to reliably identify damage. These approaches have been effectively used for experimental validation, however, they do not meet all of the service criteria in real-world applications. As a result, many kinds of research on low-cost methods are going on [7, 8]. Some techniques for SHM will be described in further subsections.

1.2.1 Vibration based SHM

The vibration-based SHM approaches are very popular. Many researchers like Boscato *et al.* have presented detailed descriptions of

Structural Health Monitoring using Vibration based approaches [9]. The goals of these SHM techniques are to determine the presence of structural defects, as well as the location and size of such damage, determine the extent of structural damage, assess the structure's safety, and estimate the remaining service life. If practicable, assess the structure and decide on a maintenance strategy. The truth is that the structure's vibration characteristics are determined by its physical parameters. Structural damage creates changes in the structure's physical/mechanical parameters, and changes in the structure's physical/mechanical parameters induce changes in the structure's vibration response. As a result, these parameters can be utilized as indications of structural health. Vibration-based SHM approaches may be divided into three groups based on the vibration parameters used: methods in the frequency domain, time domain, and time-frequency domain [10, 11].

1.2.2 Strain based SHM

Strain monitoring data from SHM activities can directly connect to stress redistribution in the region of the strain sensor, such as damage initiates and grows, hence strain-based damage detection systems are valued highly in structural health diagnostics [12]. Damage detection techniques were developed by dividing structural strain monitoring data into quasi-static and dynamic (high-frequency) components. When damage is triggered, a considerable change in strain data or strain-based feature data will arise, according to theoretical techniques [13]. The strain-based damage detection systems, however, confront the aforesaid second problem of SHM data interpretation due to the homogeneous properties of structural responses that reflect changing environmental and operating situations [14]. To prevent disguising the consequences of structural degradation, time-varying implications on strain responses should be taken into account [15].

1.2.3 Electromechanical impedance-based techniques

Measurements of electromechanical impedance (EMI) is one of the SHM approaches. As actuators and sensors, it uses piezoelectric

transducers, which are mostly built of piezoelectric ceramic. The electrical impedance of a piezoelectric transducer is measured due to the existence of electromechanical coupling of the transducer and structure [16]. In the EMI technique, a high-strength epoxy adhesive is used to embed or add a lead-zirconate titanate piezoelectric ceramic sensor (PZT) patch to the structure being monitored. In the high-frequency range (30-400 kHz), the patch's conductance characteristic is obtained. This characteristic is a reference for SHM. The conductance characteristic is re-acquired and compared to the reference conductance characteristic whenever it is necessary to assess the structure's health. Typically, a commercially available impedance analyzer is used to determine the conductance characteristic of the bonded/embedded PZT patch [17].

1.3 Guided wave (GW) and GW Imaging

Multiple physical quantities can be measured to assess a structure's health, with Guided Waves (GWs) being particularly promising. Among the benefits of GW for SHM, it should be noted that they are extremely sensitive to defects and propagate over long distances with little input energy, making them ideal for an SHM configuration (limited number of sensors). The main disadvantages of GW are that they are sensitive to boundary conditions, multi-modal and dispersive, making analysis difficult, and they are sensitive to environmental conditions. A considerable body of work exists on GW-based SHM (GW-SHM), and a variety of algorithms have been developed to identify a variety of defects in diverse structures. GW imaging (GWI) is the process of investigating structure by measuring propagating GWs [18, 19]. GWI usually, but not always, uses an array of piezoelectric (PZT) transducers, each acting sequentially as an emitter and receiver of GW to measure the propagated wave packets between each pair of sensors. In this case, for this configuration, it is anticipated that the existence of a defect will alter the wave propagation in some way, for instance, consider the presence of a reflected wave packet. Acoustic emission, pitch-catch, pulse-echo, beamforming, and tomography are some of the other GW approaches. The current limitations of GW-SHM techniques stem

from the fact that GWs are extremely sensitive to a large number of parameters, making it difficult to distinguish the influence of defects in the measured signals from the influence of other parameters such as temperature variation or different boundary conditions [20].

1.4 Overview of SHM based on elastic/ guided waves

When particles are moved in a medium, elastic waves occur in which a force proportional to the displacement acts on the particles to return them to their previous location. An elastic wave will propagate if a material possesses the property of elasticity and the particles in a specific region are placed in vibratory motion. For example, a gas, is an elastic medium (it will return its original volume if the pressure is released), and sound is conveyed through gas as an elastic wave. Guided waves are the elastic waves that propagate between surfaces. The energy of waves is concentrated near a boundary or between parallel boundaries separating different materials and that has a direction of propagation parallel to these boundaries. Lamb waves are a type of elastic wave. These propagate in thin solid plates with parallel surfaces or spheres [21]. These waves are elastic waves with particle motion in the plane containing both the wave propagation direction and the plane normal (the direction perpendicular to the plate). An infinite medium can only host two types of wave modes propagating at different velocities; however, plates may support two unlimited sets of Lamb wave modes, whose velocities are determined by the wavelength-plate-thickness connection. The Rayleigh waves, which propagate along a single surface, are referred to as Rayleigh Lamb waves. Rayleigh and Lamb waves are both bound by the elastic characteristics of the guiding surface(s). Lamb waves have come a long way in terms of knowledge and application during the 1990s, due to the fast expansion in processing capacity. Theoretical formulations by Lamb have been discovered.

In recent times, there are a large number of studies related to SHM in which the elastic waves have been analyzed, for example, in one of the recent research by Martn-Sanz *et al.*, the research team uses an

actual bridge construction to demonstrate the benefits of ultra-high-performance fibre-reinforced cement-based composites (UHPFRC). The research team was able to test the bridge both statically and dynamically before and after repair with (UHPFRC) [22]. Chakraborty *et al.* presented an approach for damage detection in multiple RC structures based on embedded ultrasonic sensors and wavelet transform [23]. Ricci *et al.* have focused on guided waves for SHM in composites [24]. Elwalwal *et al.* described the crack inspection using the guided waves [25]. In a recent research, a two-step guided waves-based damage localization technique has been presented [26]. In another recent research, guided wave propagation has been used for SHM using fibre bragg grating sensors [27]. Not only the simple analysis of waves but there is a great role of mode separation of waves for SHM, so many researchers have been performed using mode separation [28–30].

Non-destructive testing (NDT) and structural health monitoring (SHM) of engineering structures such as oil and gas pipelines, railroads, aircraft components, adhesive bonding, and perhaps much more are of significant interest to guided waves (GW). As long as the structure's cross-section is constant and the differential in acoustic impedance to the surrounding environment is high, these waves can propagate across quite long distances. Under certain conditions, it may also allow for the exploration of inaccessible portions of the building. The fundamental (i.e., first symmetric and antisymmetric) modes are primarily employed in the lower frequency range to reduce complexity in GW-based analysis since they are well understood and can be excited, measured, and analyzed without trouble. As a result, GW-based techniques are currently in development [31–34]. Ultrasonic guided wave method adopts an active manner to monitor structural conditions, which makes this method immune to most noise in condition monitoring. The ultrasonic-guided wave is excited at a well-selected frequency and interacts with the defects. The wave reflection, transmission, mode conversion, and energy loss can be used for damage detection [35].

1.5 Sensors used for SHM

Collecting precise and high-quality real-time measurements of structural element status, sharing this information with the control system, and signalling appropriate alerts if an unusual pattern is ever noticed are all critical components of SHM. Sensors for SHM are meant to make the monitoring process possible. A typical health monitoring system consists of a network of sensors that measure, strain, displacement, humidity, and temperature, among other characteristics important to the current status of the structure and its surrounding environment [36]. Here is a list of some sensors generally used in SHM: (1) Fiber Optic Sensors, (2) Linear Variable Differential Transformer, (3) Acoustic Emission Sensor, (4) Piezoelectric Sensor, (5) Electromagnetic Acoustic Transducer.

1.6 FBGs in SHM

Sensing techniques, utilizing PZT transducers have been extensively used due to the possibility of working as sensors and actuators although these are very low cost and have small mass/size. These have concerns with electromagnetic field sensitivity and high temperature (Curie temperature). Electromagnetic fields have little effect on optical fibre sensors [37]. Also, they have been considered a viable option for GW sensing owing to their tiny size, and ability to be embedded in structures. Fiber Bragg grating (FBG) optical strain sensors have the ability to be multiplexed. The multiplexing capabilities and wavelength-encoded measurement information of FBG sensors set them apart from other types of fibre optic sensors. Many tens of FBG sensors can be accommodated on a single optical fibre thread. Compared to other sensors, the FBG sensor systems are less vulnerable to signal amplitude variations since the measurement information is encoded in the wavelength, which is an absolute quantity. Other appealing aspect of FBG sensors is their intrinsic capacity to function as both a sensing element and a signal transmission medium, opening up new possibilities in the field of dependable remote SHM [28]. An FBG sensor, has all of the benefits that these devices are known for, including low loss relative

to fibre length, immunity to electromagnetic and radio frequency interference, small size and weight, intrinsically safe operation in hazardous materials environments, high sensitivity, and long-term reliability. Furthermore, fibre Bragg grating technology displays an inherent serial multiplexing capability as well as the potential to deliver absolute measurements without reference. As a result, it is a natural replacement for traditional electrical sensing methods.

1.6.1 Principles of Fiber-Optic Sensor (FOS)

The optical fibre sensor's operation includes external factor monitoring and signal transmission. Characteristic properties such as light intensity, phase, polarization state, wavelength, and frequency vary when light propagates along with an optical fibre. As a result, FOS can be classified as intensity-modulated, phase-modulated, or polarization-modulated. The modulation happens inside the FOS system, when light is delivered from the light source to the detector over an optical cable, the intrinsic FOS is what it's called [38]. Extrinsic FOS refers to modulation that happens outside of the optical fibre. According to whether the light interferes or not, FOS is separated into interferential FOS and non-interferential FOS. FOS may also be classified as point (local) FOS, quasi-distributed FOS, and distributed FOS based on the induction range. Fabry Perot fibre-optic sensor (FPFOS), fibre Bragg grating (FBG) sensor, optical time-domain reflectometer (OTDR), and long-period fibre grating (LPFG) sensor are the most widely employed in civil infrastructure [39–42].

1.7 Fiber Bragg Grating

FBG is a type of FOS sensor with Bragg grating. The Bragg grating is accomplished by transversely lighting the fibre with a UV laser beam and creating an interference pattern in its core with a phase mask. It results in a long-term alteration in the silica matrix's physical properties. This alteration is characterized by a spatially periodic modulation of the core index of refraction, which results in the formation of a resonant structure [43]. The FBG will

act as a wavelength selective mirror as a resonant structure; it is a narrow band filter. This implies that if light from a broadband source is introduced into the optical fibre, the grating will only back-reflect light with a very limited spectral breadth centered on the Bragg wavelength. The remaining light will not be lost as it travels through the optical cable to the next Bragg grating. The period of the microstructure and the index of refraction of the core determine the Bragg wavelength. Because the FBG is symmetric, it will always reflect light at the Bragg wavelength, regardless of which side it comes from.

The period of the microstructure (Λ) and the index of refraction of the core (n), generally defines the Bragg wavelength (λ).

$$\lambda = 2n\Lambda$$

Fiber Bragg grating properties make it ideal for use as a sensor. The FBG, for example, will detect strain when the fibre is stretched or compressed [44]. This occurs primarily as a result of the optical fibre's deformation, which causes a change in the microstructure's period and, as a result, the Bragg wavelength. The photoelastic effect, which is caused by changes in the refraction index, also contributes. An FBG temperature sensitivity is also inherent. The modification of the silica refraction index, generated by the thermo-optic effect, is the principal contributor to Bragg wavelength change in this situation. There is also a contribution from thermal expansion, which changes the microstructure's period. Given the low coefficient of thermal expansion of silica, this effect is minor.

1.7.1 EDGE Filtering approach for FBG

The general technique employs FBG sensors in the wavelength division multiplexing (WDM) approach, in that case, the measurement is transduced as a shift in the sensor's reflected wavelength. But in the case of the guided wave measurement, the sensitivity of this approach is very low. To overcome this issue a new trend of edge filtering approach is used. Fig. 1.1 shows the working principle of EDGE filtering. In this approach, a tunable laser with a wavelength

set on FBG's reflective slope is used. The reflective curve has a high slope but is linear in the middle section. As a result, a little change in reflectivity position causes a big change in optical power, which is sensed by the photodiode. The FBG sensors' sensitivity to GW is improved by this amplification. A large number of researchers have performed this approach for SHM because it has many advantages [45–48].

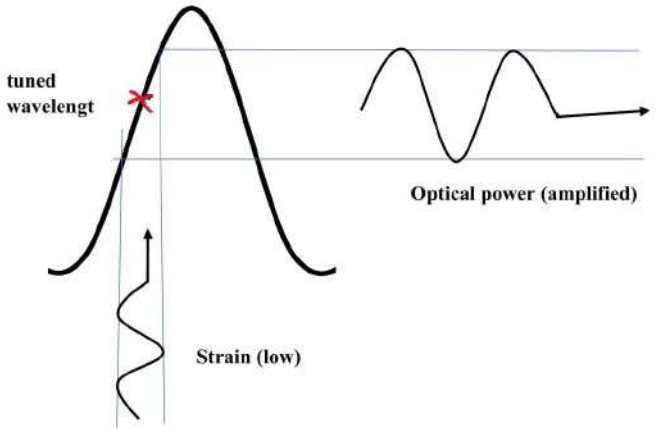


Fig. 1.1: Working principle of Edge Filtering approach.

Details of this phenomenon are explained by many researchers in their research papers [49, 50].

1.8 Methodologies for Mode separation of waves

Mode separation is a very popular technique in SHM. It can give very good results in SHM, and various advanced research studies are going on using the mode separation [51]. Multiple studies are available for mode separation or conversion [52, 53]. Numerical analysis of elastic wave mode conversion due to discontinuities was also

investigated in recent research [54]. Not only the simple mode separations but there are too many research studies on frequency-wave number-based mode filtering algorithms [55, 56]. These approaches enable mode separation as well as detecting forward or backward propagation, allowing us to determine whether a particular mode propagates forward or backward. However, full wavefield information is required to employ this approach. Signal processing approaches (like those developed by Draudviliene *et al.*) [57] attempt to circumvent the necessity for a large number of measurement points by limiting them to only two. There have been a few signal processing approaches created based on the chirplet transform [58], empirical mode decomposition [59], and its derivatives [60]. However, instead of explicitly distinguishing between the wave packets as symmetric or antisymmetric modes, these approaches essentially improve the signal's peaks by reducing high-frequency noise.

Similarly, for mode separation there is a large number of research available like Zhou *et al.* have presented a reconstruction-based mode separation method of Lamb wave for damage detection in plate structures, this paper presents a method to separate multi-mode signals of Lamb wave through reconstruction of pure mode basis signals calculated based on phase and group velocity [61]. On behalf of different review articles, Cheng and Fomel have described the Elastic wave mode separation in their research paper [62]. Here is a brief description of some methodologies which are generally used for mode separation- Kaur *et al.* presented a Fast Algorithm for Elastic Wave-Mode Separation using Deep Learning with Generative Adversarial Networks (GANs) [63]. In the Network Architecture and training section, they described that they use the paradigm of generative adversarial networks (GANs) to construct a deep learning framework for wave-mode decomposition.

The dispersion compensation method for Lamb wave mode separation is also extensively used in recent research due to its various advantages [64]. Xu *et al.* have used this method in their research. In that research, the dispersion compensation method has been explained for Lamb wave mode separation. This numerical method compensates, i.e., compresses, the individual dispersive waveforms

into temporal pulses. Thus it becomes nearly non-overlapped in frequency and time and frequency. So it could be extracted individually by rectangular time windows. The original waveform of each mode was then restored from the separated compensated pulse using an artificial dispersion technique, which was based on the reversibility of the numerical compensation method [65].

The two-dimensional fast Fourier transform (2D-FFT) has emerged as a viable approach for mode separation. Both time and space are equally measured in the 2D recorded data. The 2D-FFT is used to project the data obtained in the time distance domain onto the frequency wavenumber domain. In the frequency wavenumber domain, the Lamb wave modes are isolated and separated. The inverse 2D-FFT may also be used to reconstruct various wave modes. However, in order to detect them, a large number of sensors must be put at correctly spaced places, limiting the 2D-use FFT's in industrial applications [66]. In a research paper, the analysis of the S0/A0 elastic wave mode conversion phenomenon in glass fibre reinforced polymers has been presented [67], for that research work, the effect of the circular Teflon insert's position on S0/A0' mode conversion is explored. Teflon inserts were placed at various depths between layers of composite material. Moreover, impact damage with varying energies had an effect on the S0/A0' mode. The process of conversion was researched. Analysis of the investigated impact on the elastic wave theory is used to explain the discontinuities in S0/A0' mode conversion.

The time frequency analysis (TFA) approach is also considered as a potential tool for mode separation since it provides a good illustration of the time-frequency domain of the time-varying modal energy stream [43]. Several TFA approaches have been proposed for the signal processing of Lamb waves. On the other hand, small voltage excitations may reduce the Lamb signal's signal-to-noise ratio (SNR), resulting in a time frequency representation (TFR) with low time frequency resolution. Furthermore, a TFR with low time-frequency resolution may have a significant impact on the Lamb wave inspection's accuracy. The dispersion curve analysis method for Lamb wave mode separation is also a great approach. Hu *et al.*

described about this method in their research paper [66]. In that research, in order to solve the problem, a new signal processing method was proposed based on the dispersion curve analysis. To reduce the dispersion effect, a novel function based on the Hessian matrix was created to increase the energy concentration of the dispersion curve in the time frequency representation. Following that, the constrained penalty function algorithm was developed to detect dispersion curves. Finally, a mode reconstruction algorithm for recovering Lamb wave modes was developed. By enhancing the time frequency feature of the Lamb wave signal, the proposed method can separate overlapping wave modes and detect the crack fault. Two experiments were conducted to validate the efficacy of the proposed method for Lamb wave mode separation.

1.9 FBG and PZT sensors-based experiment for GW analysis

In this research work, the primary steps for performing mode separation based SHM have been presented. Here the signals have been successfully collected and plotted in Matlab. This will help to do a mode separation-based study by adding some advanced features. In the next section, preliminary results of elastic wave sensing using FBG and PZT are presented.

1.9.1 Experimental set-up

In order to perform experimental research related to elastic wave sensing using FBG, the experimental setup was designed. Fig. 1.2 shows the schematic sketch of the experimental setup and Fig. 1.3 shows the physical realization of the setup. As shown in Fig. 1.3, A Laser (Apex AP1000) has been used which can accept up to 8 plugs [68]. The FBGs manufactured by Femto Fibertech were used. The oscilloscope used here is of the National Instruments PXIe-5105 PXI, and an amplifier having good amplification features (Krohn Hite7500) was used [69].

The sample of interest was an Aluminum plate (100 cm x 100 cm

x 0.1 cm). Fig. 1.2 shows the connections between equipment. The connections are as follows:

1. Excitation signal is generated by wave generator and is amplified in voltage amplifier and drive PZT actuator (A). PZT generates elastic waves in the plate.
2. Two PZTs (PZT1, PZT2) working as elastic wave sensors were bonded to the plate and were connected with an oscilloscope card for elastic wave signal registering.
3. Two FBGs were connected to circulators, and then these circulators were connected to a tunable laser source and photodetectors. Photodetectors are connected also to an oscilloscope.

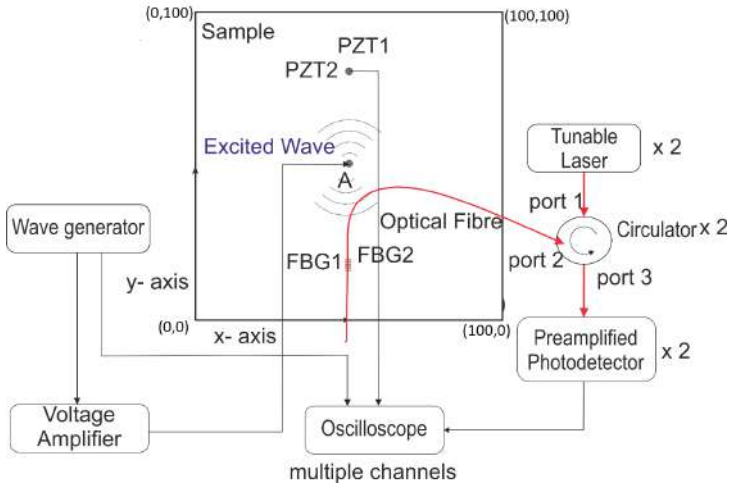


Fig. 1.2: Schematic Diagram for the complete connections.

1.9.2 Aim and the Methodology used

The aim of this research is related to the application of optical strain sensors based on Bragg gratings (FBG) for symmetric and antisymmetric elastic guided wave mode detection and mode separation for the purpose of damage detection.

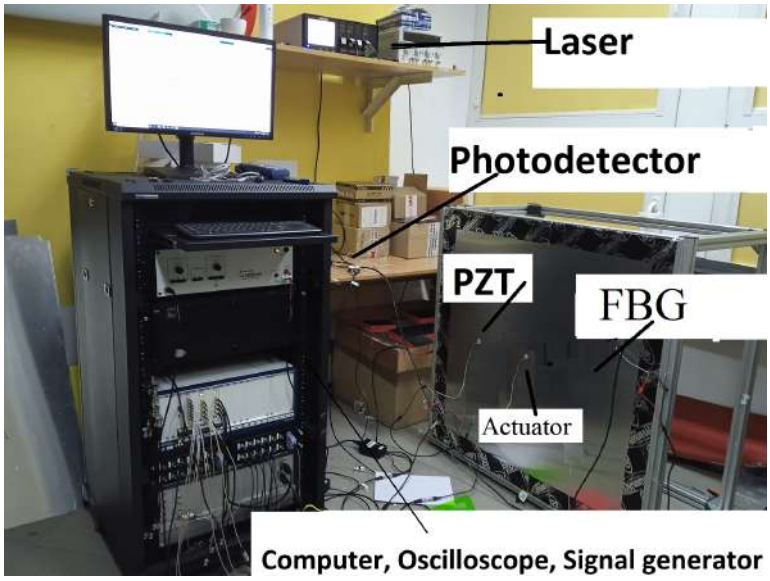


Fig. 1.3: Complete Setup.

The goal of this research is to investigate the sensitivity of FBG sensors to symmetric and antisymmetric elastic wave modes, as well as the feasibility of detecting the consequences (of these modes interacting with damage in the structure) (reflections, and especially mode conversions). The separation of symmetric and antisymmetric modes owing to the various location of FBG sensors will be the subject of research (top and bottom surface). The findings of wave sensing acquired by FBG sensors will be compared to those obtained by piezoelectric sensors, with the potential of damage detection taken into account. Connecting elastic wave sensing techniques based on piezoelectric and FBG optical sensors in such a way that the advantages of FBG sensors (directivity, lack of sensitivity to electromagnetic interferences) can be used to improve the effectiveness of damage detection methods will also be investigated. Electromagnetic interferences created during the elastic wave excitation by the piezoelectric actuator affect piezoelectric-based elastic wave sensors. Damage detection and localization were hampered as a result.

1.10 Preliminary results

In order to validate the developed measurement setup, preliminary measurements of elastic wave propagation were performed using PZTs and FBGs.

The first measurement was based on two FBGs, one at the top and the other at the bottom surface of the plate. Elastic wave excitation with frequency 110 kHz in the form of 5 cycles of sine signal modulated by Hann window was utilized. The Fig. 1.4 shows the waveform obtained in time vs amplitude (V) representation and the Fig. 1.5 shows the envelope on the waveform and peaks detection. In general, only primary peaks are used for SHM but here other peaks are also checked because these may be useful for some other applications like examining the reflection of waves.

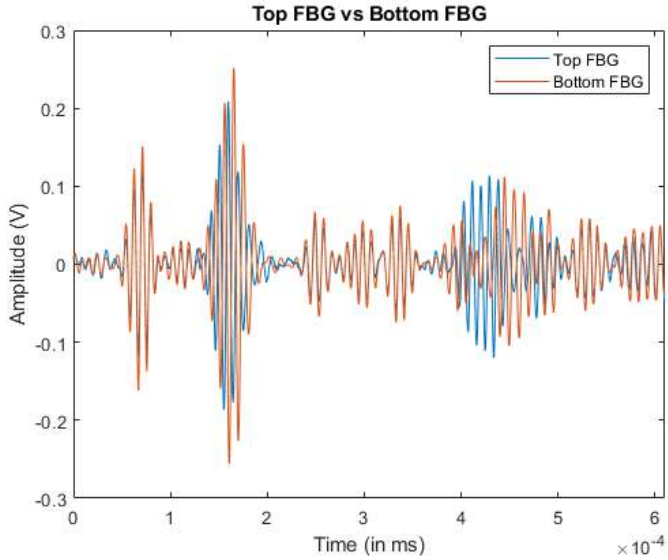


Fig. 1.4: Top FBG vs. Bottom FBG waveform.

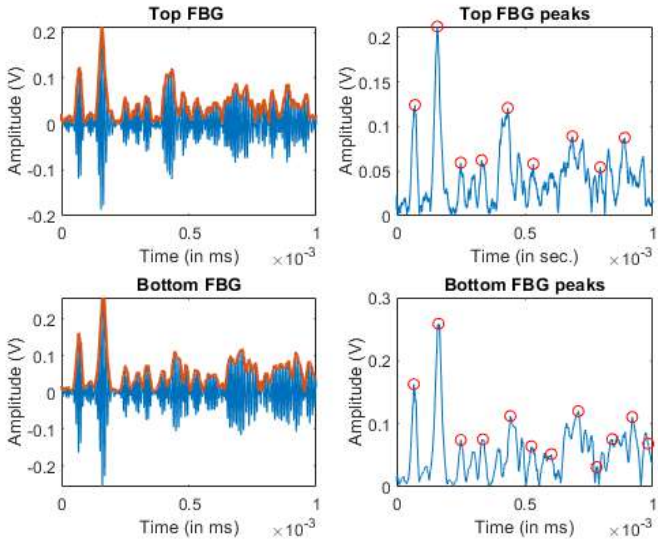


Fig. 1.5: Peak detection in the signal of both FBGs.

Due to the nature of symmetric and antisymmetric Lamb waves, the author expected that the symmetric mode signals for both FBGs will be in phase and for antisymmetric mode signals for both FBGs will be out of phase. This could be used for mode separation and is the main motivation for using 2 FBGs.

Next, the signals were gathered by two PZTs (PZT1, PZT2) which were attached to the top and bottom surface of the plate. The Fig. 1.6 shows the waveform obtained in time vs amplitude (V) representation and the Fig. 1.7 shows the envelope on the waveform and peaks detection.

In the next step, the comparison of the above two measurements has been presented by comparing PZT vs FBG. So the PZT (located at the top side) and FBG (located at the bottom side of the plate) were compared. The Fig. 1.8 shows the waveform obtained in time vs amplitude (V) representation.

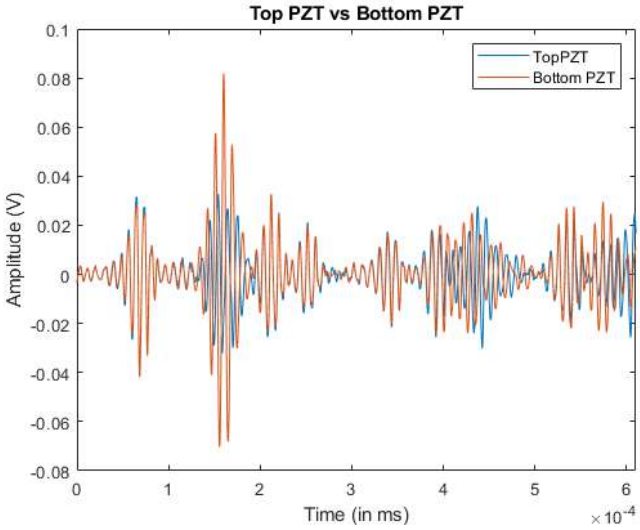


Fig. 1.6: Top PZT vs. Bottom PZT waveform.

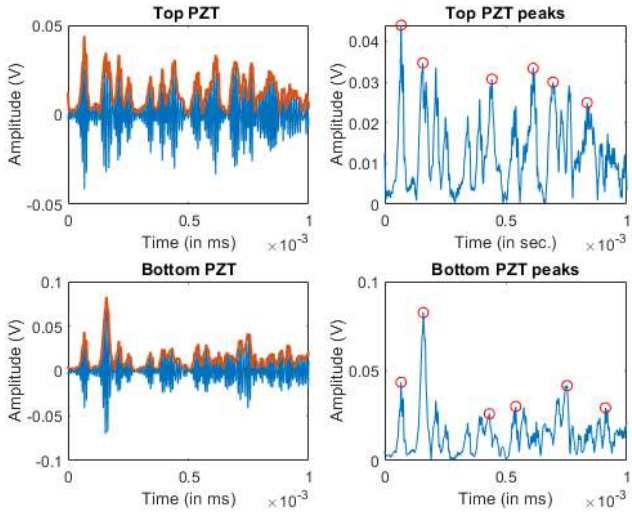


Fig. 1.7: Peak detection in the signal of both PZTs.

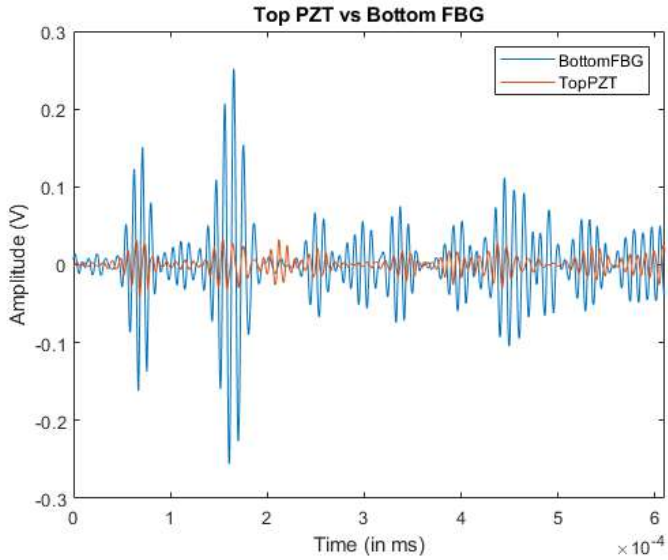


Fig. 1.8: Top FBG vs. Top PZT.

1.11 Discussion

A literature review was performed related to the techniques/ methods used for guided wave sensing and analysis for SHM. The mode separation was focused mainly due to its advantages in the research work of SHM. Major attention was focused on fibre optic Bragg grating strain sensors for elastic wave mode sensing. There is a lack of research related to the SHM based on the mode separation of the guided wave. The small number of works which were previously done by researchers were good but more and more research is needed on this topic to make the SHM process more advanced [70–72]. Due to this reason, this research has been performed and the aim of further research will be related to the problem of mode separation based on FBG sensors. Moreover, an experimental set-up for PZT and FBG-based elastic wave sensing was developed and physically realized. Preliminary measurements were performed using a developed setup. The methodology using 2 FBGs and 2 PZTs is found to be very useful, and the signals obtained are quite good

for symmetric and antisymmetric elastic guided wave mode detection and mode separation in symmetric and antisymmetric elastic guided wave modes. The methodology has been found very useful and reliable because the signals obtained are very good. Using the proposed methodology many difficulties/problems can be solved via the guided wave propagation approach in SHM. It is very feasible to identify and localize damage based on changes in wave propagation and processing. The methodology is very useful for further advanced research in the future.

Acknowledgements

The authors acknowledge the support provided by OPUS project: Study of elastic wave mode sensing and separation using FBG sensors for structural health monitoring, granted by the National Science Center, Poland under UMO-2020/39/B/ST8/01753.

Bibliography

[1] Md Khan, Nagafuchi Sunichi, and Md. Mehedi Hassan. Structural damage localization by linear technique of acoustic emission. *Open Journal of Fluid Dynamics*, 04:425–432, 01 2014.

[2] Mira Mitra and S Gopalakrishnan. Guided wave based structural health monitoring: A review. *Smart Materials and Structures*, 25(5):053001, mar 2016.

[3] Samiksha Sikarwar, Satyendra, Shakti Singh, and Bal Yadav. Review on pressure sensors for structural health monitoring. *Photonic Sensors*, 7:1–11, 08 2017.

[4] Hiroshi Tsuda. Ultrasound and damage detection in cfrp using fiber bragg grating sensors. *Composites Science and Technology*, 66(5):676–683, 2006. Reliability and Life Prediction of Composite Structures.

[5] Gayan Kahandawa Appuhamillage, Jayantha Epaarachchi, Hao Wang, and Kin tak Lau. Use of fbg sensors for shm in aerospace structures. *Photonic Sensors*, 2, 09 2012.

[6] Chuang Zhang and Leonard J. Bond. Performance evaluation of the fiber bragg grating (fbg) sensing device and comparison with piezoelectric sensors for ae detection. *AIP Conference Proceedings*, 1806(1):070007, 2017.

[7] Niusha Navabian, Sherif Beskhyroun, and Justin Matulich. Development of wireless smart sensor network for vibration-based structural health monitoring of civil structures. *Structure and Infrastructure Engineering*, 18(3):345–361, 2022.

[8] Vinoth Kanna Gopinath and Ravi Ramadoss. Review on structural health monitoring for restoration of heritage buildings. *Materials Today: Proceedings*, 43:1534–1538, 2021.

[9] G. Boscato, Luca Zanotti Fragonara, Antonella Cecchi, Emanuele Reccia, and Daniele Baraldi. Structural health monitoring through vibration-based approaches. *Shock and Vibration*, 2019:1–5, 02 2019.

[10] Claus-Peter Fritzen. Vibration-based structural health monitoring - concepts and applications. *Key Engineering Materials - KEY ENG MAT*, 293, 09 2005.

[11] D Goyal and BS Pabla. The vibration monitoring methods and signal processing techniques for structural health monitoring: a review. *Archives of Computational Methods in Engineering*, 23(4):585–594, 2016.

[12] Peng Ren and Zhi Zhou. Two-step approach to processing raw strain monitoring data for damage detection of structures under operational conditions. *Sensors*, 21:6887, 10 2021.

[13] Luigi Grassia, Michele Iannone, America Califano, and Alberto D'Amore. Strain based method for monitoring the health state of composite structures. *Composites Part B: Engineering*, 176:107253, 2019.

[14] Ajay Kesavan, Sabu John, and Israel Herszberg. Strain-based structural health monitoring of complex composite structures. *Structural Health Monitoring*, 7(3):203–213, 2008.

[15] S Saraygord Afshari, Seid H Pourtakdoust, BJ Crawford, R Seethaler, and AS Milani. Time-varying structural reliability assessment method: application to fiber reinforced composites under repeated impact loading. *Composite Structures*, 261:113287, 2021.

[16] Victor Giurgiutiu and Craig A. Rogers. Recent advancements in the electromechanical (E/M) impedance method for structural health monitoring and NDE. In Mark E. Regelbrugge,

editor, *Smart Structures and Materials 1998: Smart Structures and Integrated Systems*, volume 3329, pages 536 – 547. International Society for Optics and Photonics, SPIE, 1998.

[17] Mesut Tekkalmaz. Electro-mechanical impedance based shm of thick structures in broad-band frequency spectrum. *EskiĹźehir Technical University Journal of Science and Technology A - Applied Sciences and Engineering*, 06 2020.

[18] Tom Druet, Jean-Loup Tastet, Bastien Chapuis, and Emmanuel Moulin. Guided wave tomography for corrosion monitoring in planar structures. 2017.

[19] Andrii Kulakovskyi, Bastien Chapuis, Olivier Mesnil, Nas-Redine Bedreddine, Oscar D'almeida, and Alain Lhémery. Defect imaging on cfrp and honeycomb composite structures by guided waves generated and detected by a sparse pzt array. In *Structural Health Monitoring 2017*. DEStech Publications, Inc., 2017.

[20] Olivier Mesnil, Alexandre Imperiale, E. Demaldent, Vahan Baronian, and Bastien Chapuis. Simulation tools for guided wave based structural health monitoring. volume 1949, page 050001, 04 2018.

[21] RM Moroney, RM White, and RT Howe. Microtransport induced by ultrasonic lamb waves. *Applied physics letters*, 59(7):774–776, 1991.

[22] Henar Martin-Sanz, Konstantinos Tatsis, Domagoj Damjanovic, Irina Stipanovic, Aljosa Ĺ ajna, Ivan Duvnjak, Uros Bohinc, Eugen Bruhwiler, and Eleni Chatzi. Getting more out of existing structures: Steel bridge strengthening via uhpfrc. *Frontiers in Built Environment*, 5:26, 04 2019.

[23] Joyraj Chakraborty, Xin Wang, and Marek Stolinski. Damage detection in multiple rc structures based on embedded ultrasonic sensors and wavelet transform. *Buildings*, 11(2), 2021.

[24] Fabrizio Ricci, E. Monaco, N.D. Boffa, L. Maio, and V. Memmolo. Guided waves for structural health monitoring in composites: A review and implementation strategies. *Progress in Aerospace Sciences*, 129:100790, 2022. Impact induced dynamic response and failure behavior of aircraft structures.

[25] Hatem Elwalwal, Shahruddin Mahzan, and Ahmed Abdalla. Crack inspection using guided waves (gws)/structural health monitoring (shm): Review. *Journal of Applied Sciences*, 17:415–428, 07 2017.

[26] Rohan Soman, Kaleeswaran Balasubramaniam, Ali Golestani, MichaĹ, KarpiĹ„ski, and Pawel Malinowski. A two-step guided waves based damage localization technique using optical fiber sensors. *Sensors*, 20(20), 2020.

[27] Rohan Soman. Multi-objective optimization for joint actuator and sensor placement for guided waves based structural health monitoring using fibre bragg grating sensors. *Ultrasonics*, 119:106605, 02 2022.

[28] T.H.T. Chan, L. Yu, H.Y. Tam, Y.Q. Ni, S.Y. Liu, W.H. Chung, and L.K. Cheng. Fiber bragg grating sensors for structural health monitoring of tsing ma bridge: Background and experimental observation. *Engineering Structures*, 28(5):648–659, 2006.

[29] Yue Hu, Yanping Zhu, Xiaotong Tu, Jing Lu, and Fucai Li. Dispersion curve analysis method for lamb wave mode separation. *Structural Health Monitoring*, 19(5):1590–1601, 2020.

[30] Darryll Pines and Liming Salvino. Structural health monitoring using empirical mode decomposition and the hilbert phase. *Journal of Sound and Vibration*, 294(1):97–124, 2006.

[31] Muntazir Abbas and Mahmood Shafiee. Structural health monitoring (shm) and determination of surface defects in large metallic structures using ultrasonic guided waves. *Sensors*, 18(11):3958, 2018.

[32] Peter Cawley, Frederic Cegla, and Andrea Galvagni. Guided waves for ndt and permanently-installed monitoring. *Insight-Non-Destructive Testing and Condition Monitoring*, 54(11):594–601, 2012.

[33] Lorenzo Capineri and Andrea Bulletti. Ultrasonic guided-waves sensors and integrated structural health monitoring systems for impact detection and localization: A review. *Sensors*, 21(9):2929, 2021.

[34] Kilian Tschöke, Inka Mueller, Vittorio Memmolo, Maria Moix-Bonet, Jochen Moll, Yevgeniya Lugovtsova, Mikhail Golub, Ramanan Sridaran Venkat, and Lars Schubert. Feasibility of model-assisted probability of detection principles for structural health monitoring systems based on guided waves for fiber-reinforced composites. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 68(10):3156–3173, 2021.

[35] Xiangtao Sun, Chuanrui Guo, Lei Yuan, Qingzhao Kong, and Yiqing Ni. Diffuse ultrasonic wave-based damage detection of railway tracks using pzt/fbg hybrid sensing system. *Sensors*, 22(7):2504, 2022.

[36] Sensors for structural health monitoring - fprimec solutions inc. https://www.fprimec.com/sensors-for-structural-health-monitoring. (Accessed on 05/18/2022).

[37] Arnaldo Leal-Junior, Camilo DĂaz, Anselmo Frizera, Heeyoung Lee, Kentaro Nakamura, Yosuke Mizuno, and Carlos Marques. Highly sensitive fiber-optic intrinsic electromagnetic field sensing. *Advanced Photonics Research*, 2(1):2000078, 2021.

[38] Tiange Wu, Guowei Liu, Shenggui Fu, and Fei Xing. Recent progress of fiber-optic sensors for the structural health monitoring of civil infrastructure. *Sensors*, 20(16):4517, 2020.

[39] José Miguel López-Higuera, Luis Rodriguez Cobo, Antonio Quintela Incera, and Adolfo Cobo. Fiber optic sensors in

structural health monitoring. *Journal of lightwave technology*, 29(4):587–608, 2011.

[40] Alfredo Güemes, Antonio Fernández-López, Patricia F Díaz-Maroto, Angel Lozano, and Julian Sierra-Perez. Structural health monitoring in composite structures by fiber-optic sensors. *Sensors*, 18(4):1094, 2018.

[41] L Li, XL Tong, CM Zhou, HQ Wen, DJ Lv, K Ling, and CS Wen. Integration of miniature fabry–perot fiber optic sensor with fbg for the measurement of temperature and strain. *Optics Communications*, 284(6):1612–1615, 2011.

[42] Xiaotian Zou, Alice Chao, Ye Tian, Nan Wu, Hongtao Zhang, Tzu-Yang Yu, and Xingwei Wang. An experimental study on the concrete hydration process using fabry–perot fiber optic temperature sensors. *Measurement*, 45(5):1077–1082, 2012.

[43] What is a fiber bragg grating? — fbg — sensors — hbm. https://www.hbm.com/en/4596/what-is-a-fiber-bragg-grating. (Accessed on 05/18/2022).

[44] Tianliang Li, Jinxiu Guo, Yuegang Tan, and Zude Zhou. Recent advances and tendency in fiber bragg grating-based vibration sensor: A review. *IEEE Sensors Journal*, 20(20):12074–12087, 2020.

[45] Fengming Yu, Osamu Saito, and Yoji Okabe. Laser ultrasonic visualization technique using a fiber-optic bragg grating ultrasonic sensor with an improved adhesion configuration. *Structural Health Monitoring*, 20(1):303–320, 2021.

[46] Wuke Xu, Qi Wu, Hanqi Zhang, Chen Gong, Rong Wang, Jiyun Lu, and Ke Xiong. Debonding monitoring of cfrp t-joint using optical acoustic emission sensor. *Composite Structures*, 273:114266, 2021.

[47] Andrew Navratil, Junghyun Wee, and Kara Peters. Ultrasonic frequency response of fiber bragg grating under direct and remote adhesive bonding configurations. *Measurement Science and Technology*, 33(1):015204, 2021.

[48] Florian Lambinet and Zahra Sharif Khodaei. Development of hybrid piezoelectric-fibre optic composite patch repair solutions. *Sensors*, 21(15):5131, 2021.

[49] Qi Wu, Yoji Okabe, and Fengming Yu. Ultrasonic structural health monitoring using fiber bragg grating. *Sensors*, 18(10):3395, 2018.

[50] Rohan Soman, Junghyun Wee, and Kara Peters. Optical fiber sensors for ultrasonic structural health monitoring: a review. *Sensors*, 21(21):7345, 2021.

[51] Darryll Pines and Liming Salvino. Structural health monitoring using empirical mode decomposition and the hilbert phase. *Journal of sound and vibration*, 294(1-2):97–124, 2006.

[52] Michel Lemistre and Daniel Balageas. Structural health monitoring system based on diffracted lamb wave analysis by multiresolution processing. *Smart materials and structures*, 10(3):504, 2001.

[53] C Willberg, S Koch, G Mook, J Pohl, and U Gabbert. Continuous mode conversion of lamb waves in cfrp plates. *Smart Materials and Structures*, 21(7):075022, 2012.

[54] T Wandowski, P Kudela, and WM Ostachowicz. Numerical analysis of elastic wave mode conversion on discontinuities. *Composite Structures*, 215:317–330, 2019.

[55] Yu Wang. *Damage detection in reinforced concrete and self-healing concrete structures using non-destructive testing techniques*. PhD thesis, Cardiff University, 2020.

[56] Lingyu Yu and Zhenhua Tian. Lamb wave structural health monitoring using a hybrid pzt-laser vibrometer approach. *Structural Health Monitoring*, 12(5-6):469–483, 2013.

[57] Lina Draudvilienė, Olgirdas Tumšys, and Renaldas Raišutis. Reconstruction of lamb wave dispersion curves in different objects using signals measured at two different distances. *Materials*, 14(22):6990, 2021.

[58] Chung-Yub Kim and Kyung-Jo Park. Mode separation and characterization of torsional guided wave signals reflected from defects using chirplet transform. *NDT & E International*, 74:15–23, 2015.

[59] Javad Rostami, Jingming Chen, and Peter W Tse. A signal processing approach with a smooth empirical mode decomposition to reveal hidden trace of corrosion in highly contaminated guided wave signals for concrete-covered pipes. *Sensors*, 17(2):302, 2017.

[60] Mohammad Ali Fakih, Samir Mustapha, Jaafar Tarraf, Georges Ayoub, and Ramsey Hamade. Detection and assessment of flaws in friction stir welded joints using ultrasonic guided waves: experimental and finite element analysis. *Mechanical Systems and Signal Processing*, 101:516–534, 2018.

[61] Kai Zhou, Yuebin Zheng, Jiaqi Zhang, Xinsheng Xu, Shuyi Ma, and Zhanjun Wu. A reconstruction-based mode separation method of lamb wave for damage detection in plate structures. *Smart Materials and Structures*, 28(3):035033, 2019.

[62] Jiubing Cheng and Sergey Fomel. Fast algorithms for elastic-wave-mode separation and vector decomposition using low-rank approximation for anisotropic media. *Geophysics*, 79(4):C97–C110, 2014.

[63] Harpreet Kaur, Sergey Fomel, and Nam Pham. A fast algorithm for elastic wave-mode separation using deep learning with generative adversarial networks (gans). *Journal of Geophysical Research: Solid Earth*, 126(9):e2020JB021123, 2021.

[64] Cai-bin Xu, Zhi-bo Yang, Xue-feng Chen, Shao-hua Tian, and Yong Xie. A guided wave dispersion compensation method based on compressed sensing. *Mechanical Systems and Signal Processing*, 103:89–104, 2018.

[65] Kailiang Xu, Dean Ta, Petro Moilanen, and Weiqi Wang. Mode separation of lamb waves based on dispersion compensation

method. *The Journal of the Acoustical Society of America*, 131(4):2714–2722, 2012.

[66] Yue Hu, Yanping Zhu, Xiaotong Tu, Jing Lu, and Fucai Li. Dispersion curve analysis method for lamb wave mode separation. *Structural Health Monitoring*, 19(5):1590–1601, 2020.

[67] Tomasz Wandowski, Pawel Malinowski, and Wieslaw Ostachowicz. Analysis of s0/a0 elastic wave mode conversion phenomenon in glass fibre reinforced polymers. In *MATEC Web of Conferences*, volume 188, page 01009. EDP Sciences, 2018.

[68] Optical multitest platform ap1000-8 — apex technologies. https://apexmonitoring.com/optical-multitest-platform-ap1000-8. (Accessed on 05/18/2022).

[69] Krohn-hite corporation home page. https://www.krohn-hite.com. (Accessed on 05/18/2022).

[70] Hoon Sohn, Charles R Farrar, Francois M Hemez, and Jerry J Czarnecki. A review of structural health review of structural health monitoring literature 1996-2001. 2002.

[71] Ambarish G Mohapatra, Jaideep Talukdar, Tarini Ch Mishra, Sameer Anand, Ajay Jaiswal, Ashish Khanna, and Deepak Gupta. Fiber bragg grating sensors driven structural health monitoring by using multimedia-enabled iot and big data technology. *Multimedia Tools and Applications*, pages 1–21, 2022.

[72] L Mallet, BC Lee, WJ Staszewski, and F Scarpa. Structural health monitoring using scanning laser vibrometry: Ii. lamb waves for damage detection. *Smart Materials and Structures*, 13(2):261, 2004.

Chapter 2

Fire resistance test methods of battery energy storage systems (BESS): analysis, comparison and prospects

Daniel Darnikowski

Institute of Fluid Flow Machinery, Polish Academy of Sciences,
Fiszera 14, 80-231 Gdansk, Poland
Maritime Advanced Research Centre, Szczecińska 65, 80-392 Gdańsk, Poland

2.1 Introduction

With the wake of climate change and the efforts to reduce carbon imprint in mobility and power train, electric and hybrid vehicles sales and popularity is on the rise [1]. As a clean alternative to the internal combustion vehicles, electric vehicles (EVs) promise more eco friendly future and excellent technical performance as well as relatively reduced complexity [2]. Despite pandemic and global shortage crisis stifling car sales all around the world, EV sales prove to be resilient with increasing car stock each year [3].

Simultaneously, fire incidents involving electric car batteries are being closely scrutinised in the press and studies [4–6]. The dramatic effect of electric vehicle battery fire and inability to efficiently douse it raises great concern for the public and overall fire safety [7]. The battery fire is subjected to the phenomenon called "thermal runaway" which is defined as "a process that is accelerated by increased temperature, in turn releasing energy that further increases temperature" [8]. In terms of lithium-ion battery (LIB), one or multiple cells start releasing internal energy in form of heat, which is then transferred to another cells and modules causing a cascading effect [9]. The precise mechanism is discussed in the studies and presented in the Section 2.2 in the following article. It is also evident to assume with the increasing number of electric vehicles on roads the number of fire incidents will increase as well [10].

The literature shows LIB fire incidents pose significant danger to the surrounding, especially when encountered with internal combustion vehicles and their fuel [11]. The battery fire due to its self-sustaining effect and propagation do not responds to the conventional dousing methods, which base their mechanism on the elimination of oxidising agent – predominantly oxygen [7]. LIB burning mechanism and its characteristics is discussed further in Section 2.2. Current extinguishing methods involve state-of-art foaming agents, submerging whole vehicle in the water tanks or hours of water pouring, all with the intention to bring the battery below the critical temperature [12]. This poses enormous challenge to the fire fighting brigades, particularly if the vehicle is parked in the restricted space such as underground car parks [7, 13].

Another concern in terms of LIB fires are the hazardous gases released during the thermal runaway event. The released gases depend largely on the type of cell used in the battery and its state of charge (SOC), however, among a few repeating toxic compounds is hydrogen fluoride, carbon monoxide, hydrogen, benzine and other cancerous aromatic compounds [14]. Depending on the surrounding and volume of the room, the lethal concentration threshold for human can be surpassed multiple times for LIB fire. Similarly to the fire fighting concern, the toxicity of the gases is one of the most problematic topics for the confined spaces such as underground car parks [15].

Multiple high-quality papers of LIB failure and its mechanisms have been recently published, covering such topics as cathode and anode materials, electrolyte, heat and thermal runaway management. Following article presents on the LIB failure scenarios and relates it to the current testing procedures of battery energy storage systems (BESS). A comparative study is performed between regulatory and standardisation documents, and in addition an alternative procedure is proposed.

2.2 LIB abuse scenarios

Battery during its functioning produces heat, which cannot be removed even with normal operating conditions [16]. Temperature rise within battery may trigger other undesirable reactions, effectively causing thermal runaway event, where the heat generation cannot be reversed [17]. The literature consensus differentiates three types of abuse - depending on their type of strain [5, 8, 18]:

- Mechanical abuse – e.g. damage to shell casing, punching, strain, compression, twisting, penetration.
- Electrical abuse – e.g. over- or discharge and short-circuit.
- Thermal abuse – e.g. thermal shock, heating conditions, local heating.

Understanding the hazardous events that may lead to thermal runaway event is therefore crucial to prevent it from happening. An over-

view of the battery safety concerns has been presented in Fig. 2.1.

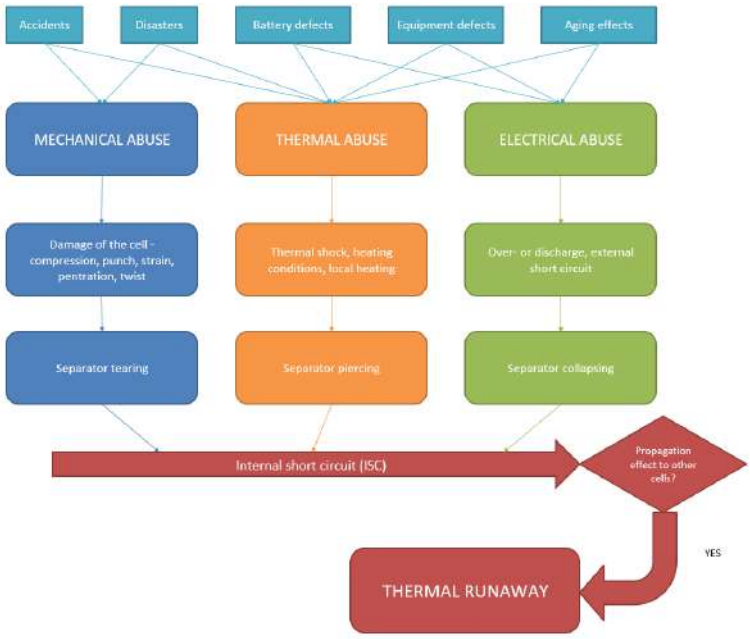


Fig. 2.1: Battery safety issues. First row represents potential causes for LIB abuse. Second row divides given incident due to type of the abuse. Third row presents the examples of the processes causing given abuse to which LIBs are subjected. Fourth row presents damage to separator, which then results in internal short circuit.

2.2.1 Mechanical abuse

With the increasing number of EVs on the roads, mechanical abuse of the LIB is expected to grow together with the number of traffic accidents and collisions involving LIB. The damage made on the impact to structure of a car may result in infringement of battery integrity and following thermal runaway. Any potential damage to shell of the cell in form of punching, strain, compression and twisting may result in internal short circuit (ISC) due to impalement of the separator [19].

Cell is built from the shell (external plastic, aluminium or steel film) and a roll of cathode, anode and separator in between. Due to external force, the weakest component of the cell is prone to damage, which may produce thermal event. Although the number of incidents including EVs is on the rise [6], there is little publicly available details on the proceedings of such events. Current research is based on theoretical and numerical studies performed on the 3D models of the batteries and such are largely regarded in the following study.

The crush model of the at the cell level has been successfully studied in the literature. Quasi-static tests has been performed, which included crush, punching, bending etc. of the cell. The cell has been built from the jelly roll and can be represented during quasi-static breaking tests [20–22]. Cell model can be applied in the collision model at module/pack level to predict behaviour of the vehicle during the traffic crash [23]. Modelling tool therefore can be successfully utilized to perform pre-homologation analysis on the specimen to ascertain its performance. In addition, these tools can be utilised to perform an analysis of the influence of potential changes (e.g. different casing shape or material) that manufacturer may be willing to make in the future, hence limiting the testing requirement.

2.2.2 Electrical abuse

The purpose of the battery is to convert chemical energy into electrical. The lithium-ion battery is of rechargeable type meaning the chemical reactions can be reversed while plugged into the power. Sometimes due to the battery degradation or malfunction of the charging equipment, the battery is exposed to overcharge.

There are multiple reasons on why the battery is overcharging. Firstly, not all the cells are the same, therefore there will be small discrepancies from one cell to another. The charging is also controlled by the ”battery management system” (BMS), which ensures the open-circuit voltage of each controlled cell is equal. Any malfunction of BMS may result in excess energy stored into the battery, surpassing their designed state of charge (SOC).

Overcharging cause electrolyte decomposition at the cathode interface and then this reaction slowly is increasing cell temperature. The cathode becomes unstable while releasing oxygen and Li^+ deposits itself on the anode in form of of Li dendrites [24]. The process continues until heat and gas causes cell to rupture and vent [25]. On the outside, the cell begins to show swelling at 120% of the SOC, while rupture and venting usually happens at 160% [5]. Past that point the cell enters thermal runaway and the reaction is irreversible.

Another point of concern in terms of electrical abuse is over-discharge of the battery. The mechanism of the over-discharge is quite similar to the overcharge. The formation of Li^+ from continuously over-discharged changes graphite structure and may lead to destruction of the solid electrolyte interphase (SEI) [26].

In addition to the above abuses, the LIB can be subjected to external short circuit. It occurs when a cathode and an anode of the same cell are in direct contact through a conductor. During such event, Li^+ transfers rapidly discharging the cell [27].

Overcharge is one of the most dangerous types of the abuse and one of the most frequently observed reasons for LIB safety accidents. It is arguably also one of the most dangerous ones due to high energy state of the battery at the time of the event [28].

2.2.3 Thermal abuse

In the event of thermal abuse, the battery experiences an exposition to either high or elevated temperature due to own self-heating or external source [29]. The battery fire may happen during charging or while in-use. A malfunctioning battery may ignite highly combustible airborne materials or flammable components of the vehicle.

High-temperature concentrated into one point of the battery is typically connected to the poor design and lack of heat dissipation [30]. Despite the fact, the battery cycling should not cause safety accidents, since the heat generation from anodic and cathodic reactions are not sufficient to cause sharp temperature increase. Instead, heat dissipation largely depends on the battery geometrics and design,

which can accumulate heat in certain points. The literature shows that the unfavourable reactions may start at 80°C [8, 31].

The exposure to external flame may happen during the accident or disaster events. During car crash an ignited pool of fire from petrol pool may cause an external heating condition to the LIB battery. During such events, a battery is expected to be exposed to temperatures between 300 and 900°C. Another external heating source are the nearby fires, however, the vast range of possible fire scenarios do not allow for the standardisation of test method.

2.3 Fire testing procedures

Following paragraph focuses on the current documents regulating or standardising fire resistance requirements of battery energy storage systems (BESS) and rechargeable energy storage systems (RESS). The tests in the paragraph are focused within the thermal abuse of device under test (DUT). DUT is exposed to the thermal shock in form of direct or indirect high temperature source for a certain amount of time and then its behaviour is evaluated, and safety risks are noted. The fire resistance criteria are provided to determine the outcome of the test. For instance GTR 20 Fire resistance test purpose is to allow for safe evacuation of the vehicle occupants [28].

While the standards involve multiple objectives to obtain for the battery in the fields of design, performance, safety design, environmental protection, classification etc., the standards set the baseline for RESS safety [32]. When products receive positive test results and are introduced to the market, the actions on the safety design become reactive after accidents or failures happen [33]. Thus, it is of utmost importance to develop high quality testing procedures to guarantee an excellent safety threshold for the users.

The standards and regulations selected for this study are: GTR 20:2018 [28], UNECE R100.02:2013 [34], UNECE R100.03:2021 [35], ISO 12405-3:2014 [36], ISO 6469-1:2019 [37], ISO 18243:2017 [38], SAE J2464:2009 [39], GB/T 31467.3-2015 [40], SAND2017-6925:2017 [41] and UL 2580:2020 [42].

Despite the fact that there multiple standards involved the procedures are often times duplicated with small changes. A following section (2.3.1) will present the procedures used broadly nowadays.

2.3.1 Open flame procedure

There are similar procedures in following standards and documents: GTR 20:2018 [28], ISO 12405-3:2014 [36], ISO 6469-1:2019 [37], ISO 18243:2017 [38], UNECE R100.02:2013 [34], UNECE R100.03:2021 [35] and GB/T 31467.3-2015 [40]. The exposure is induced by the positive-ignition pool fire contained in the pan of dimensions wider and longer by 200-500mm than the DUT. Provided DUT is not a part of a vehicle, it is set at a height of 50 cm above the surface of the fuel. The test phases are presented in Fig. 2.2 and described below:

- **Phase A: Pre-heating** – 60 seconds pre-heating with the DUT being at least 3 m away. After 60 seconds, DUT shall be exposed to fire.
- **Phase B: Direct exposure to flame** – 70 seconds of direct exposure of the DUT to flames from freely burning fuel.
- **Phase C: Indirect exposure to flame** – continuation of the DUT exposure for another 60 seconds through perforated screen made of refractory material.
- **Phase D: End of the test or post-test observation** – DUT is removed at least 3 m from the burning pan for observation. The fire is extinguished either immediately after the DUT moved away or within couple minutes. The DUT surface temperature shall reach ambient or should be decreasing for the minimum of 3 h until the end of test.

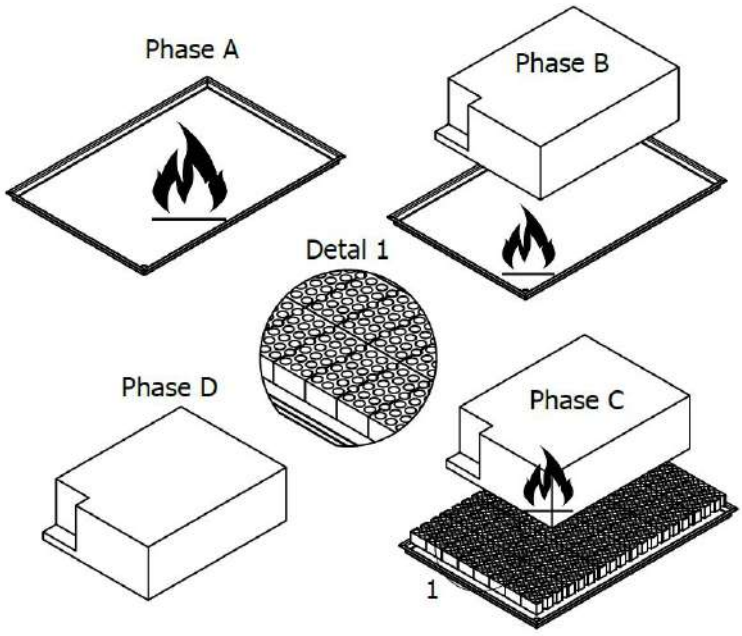


Fig. 2.2: Scheme of fire resistance procedure - phases A-D. Detail 1 shows perforated screen used in Phase C [31].

During open flame test, there are following prerequisites to be met in order to carry out the test.

- Ambient temperature must be above 0°C, wind speed above the pan must be less than 2.5 $\frac{km}{h}$ (0.7 $\frac{m}{s}$).
- State of charge must be adjusted to 100% (only UNECE R100.02 requires above 50%)
- The quantity of fuel must be sufficient to permit the flame, under free-burning conditions to burn for the whole test duration. Only ISO 18243:2017 defines specific value to 25 l/m^2.

2.3.2 LPG burner procedure

As per the newest revision of UNECE R100.03:2021 [35] and GTR 20:2018 [28], an alternative method of fire resistance testing

was introduced using LPG burners instead of petrol pool burning. The procedure then consists of three stages:

- **Heating:** average temperature needs to reach 800°C within 30 sec.
- **Exposure:** average temperature after reaching 800°C shall be maintained for 120 sec within 800-1100°C.
- **Observation:** DUT shall be observed until surface temperature reaches ambient or has been decreasing for a minimum of 3h.

The procedure additionally requires to monitor temperature 50 mm below the DUT with the minium quantity of five sensors. Such sensors are used to establish exposure conditions within 800°C - 1100°C range. Testing rig constructed by Korea Apparel Testing and Research Institute (KATRI) is shown in GTR 20:2018 [28]. It is also required for the burner flame to be 600 mm high measuring without the tested object.

2.3.3 Radiative source procedure

SAND2017-6925:2017 [41] and SAE J2464:2009 [39] documents contain fire resistant procedure based on radiative heat source (steel plate). The procedure itself is largely different to the previously presented.

DUT is exposed to radiating surface of of 890°C, which is required to be established within 90 seconds of the start of the test. Radiating surface (steel plate) must be heated unilaterally by the selected heat source with the exhaust gases not coinciding with DUT fumes (separate chamber). The inside chamber with the DUT possess separate venting system to allow to collect exhaust gases. Simplified view of the testing set up has been shown in Fig. 2.3. The DUT is exposed to radiation for 10 min. or longer.

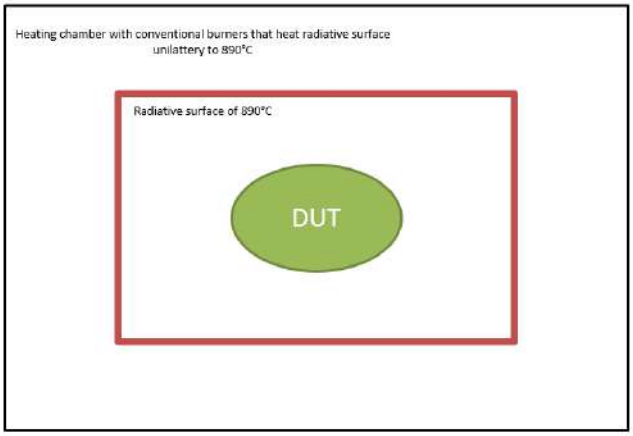


Fig. 2.3: SAND2017-6925:2017 [41] and SAE J2464:2009 [39] testing set up.

2.3.4 Testing procedure analysis and comparison

Above-mentioned procedures are summarised in the Tab. 2.1. The comparison is made in terms of heat source, main heat transfer, temperature range, tolerance and deviation, heat flux range and deviation, test length and amount of heat.

Main heat transfer is different for LPG burner method, where a convection is responsible for heat transfer to the DUT [43]. Due to large LPG flow, higher pressure and normalised burner design the flame is not prone to turbulence [44]. For open flame method, the fire usually emits mainly radiation heat due to its scale being larger than 0.2 m in diameter [45]. The flame itself however is prone to self-caused turbulence and changes in wind direction even being the 0.7 m/s required threshold [31, 46]. In addition, the large portion of the DUT is inside the flame. Radiative source procedures as the name suggest use radiation from hot plate. This form represents the most stable heating conditions, since radiative heat transfer is not prone to local turbulence [47].

Tab. 2.1: Comparison table on the RESS and BESS fire testing procedures [28, 34–39, 41].

	Open flame procedure
Documents	GTR 20:2018, ISO 12405-3:2014, ISO 6469-1:2019, ISO 18243:2017, UNECE R100.02:2013, UNECE R100.03:202
Source of heat	Petrol
Main heat transfer	Radiative/Convective
Temperature range	250-900°C
Temperature tolerance	None
Heat flux range	25-50 kW/m^2
Test length	130 seconds
Heat amount range	3.25-6.5 MJ/m^2
Temperature deviation	260%
Heat flux deviation	100%
	LPG burner procedure
Documents	UNECE R100.03:2021, GTR 20:2018
Source of heat	LPG burner
Main heat transfer	Convective
Temperature range	800-1100°C
Temperature tolerance	950±150°C
Heat flux range	30-40 kW/m^2
Test length	120 seconds
Heat amount range	3.6-4.8 MJ/m^2
Temperature deviation	38%
Heat flux deviation	33%
	Radiative source procedure
Documents	SAND2017-6925:2017, SAE J2464:2009
Source of heat	Hot plate
Main heat transfer	Radiative
Temperature range	890°C
Temperature tolerance	5% (890±44.5°C)
Heat flux range	56-77 kW/m^2
Test length	600 seconds
Heat amount range	33.6-46.2 MJ/m^2
Temperature deviation	5%
Heat flux deviation	38%

Temperature range for radiative heat transfer is fixed at 890°C with tolerance of 5%. This constitutes to the most stable conditions in respect to temperature range and deviation. The LPG burner method requires temperature range of 800°C - 1100°C (38% deviation allowed), which is still relatively large when referred with short exposure time. The open flame method does not specify allowed temperature range, however, the literature shows the range of 250°C - 900°C [31]. In addition, the open flame procedure does not require any tolerance, therefore a deviation up to 260% in temperature can be permitted.

Depending on the testing conditions, the open flame method also shows the largest variability in terms of heat flux range between 25 and 50 kW/m^2 that constitutes up to 100% deviation [4, 28]. LPG burners have significantly less fluctuations at the range of 30-40 kW/m^2 (33% deviation) [28]. The radiative source procedure despite the fact that it allows for miniscule temperature tolerance, it contributes to notable range in heat flux of 56-77 kW/m^2 (38% deviation). Heat flux of the hot plate was calculated on the basis of net radition loss rate and the following equation:

$$\dot{Q} = \epsilon \times \sigma \times (T^4_{h_{min-max}} - T^4_{c_{min-max}}) \times A_h$$

where $\dot{Q}$ is heat flux, ϵ is an emissivity coefficient of the object (estimated for steel at 0.64), σ is Stefan-Boltzmann constant, $T^4_{h_{min-max}}$ is a hot body absolute minimum or maximum temperature, $T^4_{c_{min-max}}$ is a cold surroundings absolute temperature, A_h is an area of an object (standardised for 1 m^2).

In terms of total amount of heat, the large difference between open flame and LPG burner procedures and radiative source procedure is due to length of the test. The amount of heat is comparable between LPG burner and open flame procedures as per the literature data.

In the wake of the comparison of fire testing procedures, a major deviations can be found within each method in term of heat flux. By far the open flame procedure proved to be the least reliant in terms of predicted repeatability of the exposure conditions.

2.4 Alternative procedure

Due to variability found in the current test methods of BESS fire resistance, an alternative solution was studied. EN 1363-1 [48] is a fire resistance standard of passive elements, that uses large scale furnace to conduct tests on elements such as doorsets, windows, walls and other. It uses a set of burners (usually fuelled LPG or LNG) located inside the chamber to produce standardised heating curve. The temperature inside the chamber is strictly controlled by the set of thermocouples located at different heights. The test scheme was presented in Fig. 2.4.

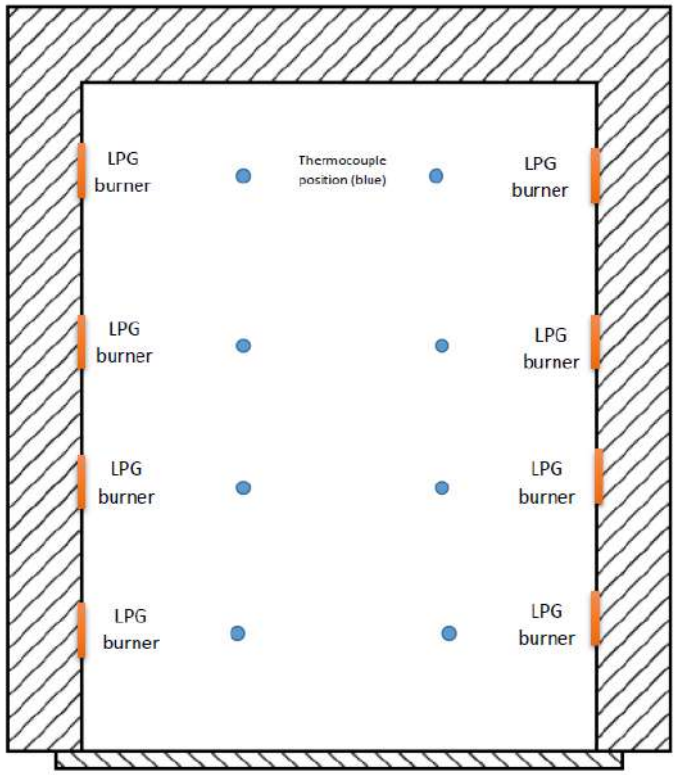


Fig. 2.4: EN 1363-1 testing set up.

Inside such furnace a possibility of testing large scale BESS or RESS is highly anticipated. The control over temperature inside the testing chamber and the event temperature distribution create promising prospects to reduce variability of the above mentioned methods. In order to check the repeatability a series of two tests were performed. A comparison of temperature curves and deviation was made.

The tests were carried out in a chamber 3 m wide and high and 2 m deep with four LPG burners on the each side. Two sets of four thermocouples on each side was placed in the furnace. Thermocouples were distanced 100 cm from the chamber wall and spaced at 60 cm from each other at different heights.

The standardised temperature-time curve was selected from ISO 834-1 as following equation [49]:

$T = T_0 + 345 \times \log_{10}(8 \times t + 1)$

where T is temperature inside the furnace (°C), T_0 is the initial furnace temperature (°C), t is time (minutes). Furnace thermocouples were enclosed by the INCONEL 600 plates of 0.7 mm thickness and insulation material, as prescribed in EN 1363-1 [50]. The enclosure of the thermocouple allows to exclude the influence of turbulence and free-burning of the sample on the measuring value of the thermocouple, while simultaneously to provide reference for burner control system. The measuring tip of the thermocouple was inserted in between the plate and insulation material, thus there are no significant losses expected in temperature rise rate (i.e. $<0.1°C/s$).

Temperature results of the tests are shown in Fig. 2.5 and deviation from the standard curve is shown in Fig. 2.6. Test one was referred as "T1" and test two as "T2".

Temperature graph shows significant deviation from the standard temperature in the initial two minutes of the test. This may be related to the large scale of the chamber. In addition measuring tips of the thermocouples were covered by the insulation material to reduce influence of the turbulence on LPG burner control system. This, however, added additional inertia to the thermocouple readouts. Past two minutes, the temperatures in both tests

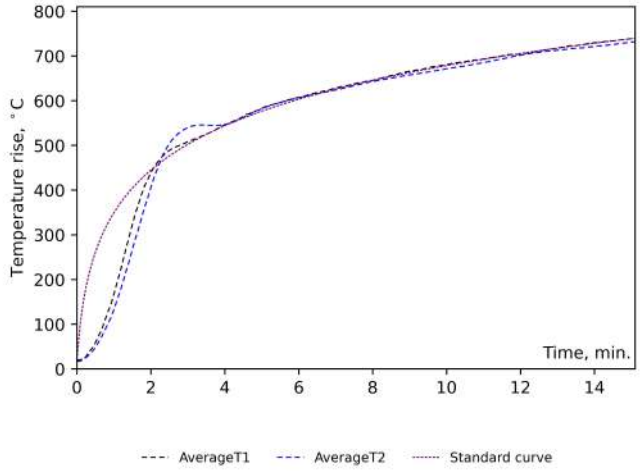


Fig. 2.5: Temperature time heating curve from two tests – T1 and T2.

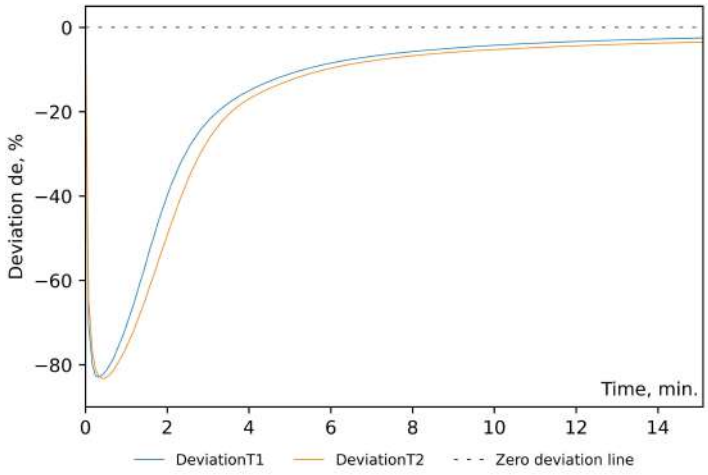


Fig. 2.6: Deviation from the standard heating curve from two tests – T1 and T2.

were following closely the designed standard curve, until the end at 15 minutes. Additionally, in both tests the curves had almost identical course, which confirm the repeatability of the method.

Deviation graph corresponds to the temperature results – we can observe large deviation at the start of the test due to the initial lower recorded temperatures than the standard curve. With the progression of the tests, the deviation diminished to less than 2%.

The summary of differences between the test is presented in the Tab. 2.2. The highest recorded temperature difference was 56.5°C, while the maximum deviation difference was at 9.50%. Such results show the repeatability and accuracy of the results may be significantly higher than the methods discussed in Section 2.3.

Tab. 2.2: Tests comparison – maximum and minimum difference between T1 and T2 in terms of temperature and deviation.

Average Temperature Maximum Difference	56.5°C
Average Temperature Minimum Difference	0.1°C
Deviation Maximum Difference	9.50%
Deviation Minimum Difference	0.12%

In the following studies, the temperature-time curve will be adjusted to obtain optimal temperature rise at the beginning of the test. In addition, the method to control the temperatures at the start of the test will be optimised to mitigate the inertia.

2.5 Conclusion

The paper presents potential LIB abuse scenarios, which may happen during everyday service. The abuse may be categorised with the type of the strain induced on the LIB - that is mechanical, electrical or thermal abuse. It investigates fire testing procedures currently in force that regulate or standardise BESS and presents their specifics. The comparison on the procedures is made in terms of exposure conditions and differences in testing procedure.

Lastly, the alternative fire resistance testing method is tested that may be implemented into fire resistance testing of BESS. The repeatability of the alternative method promise a significant improvement on the safety testing of BESS. Despite the initial deviation from the standard temperature, the heating curve optimum was reached in the 4. minute of the test and kept for the remaining duration. The initial temperature rise needs to be adjusted to obtain faster designed value, however, a following work will be a subject of next study. In addition, the DUT being kept in a insulated chamber creates safety working environment with potentially dangerous samples.

Based on the foregoing results, the following conclusions were drawn.

1. The LIB abuse types are mechanical, electrical and thermal. The abuse lead to the event of thermal runaway, which is a cause of a vehicle fire.
2. Among three types of fire resistant procedures: open flame, LPG burner and radiative source - the radiative source procedure showed the most accuracy and repeatability. The highest deviation was noted in the open flame methods, which is currently in force.
3. EN 1363-1 alternative fire resistance testing procedure apparatus can be used to simulate designed fire scenarios with improved repeatability. The maximum difference between tests in terms of temperatures was 56.5°C.
4. Alternative fire resistance testing procedure based on EN 1363-1 passive fire resistance testing showed improved repeatability and accuracy in terms of achieving designed exposure conditions. The method can be adapted to simulate appropriate testing conditions to test BESS fire resistance.

Acknowledgements

The research was carried out within the fourth edition of the Program of the Polish Ministry of Science and Higher Education entitled ”Doktorat wdrożeniowy” (Industrial Doctoral Program).

Bibliography

[1] IEA. Global ev outlook 2021. Report, International Energy Agency, 2021. [Online; accessed 14 February 2022].

[2] Lucien Mathieu. Electric surge: Carmakers' electric car plans across europe 2019-2025. Report, European Federation for Transport and Environment AISBL.

[3] Organisation Internationale Des Constructeurs Automobiles. Global auto production in 2020 severely hit by covid-19 crisis with a 16% drop in world auto production, 2022. [Online; accessed 14 February 2022].

[4] Daniel Darnikowski, Magdalena Mieloszyk, and Mateusz Weryk. Fire resistance of construction elements. In *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2022*, volume 12046, pages 203–218. SPIE, 2022.

[5] Xuning Feng, Minggao Ouyang, Xiang Liu, Languang Lu, Yong Xia, and Xiangming He. Thermal runaway mechanism of lithium ion battery for electric vehicles: A review. *Energy Storage Materials*, 10:246–267, 2018.

[6] Deloitte. 2020 global automotive consumer study. Report, Deloitte, 2020. [Online; accessed 14 February 2022].

[7] Mohammadmahdi Ghiji, Vasily Novozhilov, Khalid Moinuddin, Paul Joseph, Ian Burch, Brigitta Suendermann, and Grant Gamble. A review of lithium-ion battery fire suppression. *Energies*, 13(19), 2020.

[8] Yuqing Chen, Yuqiong Kang, Yun Zhao, Li Wang, Jilei Liu, Yanxi Li, Zheng Liang, Xiangming He, Xing Li, Naser Tavajohi, and Baohua Li. A review of lithium-ion battery safety concerns: The issues, strategies, and testing standards. *Journal of Energy Chemistry*, 59:83–99, 2021.

[9] Huang Li, Qiangling Duan, Chunpeng Zhao, Zonghou Huang, and Qingsong Wang. Experimental investigation on the thermal runaway and its propagation in the large format battery module with li (ni1/3co1/3mn1/3) o2 as cathode. *Journal of hazardous materials*, 375:241–254, 2019.

[10] Peiyi Sun, Roeland Bisschop, Huichang Niu, and Xinyan Huang. A review of battery fires in electric vehicles. *Fire technology*, 56(4):1361–1410, 2020.

[11] Adam Dorsz and Miroslaw Lewandowski. Analysis of fire hazards associated with the operation of electric vehicles in enclosed structures. *Energies*, 15(1):11, Dec 2021.

[12] Sofia Ubaldi, Cinzia Di Bari, Armando De Rosa, Mazzaro Michele, and Paola Russo. Investigation on effective fighting technology for lib fire. *Chemical Engineering Transactions*, 91:505–510, 2022.

[13] Cagri Un and Kadir Aydın. Thermal runaway and fire suppression applications for different types of lithium ion batteries. *Vehicles*, 3(3):480–497, 2021.

[14] Fredrik Larsson, Petra Andersson, Per Blomqvist, and Bengt-Erik Mellander. Toxic fluoride gas emissions from lithium-ion battery fires. *Scientific reports*, 7(1):1–13, 2017.

[15] Antonio Nedjalkov, Jan Meyer, Michael Köhring, Alexander Doering, Martin Angelmahr, Sebastian Dahle, Andreas Sander, Axel Fischer, and Wolfgang Schade. Toxic gas emissions from damaged lithium ion batteries – analysis and safety enhancement solution. *Batteries*, 2(1):5, 2016.

[16] Khalil Amine, Ilias Belharouak, Zonghai Chen, Taison Tran, Hiroyuki Yumoto, Naoki Ota, Seung-Taek Myung, and Yang-Kook Sun. Nanostructured anode material for high-power battery system in electric vehicles. *Advanced materials*, 22(28):3052–3057, 2010.

[17] Qingsong Wang, Ping Ping, Xuejuan Zhao, Guanquan Chu, Jinhua Sun, and Chunhua Chen. Thermal runaway caused fire and explosion of lithium ion battery. *Journal of power sources*, 208:210–224, 2012.

[18] Roeland Bisschop, Ola Willstrand, and Max Rosengren. Handling lithium-ion batteries in electric vehicles: preventing and recovering from hazardous events. *Fire technology*, 56(6):2671–2694, 2020.

[19] Elham Sahraei, Joseph Meier, and Tomasz Wierzbicki. Characterizing and modeling mechanical properties and onset of short circuit for three types of lithium-ion pouch cells. *Journal of Power Sources*, 247:503–516, 2014.

[20] Wei-Jen Lai, Mohammed Yusuf Ali, and Jwo Pan. Mechanical behavior of representative volume elements of lithium-ion battery cells under compressive loading conditions. *Journal of Power Sources*, 245:609–623, 2014.

[21] Elham Sahraei, Rich Hill, and Tomasz Wierzbicki. Calibration and finite element simulation of pouch lithium-ion batteries for mechanical integrity. *Journal of Power Sources*, 201:307–321, 2012.

[22] Lars Greve and Clemens Fehrenbach. Mechanical testing and macro-mechanical finite element simulation of the deformation, fracture, and short circuit initiation of cylindrical lithium ion battery cells. *Journal of Power Sources*, 214:377–385, 2012.

[23] Yong Xia, Tomasz Wierzbicki, Elham Sahraei, and Xiaowei Zhang. Damage of cells and battery packs due to ground impact. *Journal of Power Sources*, 267:78–97, 2014.

[24] Sheng Shui Zhang. A review on electrolyte additives for lithium-ion batteries. *Journal of Power Sources*, 162(2):1379–1394, 2006.

[25] Diego Lisbona and Timothy Snee. A review of hazards associated with primary lithium and lithium-ion batteries. *Process safety and environmental protection*, 89(6):434–442, 2011.

[26] Hui-Fang LI, Jun-Kui GAO, and Shao-Li ZHANG. Effect of overdischarge on swelling and recharge performance of lithium ion cells. *Chinese journal of chemistry*, 26(9):1585–1588, 2008.

[27] Akos Kriston, Andreas Pfrang, Harry Döring, Benjamin Fritsch, Vanesa Ruiz, Ibtissam Adanouj, Theodora Kosmidou, Jürgen Ungeheuer, and Lois Boon-Brett. External short circuit performance of graphite-lini1/3co1/3mn1/3o2 and graphite-lini0. 8co0. 15al0. 05o2 cells at different external resistances. *Journal of Power Sources*, 361:170–181, 2017.

[28] Economic Commission for Europe of the United Nations. Global technical regulation on the electric vehicle safety (evs). Regulation, March 2016.

[29] Zhi Wang, Han Yang, Yan Li, Guo Wang, and Jian Wang. Thermal runaway and fire behaviors of large-scale lithium ion batteries with different heating methods. *Journal of hazardous materials*, 379:120730, 2019.

[30] Pengjie Liu, Chaoqun Liu, Kai Yang, Mingjie Zhang, Fei Gao, Binbin Mao, Huang Li, Qiangling Duan, and Qingsong Wang. Thermal runaway and fire behaviors of lithium iron phosphate battery induced by over heating. *Journal of Energy Storage*, 31:101714, 2020.

[31] Daniel Darnikowski and Magdalena Mieloszyk. Investigation into the lithium-ion battery fire resistance testing procedure for commercial use. *Batteries*, 7(3):44, 2021.

[32] Noshin Omar, Mohamed Daowd, Omar Hegazy, Grietus Mulder, Jean-Marc Timmermans, Thierry Coosemans, Peter

Van den Bossche, and Joeri Van Mierlo. Standardization work for bev and hev applications: Critical appraisal of recent traction battery documents. *Energies*, 5(1):138–156, 2012.

[33] Dongxu Ouyang, Mingyi Chen, Que Huang, Jingwen Weng, Zhi Wang, and Jian Wang. A review on the thermal hazards of the lithium-ion battery and the corresponding countermeasures. *Applied Sciences*, 9(12), 2019.

[34] Uniform provisions concerning the approval of vehicles with regard to specific requirements for the electric power train. Regulation, Economic Commission for Europe of the United Nations, August 2013.

[35] Uniform provisions concerning the approval of vehicles with regard to specific requirements for the electric power train. Regulation, Economic Commission for Europe of the United Nations, August 2021.

[36] Electrically propelled road vehicles – Test specification for lithium-ion traction battery packs and systems – Part 3: Safety performance requirements. Standard, International Organization for Standardization, Geneva, CH, May 2014.

[37] Electrically propelled road vehicles – Safety specifications – Part 1: Rechargeable energy storage system (RESS). Standard, International Organization for Standardization, Geneva, CH, April 2019.

[38] Electrically propelled mopeds and motorcycles – Test specifications and safety requirements for lithium-ion battery systems. Standard, International Organization for Standardization, Geneva, CH, April 2017.

[39] Electric and Hybrid Electric Vehicle Rechargeable Energy Storage System (RESS) Safety and Abuse Testing. Standard, Society of Automotive Engineers, Warrendale, US, April 2009.

[40] Lithium-ion traction battery pack and system for electric vehicles – Part 3:Safety requirements and test methods. Stand-

ard, Standardization Administration of the People's Republic of China, Shenzhen, CN, May 2015.

[41] Recommended Practices for Abuse Testing Rechargeable Energy Storage Systems (RESSs). Standard, Sandia National Laboratories, Albuquerque, US, July 2017.

[42] Batteries for Use In Electric Vehicles. Standard, Underwriters Laboratories, Northbrook, US, March 2020.

[43] S Mahesh and DP Mishra. Flame structure of lpg-air inverse diffusion flame in a backstep burner. *Fuel*, 89(8):2145–2148, 2010.

[44] Hyuk Jung, Bohyun Moon, and Gwang Goo Lee. Development of experimental apparatus for fire resistance test of rechargeable energy storage system in xev. *Energies*, 13(2), 2020.

[45] Vytenis Babrauskas. Estimating large pool fire burning rates. *Fire technology*, 19(4):251–261, 1983.

[46] V.I. Bilnov and G.N. Khudiakov. Diffusion burning of liquids. *Diffusion Burning of Liquids; US Army Translation, NTIS*, AD296762, 1974.

[47] Juan Carlos Cuevas and Francisco J García-Vidal. Radiative heat transfer. *Acs Photonics*, 5(10):3896–3915, 2018.

[48] Polish Committee for Standardization. *EN 1363-1:2020 - Fire reistance tests - Part 1: General requirements.*

[49] International Organization for Standardization. *ISO 834-1:1999 - Fire-resistance tests - Elements of building construction - Part 1: General requirements.*

[50] Polish Committee for Standardization. *EN 1363-1:2020 - Fire reistance tests - Part 1: General requirements.*

Chapter 3

Deep Learning based Damage Imaging Techniques

Abdalraheem A. Ijjeh

Institute of Fluid Flow Machinery, Polish Academy of Sciences,
Fiszera 14, 80-231 Gdansk, Poland

3.1 Introduction

Composite materials have gained popularity in recent years in a variety of industries, including aerospace, automobiles, construction, marine, and others, due to their light-weight, excellent fatigue, and corrosion resistance. However, composite materials may suffer from matrix cracks, fibre breakage, debonding, and delamination [1, 2]. Delamination (separation of layers in a laminate composite) is one of the most dangerous of these flaws since it occurs largely below top surfaces and is scarcely noticeable [3].

Delamination in composite materials can be caused by a variety of factors, including manufacturing flaws, notches, and impact events. As a result, delamination can diminish the engineering structure's strength and performance. To avoid such complications, real-time delamination detection is necessary. In order to monitor the integrity of structures, numerous physics-based approaches for damage detection and localization have been developed in the disciplines of structural health monitoring (SHM) and non-destructive testing (NDT).

Guided waves, namely Lamb waves, are used in a well-known physics-based technique in the field of SHM for damage diagnosis. Lamb waves are elastic waves that propagate through thin plates and shells that are bounded by stress-free surfaces [4]. Lamb waves are notable for their strong sensitivity to discontinuities (cracks, delaminations) and low amplitude loss, particularly in metallic elements of structures [5].

An array of Piezoelectric transducers (PZT's) can be used to excite the examined structure and generate Lamb waves, which can subsequently be registered as reflected waves from damage. After that, a damage influence map is generated. The number of sensing points determines the accuracy of the damage influence map in displaying damage location. As a result, damage localization resolution may be limited. Alternatively, a scanning laser Doppler vibrometer (SLDV) is used to measure Lamb waves in a dense grid of points over the structure under investigation. The collected data represents a complete wavefield propagation, resulting in high-resolution

damage influence maps. Damage detection systems that use whole wavefield signals can accurately estimate the extent and location of damage [6, 7].
Support vector machine (SVM) [8–10], K-Nearest Neighbor (KNN) [11], decision tree [12], particle swarm optimization (PSO) [13, 14], principle component analysis (PCA) [15–17] are considered a conventional machine learning methods that are utilised for damage detection with SHM. Such methods among others have proven to be capable of detecting damage in the structures under investigation. However, when dealing with large amounts of data, such methods have drawbacks since they demand a sophisticated feature engineering calculation [18], which also necessitates a high level of expertise and abilities to extract damage-sensitive features for specific SHM applications. As a result, in recent years, a data-driven solutions for SHM applications have emerged in the form of deep learning (DL) end-to-end approaches, which automate the feature engineering and classification processes.

High- and abstract-level features can be translated into a hierarchical order of simple- and low-level obtained features using deep learning techniques [19]. As a result, DL approaches may handle big issues by breaking them down into a huge number of basic blocks. Another important benefit of applying deep learning techniques is so-called transfer learning, which refers to the ability to reuse a pre-trained model created for one task in another.

Due to rapid advancements in computer hardware and software, big data, and cloud-based computations, DL techniques in many SHM domains have gotten increased attention in recent years [20]. For SHM of civil engineering structures, several DL-based appro-aches were used for damage detection and localization [21, 22], corrosion detection [23], and concrete crack detection [24].

De Assis et al. [25] presented a comparative study based on modal responses for crack identification in laminated composites. Furthermore, the authors of this paper solved an inverse crack identification problem using the metaheuristic sunflower optimization (SFO) algorithm, artificial neural networks (ANNs), and the response surface approach. The results showed that the SFO and

ANN approaches may both be utilized to estimate a crack's location, size, and orientation. Furthermore, Oliver et al. [26] developed ANNs for delamination detection in composite laminates that were merely trained on frequency fluctuation values as inputs. The proposed model had a damage quantification success rate up to 95%.

In comparison to vibration-based SHM, DL applications for guided wave-based SHM have received less attention in the literature.

Several DL approaches for guided wave-based damage identification and localization are discussed in the following sections. For damage detection in curved composite panels, Chetwynd et al. [27] suggested a multi-layer perceptron (MLP) network. A force applicator with a circular tip loaded by a mass was used to imitate the damage. Lamb waves propagating through the panel were also generated and registered using an array of PZT. In addition, for each Lamb wave response, a novelty index was calculated and compared to specific threshold values. As a result, if the index exceeds the threshold, it shows that the structure is damaged. The generated novelty indices were fed into the MLP network, which performed two operations: classification and regression. The classification network was created to assess whether the panel was damaged or not by defining three convex sections of the panel. The regression network, on the other hand, can pinpoint the specific area of the damage.

In addition, the authors [28] presented an ANN model for detecting damage in plates comprised of aluminum alloys and composites. The ANN was trained using synthetic data generated with the finite element approach. Furthermore, the researchers calculated damage indexes using the data collected from propagating Lamb waves.

Authors in [29] developed a CNN model for damage state prediction based on full wavefield scans of thin aluminum plates. When compared to SVM, the model achieved superior accuracy in terms of damage (99.98%) compared to SVM (62%).

Ewald et al. [30] presented a CNN model for signal categorization using Lamb waves (DeepSHM). The model uses reflected signals acquired by sensors to perform an end-to-end strategy for SHM. The authors also used the wavelet transform to preprocess response data

in order to generate the wavelet coefficient matrix (WCM), which was then input into the CNN model.

Liu and Zhang [31] developed a CNN model for detecting damage in thin aluminum plates. Analytical formulas have been derived for generating Lamb wave signals subsequently used for training and validation of the proposed crack damage detection technique. Furthermore, the authors validated their model by putting it to the test on real-world data using a notch crack to simulate damage.

In addition, Esfandabadi et al. [32] studied the use of compressive sensing technique [33] in conjunction with super-resolution techniques to obtain high-resolution wavefield scans using neural networks trained on various aluminum and carbon fibre reinforced polymer (CFRP) plates. To recover high spatial frequency information from low-resolution wavefield scans, the authors implemented two types of CNN architecture: Super-Resolution Convolutional Neural Networks (SRCNNs) and Very-Deep Super Resolution (VDSR) using compressive sensing. Enhancing the resolution, on the other hand, has a negative impact on the affected region, changing the damage characteristics.

Current SHM approaches are able to effectively localize impact events [34], or damage [35]. However, with a sparse array of sensors, determining the magnitude and shape of damage is challenging. To solve this issue, full wavefield ultrasonic methods can be used locally to assess damage severity and then estimate damage prognosis.

In this chapter, we compare five DL models for semantic image segmentation that have been used in CFRP for delamination detection, localization, and size estimation. To demonstrate their generalization ability, the models were tested using numerical and experimental data.

3.2 Methodology

3.2.1 Dataset

The synthetically generated dataset [36] was utilised for training several DL models. The dataset resembles velocity measurements

obtained by SLDV in the transverse direction (perpendicular to the plate surface). The dataset contains 475 simulated full wavefield scans of propagating Lamb waves in an eight-layer CFRP plate with a total thickness of 3.9 mm interacting with delamination.

Delamination is represented by two shell layers that are only connected at the damaged area's boundary. The corresponding, reduced, number of laminae and their offset from the plate's neutral axis are used to calculate these elements. This method is described in detail in [37], has the advantage of keeping the computing efficiency of the model as compared to solid elements, as the increase in nodes is usually small. In addition, the mass matrix's favorable attribute of being diagonal is maintained. Each simulated scenario depicts the interaction of Lamb waves with a single delamination, which was modeled using random parameters as like spatial location, size, and orientation. Fig. 3.1 shows a plate with 475 delamination cases overlayed.

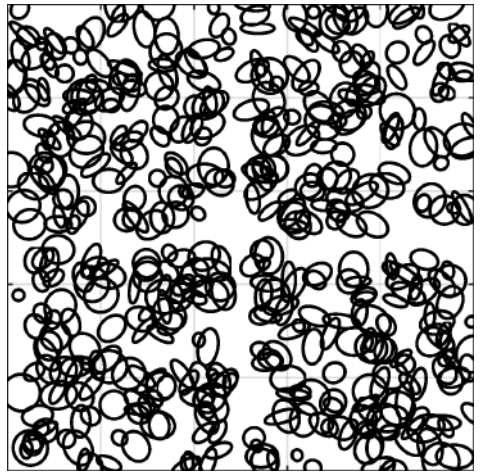


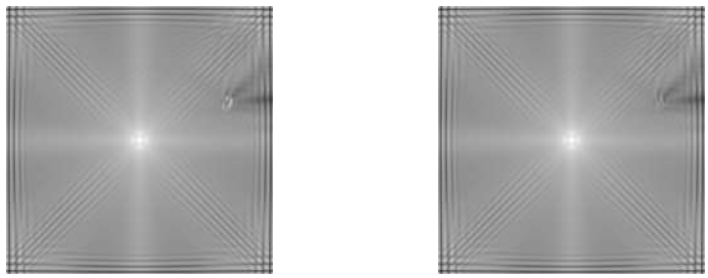
Fig. 3.1: The plate with 475 cases of random delaminations.

The amplitudes of the propagating waves at location (x, y) and time (t) are stored in each simulated case in the dataset generated from the wave propagation model. As a result, these matrices can be viewed as animated frames of propagating waves at dis-

crete time (t_k). Furthermore, it should be noted that the simulated delaminations were closer to the top of the plate surface. As a result, utilizing the simulated full wavefield on the top surface of the plate instead of the bottom surface makes it easier to detect delamination. However, we used the more difficult situation of registering the entire wavefield at the bottom surface of plate to train several models. The root mean square (RMS) depicted in Eq. 3.1 was applied to the full wavefield frames in order to improve the visualisation and shrink data size:

$$\hat{s}(x,y) = \sqrt{\frac{1}{N}\sum_{k=1}^{N} s(x,y,t_k)^2}, \tag{3.1}$$

where N denotes the number of sampling points (N=512). The results of applying RMS to the full wavefield from the top and bottom surfaces of the plate can be seen in 3.2a(a) and 3.2a(b).



(a) top (b) bottom

Fig. 3.2: RMS of the full wavefield from the top surface of the plate (a) and the bottom surface of the plate (b).

3.2.2 Data preprocessing

To improve the performance of the optimizer throughout the training phase, the colour scale values were standardized to a range of (0-1) instead of the initial scale which was in the range of (0-255). Additionally, the dataset was augmented by flipping the images horizontally, vertically, and diagonally. As a result, the dataset expanded four times in size, yielding (1 900) image frames. Further,

the dataset was divided into two parts: the training set (80%) and the testing set (20%). Moreover, a K-folds cross-validation method was applied to the training set to reduce the overfitting which happens when the model is able to fit on the training data, while it poorly fit on the new unseen data. In other words, the model only learns the patterns of the training data therefore the model will not generalise well. Figure. 3.3 illustrates the K-fold cross validation (CV) technique. In this technique, we have split the training set into K small sets (folds), hence the name K-folds. Therefore, we iterate over the training set K iterations. During each iteration, the model uses $K-1$ folds for training and the remaining fold is used for validation. In our models, we have chosen $K=5$, therefore, we have 5 iterations of training. For each iteration, we computed the performance of the model. Finally, the iteration with highest performance is selected. The main advantage of the K-folds method versus a regular train/test split is to reduce the overfitting by utilising data more efficiently as every data sample is used in both training and validation. Therefore, by using this technique, we aim to improve the ability of the model to generalise and reduce the overfitting.

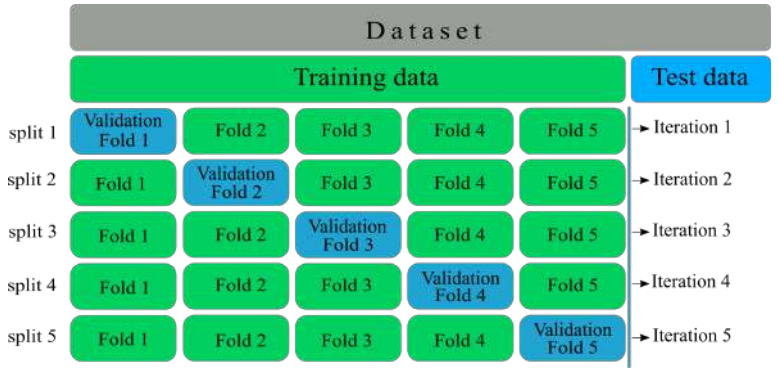


Fig. 3.3: K-fold cross validation, K=5.

3.2.3 Semantic segmentation models

DL approaches have advanced quickly in recent years in a variety of real-world applications, such as computer vision. The technique of image segmentation is well-known in the field of computer vision. This technique aims to assign a class to each pixel in the input image, and it is used in a variety of real-world applications like self-driving automobiles, medical imaging, traffic management systems, video surveillance, and more. In this chapter, we compared five DL models based on fully convolutional networks (FCN) [38] in order to identify and localize delamination in composite plates. Similarly, these models attempt to perform pixel-wise segmentation by classifying each pixel in the input image as damaged or undamaged. FCN can be generated by combining convolutional layers and skipping dense layers in an encoder-decoder architecture. The encoder aims to produce compressed feature maps from the input image at various scale levels using cascaded convolutions and downsampling operations. While the decoder is responsible for upsampling the condensed feature maps to the original input shape.

The softmax activation function was used at the output layer for all constructed models. For every single pixel, the softmax estimates the probability of each predicted output being damaged or undamaged, implying that the sum of the two probabilities must be one. The softmax activation function is depicted by Eq. (3.2), where $P(x)_i$ is the probability of each target class x_j across all potential target classes x_j, C in our instance being two classes (damaged and undamaged). An argmax function is used to find the maximum probability between each of them in order to predict the label of the output (y_{pred}).

$$P(x)_i = \frac{e^{x_i}}{\sum_j^C e^{x_j}}, \tag{3.2}$$

$$y_{pred} = \text{argmax}_i \left(P(x)_i \right), \tag{3.3}$$

Choosing the right loss function is crucial since it determines how effectively the model learns and performs. As a result, the categorical cross-entropy (CCE) loss function [39], commonly known as the "softmax loss function", was utilised. The difference between the

actual damage (ground truth) and the expected damage is estimated using CCE as the objective function. Furthermore, because there are only two classes to predict, a sigmoid activation function at the output layer can be combined with a binary cross-entropy (BCE) without affecting the anticipated outputs. The CCE is illustrated by Eq. (3.4), where $P(x)_i$ refers to the softmax value of the target class.

$$CCE = -\log\left(P(x)_i\right), \tag{3.4}$$

Further, it is vital to choose an appropriate accuracy metric for the training purposes. Hence, intersection over union IoU (Jaccard index) [40] was utilised to measure the accuracy of prediction during training process. The intersection area between the ground truth and the predicted output is used to estimate IoU. In this work, the (IoU) is calculated solely for the damaged class label. Eq. (3.5) defines the (IoU) metric as follows:

$$IoU = \frac{Intersection}{Union} = \frac{\hat{Y} \cap Y}{\hat{Y} \cup Y}, \tag{3.5}$$

where $\hat{Y}$ refers to the predicted tensor values, while Y refers to the tensor of ground truth values.

Moreover, the Adam optimizer was used as our optimizer to improve the (IoU) and decrease the loss during training. The implemented DL models for pixel-wise semantic segmentation for delaminations identification are depicted in Fig. 3.4. In the following, five DL models will be illustrated.

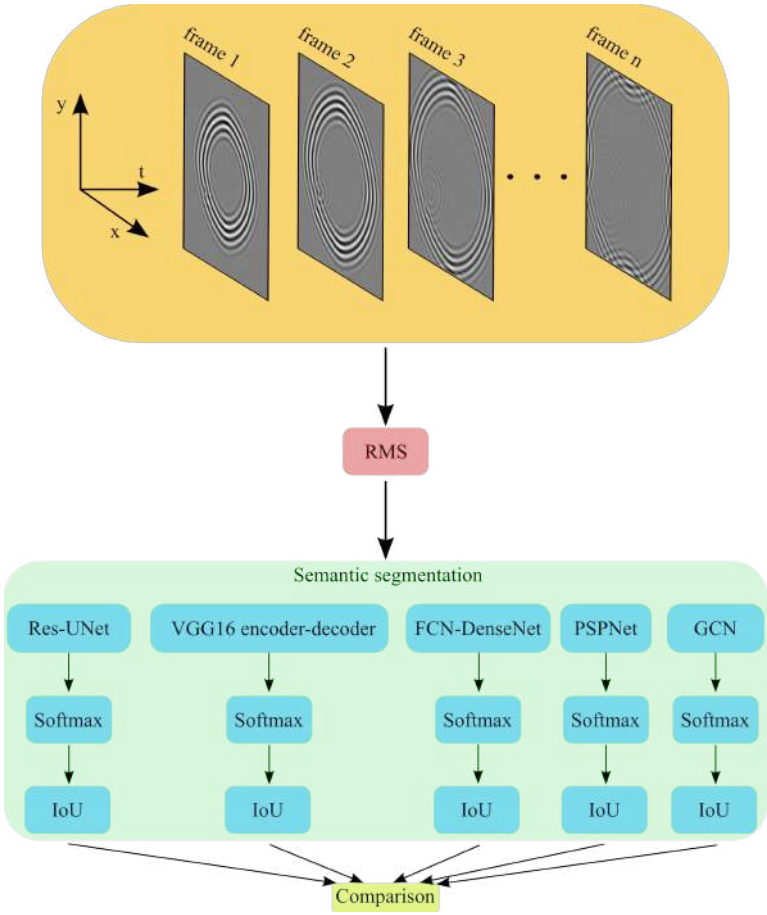


Fig. 3.4: Schematic diagram of the approach used for comparison of semantic segmentation methods accuracy.

Residual UNet model

The Residual UNet (Res-UNet) model was inspired based on residual learning [41] and UNet approaches [42]. The Res-UNet architecture is depicted in Fig. 3.5. The encoder (compressive) path aims to capture the detailed features of an input image, whereas the decoder (decompressive) path aims to perform exact localization. As a result, residual connections were established at two levels in

order to prevent the spatial and contextual information from the preceding layers from being lost:

- at each step of the encoder and decoder paths,
- between the encoder parts and their corresponding decoder parts (skip connections) which ensures that the feature maps which were learned during the downsampling will be utilized in the reconstruction.

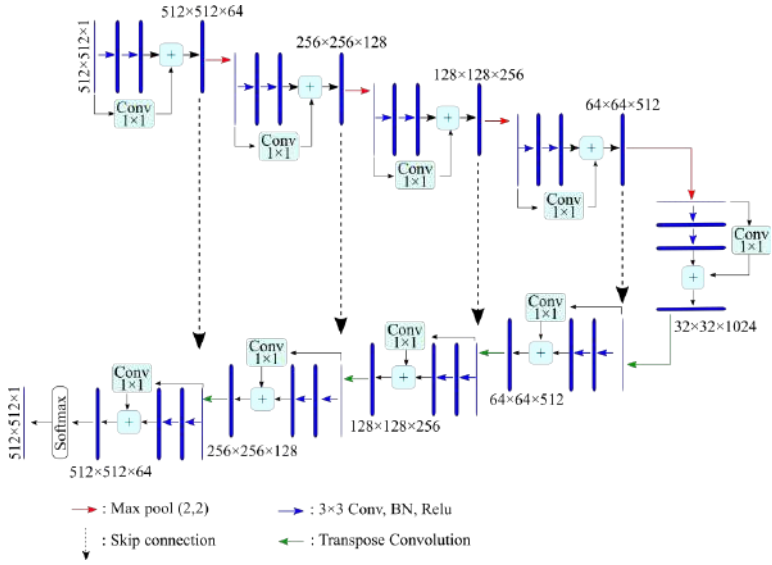


Fig. 3.5: Res-UNet architecture.

Several downsampling (Max-pool) blocks are used in the encoder section. Each block applies two convolutional layers followed by a (2×2) max pooling with a (2×2) strides that selects the maximum value in a local pool filter in one feature map (or n-feature maps), resulting in a reduction in the dimension of feature maps [43], as a result, a reduction in computation complexity. Each convolutional layer does (3×3) convolution operations, then batch normalization (BN), finally a Relu activation function. Furthermore, after each downsampling block, the number of convolutional filters is increased, allowing the model to learn complex patterns successfully.

The bottleneck layer is a joining point in the model's deepest layer, located between the encoder and the decoder. Two convolutional layers with (1024) filters make up the bottleneck, which aids the model in learning and recognizing complex features.

The decoder is composed of a number upsampling blocks that function together to recover original input dimensions and improve resolution. As in the downsampling block, each upsampling block transmits the input through two convolution layers, followed by a transmission up layer consisting of a transposed convolutional layer (upsampling). The transposed convolutional layer varies from the standard upsampling function in that it introduces learnable parameters for the transposed convolution filters, which improve the model's learning process. Furthermore, the number of filters used by the convolutional layer is reduced by half after each upsampling operation to keep the model symmetrical.

VGG16 encoder-decoder

The application of the VGG16 [44] architecture as a backbone encoder to the UNet [42] approach is addressed in this model. VGG16 is a classification algorithm that consists of 13 convolutional layers, pooling layers, and 3 dense layers. The dense layers were removed form the original VGG16 model, and a 13 convolutional layers were applied resulting in an encoder-decoder scheme for pixel-wise image segmentation. The architecture of the VGG16 encoder-decoder model is shown in Fig. 3.6. The model is U-shaped like, and consists of two parts: encoder and decoder. The encoder is made up of (five) convolutional blocks with a total of (13) (3×3) convolutional layers, followed by BN and Relu as the activation function. After each convolutional block, a Max pool operation with a pool size of (2×2) is conducted, followed by dropout. The upsampling process is used to retrieve spatial resolution, and it contains 5 convolutional blocks of total 13 convolutional layers. Bilinear interpolation with (2×2) kernel size is used for upsampling. In order to improve recovering fine-grained information, skip connections were added between downsampling blocks and the matching upsampling blocks, allowing feature re-usability from earlier layers.

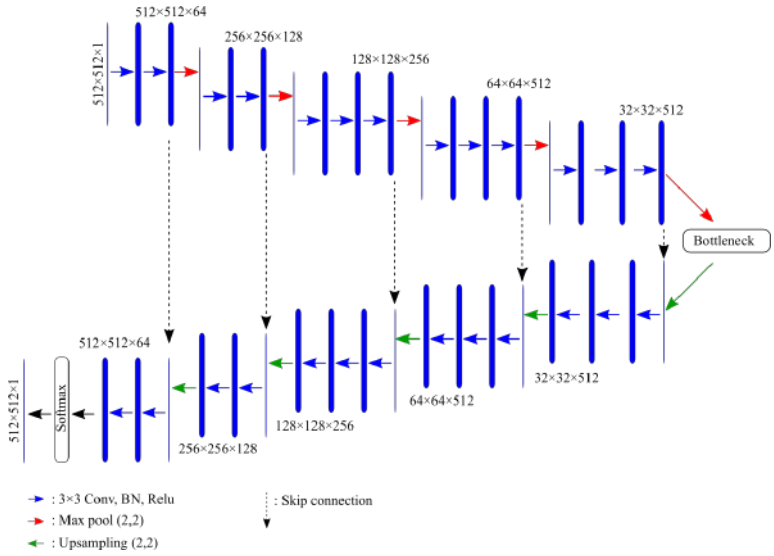


Fig. 3.6: VGG16 encoder decoder architecture.

FCN-DenseNet model

FCN-DenseNet is a pixel-wise image segmentation algorithm that was first introduced in [45]. To boost the resolution of the final feature map, FCN-DenseNet uses a U-shaped encoder-decoder architecture with skip connections between downsampling and upsampling channels. Hence, FCN-DenseNet introduced a dense block representing its main component. The dense block is made up of n layers, each of which is made up of a set of operations, as given in Tab. 3.1. The purpose of the dense block is to concatenate the input (feature maps) of a layer with its output (feature maps) to emphasize spatial details information. The architecture of the dense block is presented in Fig. 3.7.

A transition down layer was added to execute a (1×1) convolution followed by a (2×2) Maxpooling operation to minimize the spatial dimensionality of the resulting feature maps. As a result, a transition-up layer was added to recover the spatial resolution. FCN-DenseNet essentially upsamples feature maps from the previous layer using a transpose convolution technique. Upsampled

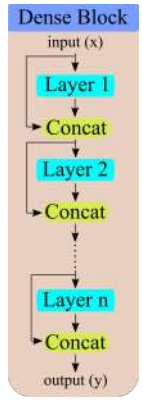


Fig. 3.7: Dense block architecture.

feature maps are concatenated with those produced by the skip connection to provide the input to a new dense block.

As the upsampling approach expands the spatial resolution of the feature maps, the input to the dense block is not concatenated with its output during upsampling to avoid the overhead of memory shortage. The FCN-DenseNet architecture for image segmentation utilized for delamination detection is shown in Fig. 3.8.

Tab. 3.1 presents the architecture of a single layer, the transition down and transition up layers in details.

Tab. 3.1: Layer, Transition Down and Transition Up layers.

Layer	**Transition Down**	**Transition Up**
Batch Normalization	Batch Normalization	3×3 Transposed Convolution
ReLU	ReLU	strides $= (2 \times 2)$
(3×3) Convolution	(1×1) Convolution	
Dropout $p = 0.2$	Dropout $p = 0.2$	
	(2×2) Maxpooling	

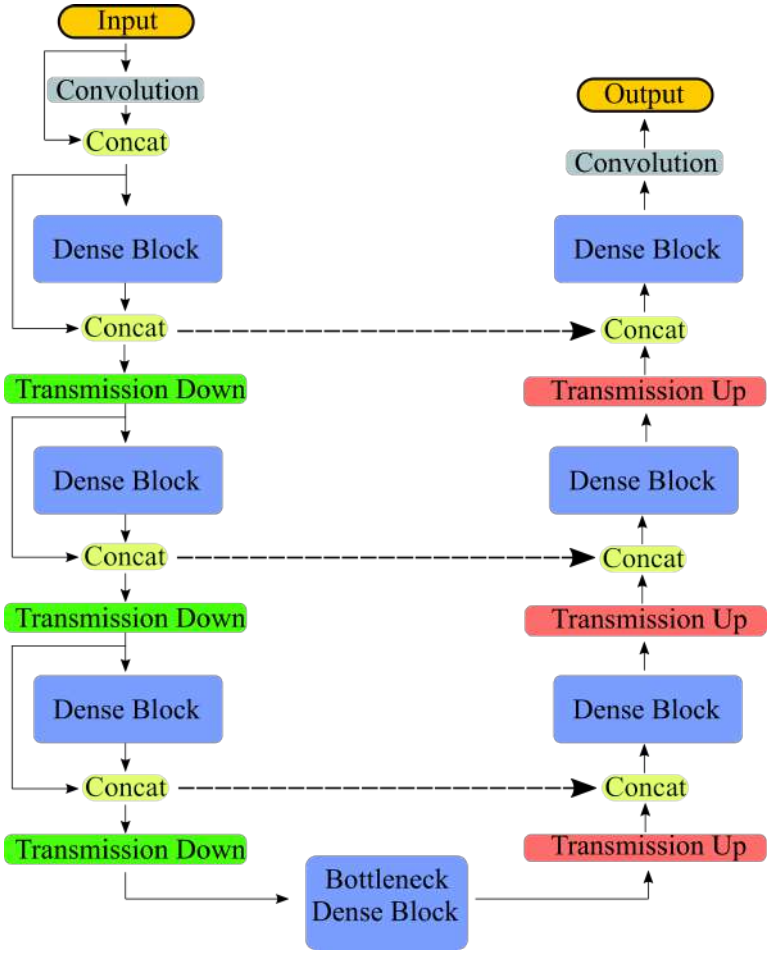


Fig. 3.8: FCN-DenseNet architecture.

Pyramid Scene Parsing Network

The main idea of PSPNet [46] is to combine local and global features to give appropriate global contextual information for pixel-level scene parsing. As a result, a spatial pyramid pooling module was developed to execute four different layers of pooling with four different pooling sizes and strides. The pyramid pooling module is able to capture contextual features from many scales in this way.

To enhance the PSPNet model a ResNet-50 model [41] was added. It works as a backbone for feature map extraction with dilation at the last two layers of ResNet. The implemented PSPNet architecture is shown in Fig. 3.9. Hence, a pyramid pooling module was utilised at 4 pooling levels. The coarsest level of a single bin output depicted in the red box was generated using global average pooling. (2×2), (4×4), and (8×8) are the pooling sizes for the other three sub-region levels, respectively. To minimize the dimensionality of the generated feature maps, a (1×1) convolutional layer was applied, followed by a BN and Relu. Subsequently, bilinear interpolation was used to upsample the feature maps created at each level. Furthermore, the upsampled features are combined with the output of ResNet-50 to produce both local and global context information. The pixel-wise segmentation predictions were then generated using two cascaded convolutional layers.

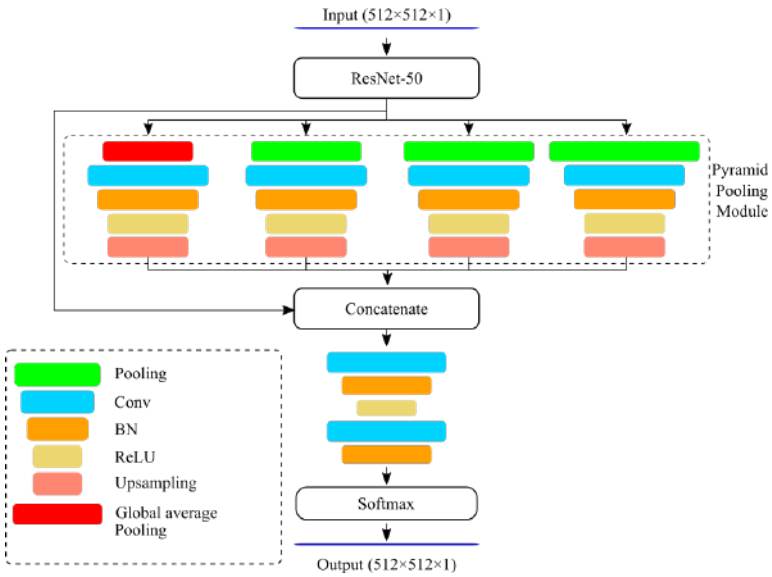


Fig. 3.9: PSPNet architecture.

Global Convolutional Network

Peng et al. [47] introduced the Global Convolutional Network (GCN) to address the importance of having large kernels for both localization and classification operations for semantic segmentation in order to increase the size of respective fields. However, when performing classification and localization tasks, a contradiction emerges due to the fact that classification tasks necessitate invariant models for various transformations such as rotation and translation while localisation tasks necessitate models that are sensitive to any modification and appropriately assign each pixel to its semantic category. To alleviate such contradiction, two design principles were proposed: (1) For the classification task, in order to improve the capability of the model to handle different transformations, a large kernel size must be used to enable dense connections between feature maps and per-pixel classifiers; (2) for localisation task, the model must be fully convolutional. Additionally, fully connected or global pooling layers are not applied as these layers will discard the localisation information.

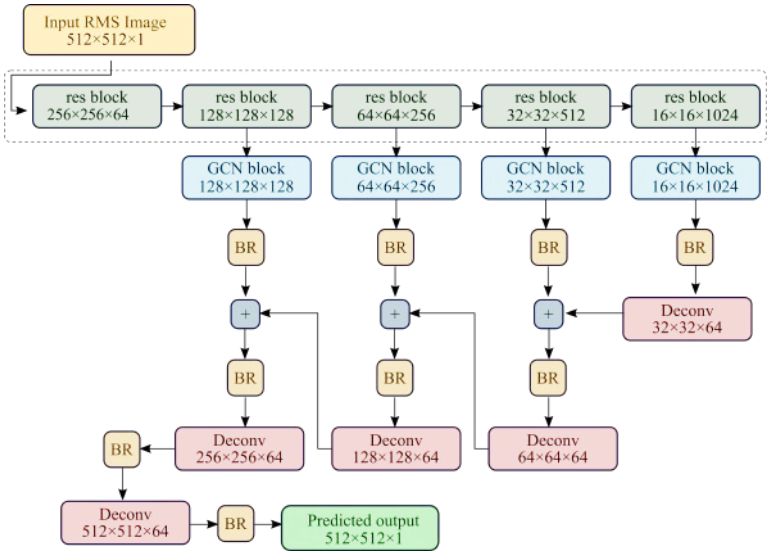


Fig. 3.10: Global Convolution Network whole architecture.

The implemented GCN technique for semantic segmentation is shown in Fig. 3.10. A residual network was used as a backbone for improving the feature extraction process, as demonstrated in Fig. 3.10, further, the residual block is presented in Fig. 3.11a. A GCN block presented in Fig. 3.11b is placed after each residual block, which employs a mix of $(1\times k)+(k\times 1)$ and $(k\times 1)+(1\times k)$ convolutions to establish dense connections within $(k\times k)$ region in the feature map. The boundary refinement (BR) block, depicted in Fig. 3.11c, is then used to improve the predictions along the object borders, resulting in a lower resolution score map. Furthermore, the upsampling operation is done recursively, it upsamples the low resolution score maps then concatenate it with a higher one to produce a new score maps. The deconvolution operation is repeated until the original image size is obtained.

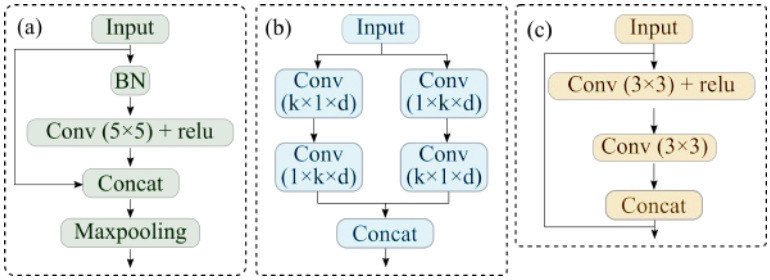


Fig. 3.11: (a) Residual block, (b) Global Convolution Network block, (c) Boundary Refinement.

3.3 Results and discussions

In this section, five DL models of semantic segmentation approach were evaluated on exemplary three damage scenarios of an RMS of the numerically calculated full wavefield interpolated at the bottom surface of the plate in order to identify the delamination, including Res-UNet, VGG16 encoder-decoder, FCN-DenseNet, PSPNet, and GCN. In addition, an experimental scenario was used to measure the performance of the models and further to demonstrate the generalization capabilities. The mean and maximum (IoU) for each

model is calculated and displayed as a comparative metric. Furthermore, for each model, the accuracy of classification, precision, recall, and F1 score were determined using Eqs. (3.6 - 3.9):

$$\text{Accuracy } = \frac{\text{TP} + \text{TN}}{(\text{Total number of tested samples})}, \tag{3.6}$$

$$\text{Pricision } = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{3.7}$$

$$\text{Recall } = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{3.8}$$

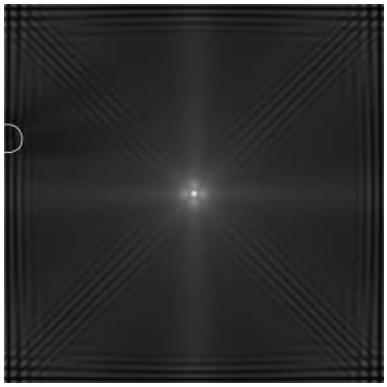
$$\text{F1 score} = \frac{2 \times (\text{Pricision} \times \text{Recall})}{(\text{Pricision} + \text{Recall})}, \tag{3.9}$$

where the Positive/Negative refers to the predicted output as (damage or non-damage) respectively, True Positive (TP) and True Negative (TN) represent the correct classification, and the False Positive (FP) and False Negative (FN) represent the incorrect classification.

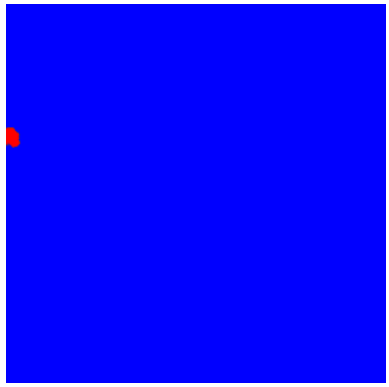
For training purposes, a K-fold cross-validation technique was implemented with five folds. As a result, each model has cycled over 5 iterations of training. Furthermore, for each iteration, the number of epochs equals 20.

3.3.1 Numerical scenarios

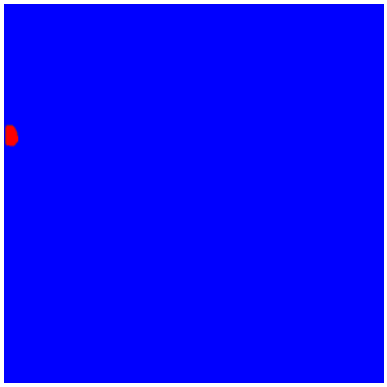
Three numerical data scenarios for delamination at various locations, shape, orientations and angles are discussed in the following. Fig. 3.12 depicts the first studied delamination scenario. As illustrated in Fig. 3.12a, the delamination lies near the left boundary of the plate and is outlined by line to represent its shape, orientation and location. The predictions of Res-UNet, VGG16 encoder-decoder, PSPNet, FCN-DenseNet, and GCN models are shown in Figs. 3.12b - 3.12f. The location of the delamination is correctly indicated in all models. Furthermore, delamination-related pixels are grouped into a single region with no extra noise. Other models, such as the VGG16 encoder-decoder and GCN, are more visually similar to the actual shape of the delamination.



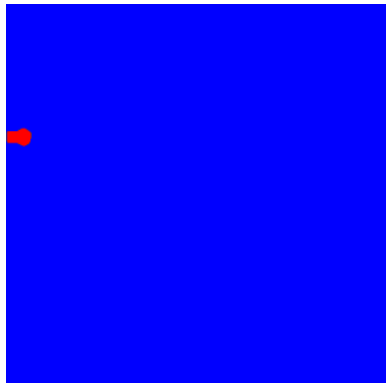
(a) RMS bottom with label.



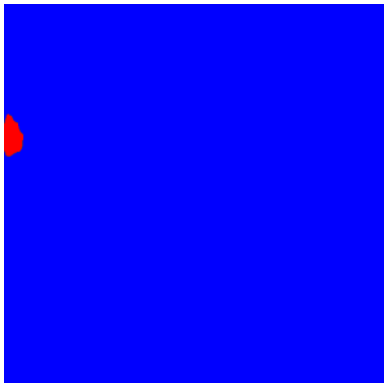
(b) Res-UNet.



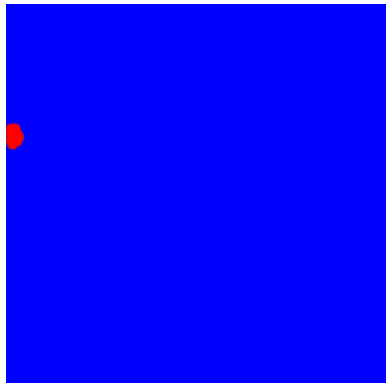
(c) VGG16 encoder-decoder.



(d) PSPNet.



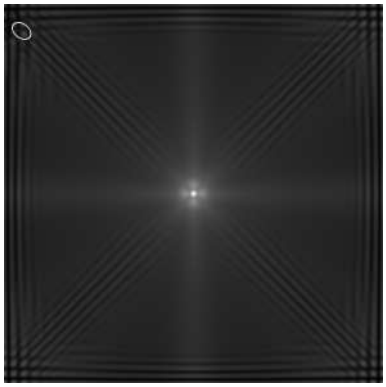
(e) FCN-DenseNet.



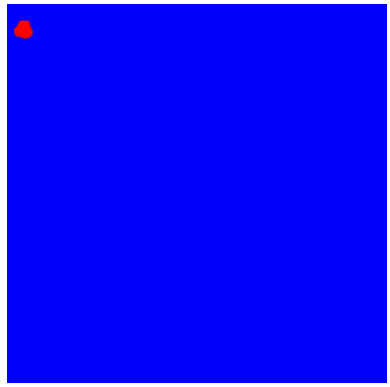
(f) GCN.

Fig. 3.12: First delamination scenario based on numerical data.

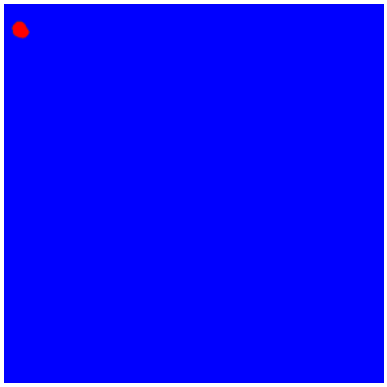
The delamination is positioned at the top left corner of the plate in the second delamination scenario, as shown in Fig. 3.13, and it is surrounded by an oval line to depict its shape and location, as shown in Fig. 3.13a. As the reflections from plate edges overshadow reflections from the delamination, this is the most difficult damage scenario. As a result, RMS pattern changes are barely distinguishable. Figs. 3.13b - 3.13f show the predicted output of the Res-UNet, VGG16 encoder-decoder, PSPNet, FCN-DenseNet and GCN models, respectively. Considering the challenging damage scenario, all models perform reasonably well.



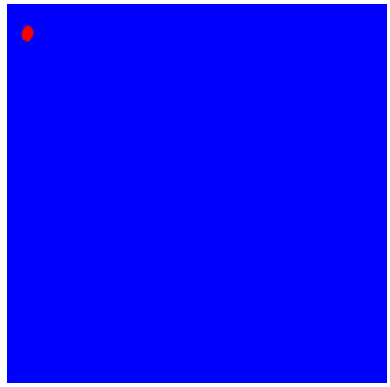
(a) RMS bottom with label.



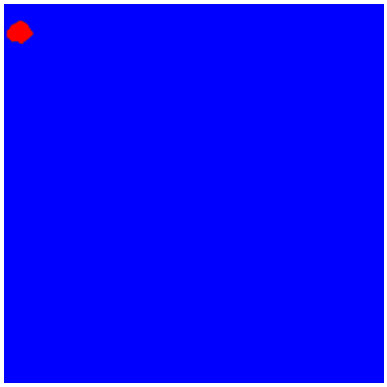
(b) Res-UNet.



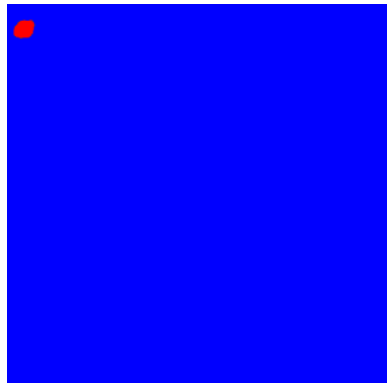
(c) VGG16 encoder-decoder.



(d) PSPNet.



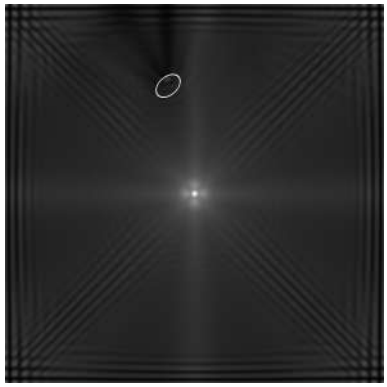
(e) FCN-DenseNet.



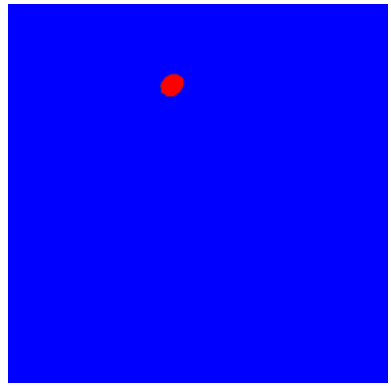
(f) GCN.

Fig. 3.13: Second delamination scenario based on numerical data.

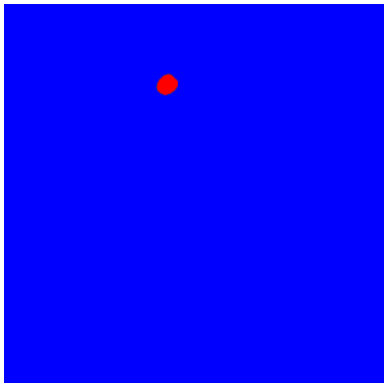
Fig. 3.14 depicts the third delamination scenario. As illustrated in Fig. 3.14a, the delamination lies in the upper middle of the plate and is surrounded by an ellipse to represent its shape, location and orientation. Figs. 3.14b - 3.14f show the prediction outputs of Res-UNet, VGG16 encoder-decoder, PSPNet, FCN-DenseNet, and GCN models, respectively. Res-UNet and FCN-DenseNet are the best in preserving the elliptical shape of delamination in this case. GCN, on the other hand, gets the highest IoU value (see Tab. 3.2).



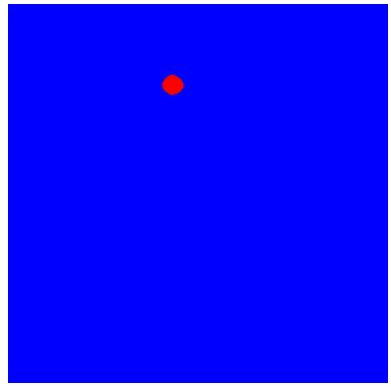
(a) RMS bottom with label.



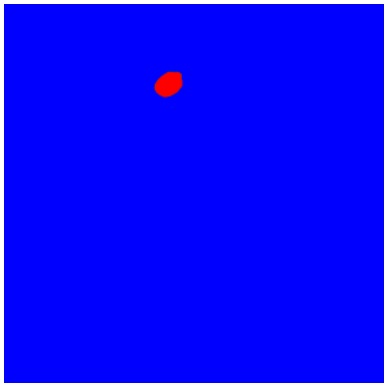
(b) Res-UNet.



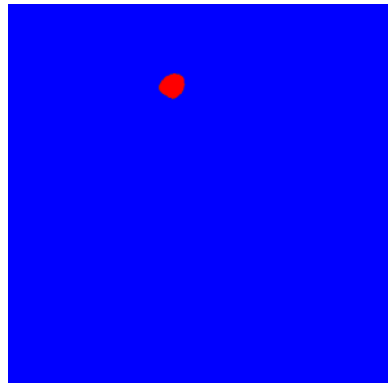
(c) VGG16 encoder-decoder.



(d) PSPNet.



(e) FCN-DenseNet.



(f) GCN.

Fig. 3.14: Third delamination scenario based on numerical data.

Tab. 3.2: *IoU* of numerical scenarios.

Model	**1st scenario**	**2nd scenario**	**3rd scenario**
Res-UNet	0.498	0.782	0.816
VGG16 encoder-decoder	0.512	0.787	0.662
FCN-DenseNet	0.734	0.612	0.866
PSPNet	0.389	0.496	0.646
GCN	0.791	0.696	0.875

Tab. 3.3: Analysis of numerical data.

Model	**mean** *IoU*	**max** *IoU*
Res-UNet	0.664	0.888
VGG16 encoder-decoder	0.572	0.841
FCN-DenseNet	0.680	0.920
PSPNet	0.549	0.914
GCN	0.763	0.931

Tab. 3.2 shows the (IoU) values for all models with respect to predicted damage. The GCN model has the highest (IoU) compared to the other models in the first and third scenarios, and the VGG16 encoder-decoder model has the highest (IoU) compared to the other models in the second scenario. Additionally, there is no noise in the predicted outputs for delamination identification in all the models.

The mean and maximum values of (IoU) obtained for the previously unseen numerical test set (380 cases) for all models are presented in Tab. 3.3. Additionally, Tab. 3.3 demonstrates that all models have a high (IoU), showing their capability to identify damage.

Tab. 3.4 presents the TP, TN, FP, and FN for all models with respect to the test set.

As additional evaluation metrics, Tab. 3.5 displays the classification accuracy, precision, recall, and F1-score values for all models. All models show high classification accuracy, as shown in Tab. 3.5, indicating that all of the offered models are capable of predicting

Tab. 3.4: Model classification performance.

Model	TP	TN	FP	FN
Res-UNet	376	376	4	0
VGG16 encoder-decoder	373	373	7	0
FCN-DenseNet	378	378	2	0
PSPNet	368	368	12	0
GCN	380	380	0	0

Tab. 3.5: Evaluation metric.

Model	Accuracy	Precision	Recall	F1-Score
Res-UNet	0.994	0.989	1.00	0.994
VGG16 encoder-decoder	0.991	0.981	1.00	0.991
FCN-DenseNet	0.997	0.994	1.00	0.994
PSPNet	0.984	0.968	1.00	0.984
GCN	1.00	1.00	1.00	1.00

the existence of delamination in all numerically generated cases. Moreover, it can be concluded that the GCN model outperformed the other implemented models.

Furthermore, the total number of parameters in each DL model is the sum of trainable (convolution filter weights) and non-trainable parameters (biases and pooling filters). The trainable parameters are updated continuously until the minimal loss value is reached, whereas the non-trainable parameters remain unchanged during the training process. The total number of parameters for all implemented models is shown in Tab. 3.6. In addition, the total number of parameters can show the computational complexity of the model. It should be noted that as the number of total parameters increases, the required time for training increases.

3.3.2 Experimental scenario

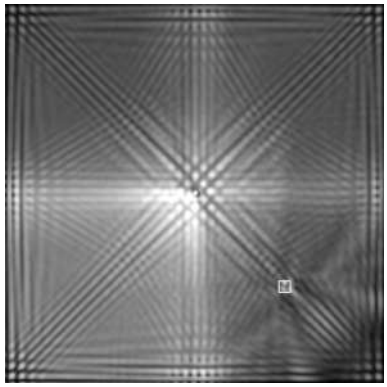
An experimental case of CFRP with Teflon insert as artificial delamination is examined in this scenario which is shown in Fig. 3.15.

Tab. 3.6: Model parameters.

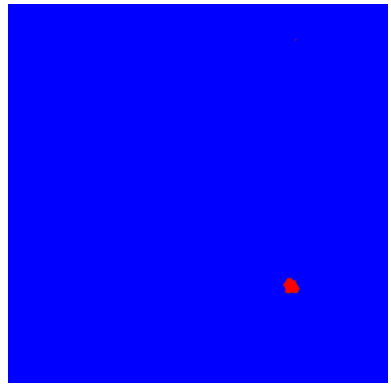
Model	**Total parameters** ($\approx$)
Res-UNet	52×10^6
VGG16 encoder-decoder	37.3×10^6
FCN-DenseNet	2.5×10^6
PSPNet	6.6×10^6
GCN	36×10^6

A frequency of 50 kHz was used to stimulate a signal in a transducer placed in the center of the plate, similar to the synthetic data set. At this frequency, the A0 mode wavelength for this particular CFRP material is around 20 mm. The measurements were conducted with a Polytec PSV-400 SLDV on the bottom surface of plate, that sized 500×500 mm. The sampling frequency was 512 kHz and the measuring grid spacing was 1 mm. An energy compensated RMS was used to further process the obtained full wavefield, considering wave attenuation. Fig. 3.15a shows the results of such a procedure.

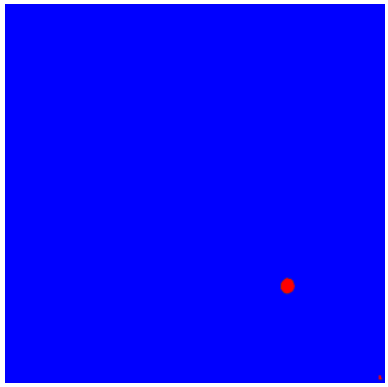
A square frame surrounds the delamination, indicating its shape and location. Delamination prediction maps for Res-UNet, VGG16 encoder-decoder, PSPNet, FCN-DenseNet, and GCN models are shown in Figs. (3.15b - 3.15f), respectively. The performance of the models was assessed using the IoU measure, which considers not only the location of the damage but also its shape and size. The Res-UNet $IoU = 0.577$, the VGG16 encoder-decoder $IoU = 0.624$, the PSPNet (IoU = 0.488), the FCN-DenseNet $IoU = 0.537$, and the GCN $IoU = 0.723$ were the results achieved. The best accuracy was achieved using GCN, the same as it was with the numerical test data.



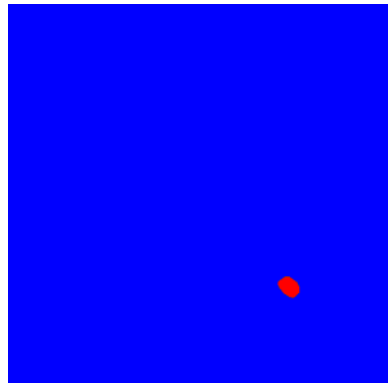
(a) ERMS.



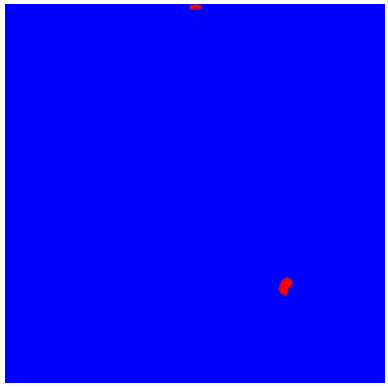
(b) Res-UNet.



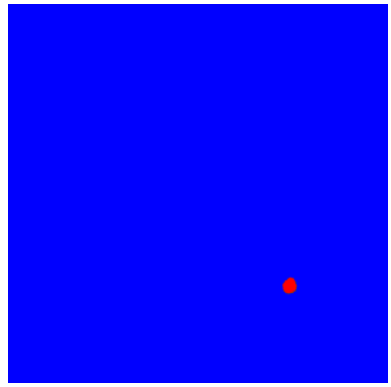
(c) VGG16 encoder-decoder.



(d) PSPNet.



(e) FCN-DenseNet.



(f) GCN.

Fig. 3.15: Experimental results.

As illustrated, the models are capable of detecting and localizing the delamination with high precision accuracy. It cat be noticed that the models can recognize delamination with nearly no noise, indicating that the models can generalize and detect delamination on previously unseen data. Given that the provided models were only trained on a numerically generated data, they had a high degree of generalisation capability. It is expected that when the models are trained on experimental and numerical data, additional features are learned, and their performance can be improved even further.

Summary

In this chapter, five different DL models: Res-UNet, VGG16 encoder-decoder, PSPNet, FCN-DenseNet, and GCN have been introduced and trained to perform image semantic segmentation in order to identify delamination in CFRP materials. The models were trained on a numerically generated dataset simulating a full wavefield of propagating guided waves. The results of DL models in identifying various forms of delaminations in terms of their locations, shapes, sizes, and angles are encouraging. Furthermore, the models demonstrate good generalization when it comes to predicting delamination in numerically acquired data which has not been observed before. Furthermore, the models demonstrate their generalisation ability by detecting delamination in experimentally measured data.

Regarding the numerical test data and the experimental scenario, it can be stated that the GCN model has the highest identification accuracy among all implemented models. Furthermore, the PSPNet model had the lowest identification accuracy of all the models. As a result, among all applied DL models, the GCN model has the best performance. Furthermore, training DL models on experimental data will allow them to learn new complex patterns, hence, their performance will increase.

Bibliography

[1] Kim-Ho Ip and Yiu-Wing Mai. Delamination detection in smart composite beams using lamb waves. *Smart Materials and Structures*, 13(3):544, 2004.

[2] RA Smith. Composite defects and their detection. *Materials science and engineering*, 3(1):103–143, 2009.

[3] Jian Cai, Lei Qiu, Shenfang Yuan, Lihua Shi, PeiPei Liu, and Dong Liang. Structural health monitoring for composite materials. In *Composites and their applications*. IntechOpen, 2012.

[4] Mira Mitra and S Gopalakrishnan. Guided wave based structural health monitoring: A review. *Smart Materials and Structures*, 25(5):053001, 2016.

[5] Casey J Keulen, M Yildiz, Afzal Suleman, et al. Damage detection of composite plates by lamb wave ultrasonic tomography with a sparse hexagonal network using damage progression trends. *Shock and Vibration*, 2014, 2014.

[6] Donato Girolamo, Huan-Yu Chang, and Fuh-Gwo Yuan. Impact damage visualization in a honeycomb composite panel through laser inspection using zero-lag cross-correlation imaging condition. *Ultrasonics*, 87:152–165, 2018.

[7] Pawel Kudela, Maciej Radzienski, and Wieslaw Ostachowicz. Impact induced damage assessment by means of lamb wave image processing. *Mechanical Systems and Signal Processing*, 102:23–36, 2018.

[8] MN Noori, Y Cao, Z Hou, and S Sharma. Application of support vector machine for reliability assessment and structural health monitoring. *International Journal of Engineering Under Uncertainty: Hazards, Assessment and Mitigation*, 2(3-4):89–98, 2010.

[9] Nguyen LD Khoa, Bang Zhang, Yang Wang, Fang Chen, and Samir Mustapha. Robust dimensionality reduction and damage detection approaches in structural health monitoring. *Structural Health Monitoring*, 13(4):406–417, 2014.

[10] Ramin Ghiasi, Peyman Torkzadeh, and Mohammad Noori. A machine-learning approach for structural damage detection using least square support vector machine based on a new combinational kernel function. *Structural Health Monitoring*, 15(3):302–316, 2016.

[11] Jaime Vitola, Francesc Pozo, Diego A Tibaduiza, and Maribel Anaya. A sensor data fusion system based on k-nearest neighbor pattern classification for structural health monitoring applications. *Sensors*, 17(2):417, 2017.

[12] Giulio Mariniello, Tommaso Pastore, Costantino Menna, Paola Festa, and Domenico Asprone. Structural damage detection and localization using decision tree ensemble and vibration data. *Computer-Aided Civil and Infrastructure Engineering*, 36(9):1129–1149, 2021.

[13] Samir Khatir, Kevin Dekemele, Mia Loccufier, Tawfiq Khatir, and Magd Abdel Wahab. Crack identification method in beam-like structures using changes in experimentally measured frequencies and particle swarm optimization. *Comptes Rendus Mécanique*, 346(2):110–120, 2018.

[14] MR Nouri Shirazi, H Mollamahmoudi, and Seyed Mohammad Seyedpoor. Structural damage identification using an adaptive multi-stage optimization method based on a modified particle swarm algorithm. *Journal of Optimization Theory and Applications*, 160:1009–1019, 2014.

[15] Cheng Wang, Wei Guan, Jin Gou, Feng Hou, Junqing Bai, and Guirong Yan. Principal component analysis based three-dimensional operational modal analysis. *International Journal of Applied Electromagnetics and Mechanics*, 45(1-4):137–144, 2014.

[16] Viet Ha Nguyen and Jean-Claude Golinval. Fault detection based on kernel principal component analysis. *Engineering Structures*, 32(11):3683–3691, 2010.

[17] Yi Liu, Ke Li, Shimin Song, Yi Sun, Yong Huang, and Jun Wang. The research of spacecraft electrical characteristics identification and diagnosis using pca feature extraction. In *2014 12th International Conference on Signal Processing (ICSP)*, pages 1413–1417. IEEE, 2014.

[18] Nur Sila Gulgec, Martin Takáč, and Shamim N Pakzad. Convolutional neural network approach for robust structural damage detection and localization. *Journal of computing in civil engineering*, 33(3):04019005, 2019.

[19] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.

[20] Mohsen Azimi, Armin Dadras Eslamlou, and Gokhan Pekcan. Data-driven structural health monitoring and damage detection through deep learning: State-of-the-art review. *Sensors*, 20(10):2778, 2020.

[21] Young-Jin Cha, Wooram Choi, Gahyun Suh, Sadegh Mahmoudkhani, and Oral Büyüköztürk. Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types. *Computer-Aided Civil and Infrastructure Engineering*, 33(9):731–747, 2018.

[22] Xiangxiong Kong and Jian Li. Vision-based fatigue crack detection of steel structures using video feature tracking. *Computer-Aided Civil and Infrastructure Engineering*, 33(9):783–799, 2018.

[23] Deegan J Atha and Mohammad R Jahanshahi. Evaluation of deep learning approaches based on convolutional neural networks for corrosion detection. *Structural Health Monitoring*, 17(5):1110–1128, 2018.

[24] Cao Vu Dung et al. Autonomous concrete crack detection using deep fully convolutional neural network. *Automation in Construction*, 99:52–58, 2019.

[25] Felipe Mouallem de Assis and Guilherme Ferreira Gomes. Crack identification in laminated composites based on modal responses using metaheuristics, artificial neural networks and response surface method: a comparative study. *Archive of Applied Mechanics*, 91:4389–4408, 2021.

[26] Guilherme Antonio Oliver, Antonio Carlos Ancelotti, and Guilherme Ferreira Gomes. Neural network-based damage identification in composite laminated plates using frequency shifts. *Neural Computing and Applications*, 33:3183–3194, 2021.

[27] D Chetwynd, F Mustapha, K Worden, JA Rongong, SG Pierce, and JM Dulieu-Barton. Damage localisation in a stiffened composite panel. *Strain*, 44(4):298–307, 2008.

[28] Angelo De Fenza, Assunta Sorrentino, and Pasquale Vitiello. Application of artificial neural networks and probability ellipse methods for damage detection using lamb waves. *Composite Structures*, 133:390–403, 2015.

[29] Joseph Melville, K Supreet Alguri, Chris Deemer, and Joel B Harley. Structural damage detection using deep learning of ultrasonic guided waves. In *AIP conference proceedings*, volume 1949. AIP Publishing, 2018.

[30] Vincentius Ewald, Roger M Groves, and Rinze Benedictus. Deepshm: A deep learning approach for structural health monitoring based on guided lamb wave technique. In *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2019*, volume 10970, pages 84–99. SPIE, 2019.

[31] Heng Liu and Yunfeng Zhang. Deep learning based crack damage detection technique for thin plate structures using guided lamb wave signals. *Smart Materials and Structures*, 29(1):015032, 2019.

[32] Yasamin Keshmiri Esfandabadi, Maxime Bilodeau, Patrice Masson, and Luca De Marchi. Deep learning for enhancing wavefield image quality in fast non-contact inspections. *Structural Health Monitoring*, 19(4):1003–1016, 2020.

[33] Emmanuel J Candes, Justin K Romberg, and Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 59(8):1207–1223, 2006.

[34] Francesco Ciampa, Michele Meo, and Ettore Barbieri. Impact localization in composite structures of arbitrary cross section. *Structural Health Monitoring*, 11(6):643–655, 2012.

[35] A Nokhbatolfoghahai, HM Navazi, and RM Groves. Using the hybrid das-sr method for damage localization in composite plates. *Composite Structures*, 247:112420, 2020.

[36] Abdalraheem A Ijjeh, Saeed Ullah, and Pawel Kudela. Full wavefield processing by using fcn for delamination detection. *Mechanical Systems and Signal Processing*, 153:107537, 2021.

[37] P Kudela and W Ostachowicz. A multilayer delaminated composite beam and plate elements: reflections of lamb waves at delamination. *Mechanics of Advanced Materials and Structures*, 16(3):174–187, 2009.

[38] Evan Shelhamer, Jonathan Long, Trevor Darrell, et al. Fully convolutional networks for semantic segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 39(4):640–651, 2017.

[39] Giuseppe Bonaccorso. *Mastering Machine Learning Algorithms: Expert techniques for implementing popular machine learning algorithms, fine-tuning your models, and understanding how they work*. Packt Publishing Ltd, 2020.

[40] Jeroen Bertels, Tom Eelbode, Maxim Berman, Dirk Vandermeulen, Frederik Maes, Raf Bisschops, and Matthew Blaschko. Optimizing the dice score and jaccard index for medical image segmentation. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part II 22*, pages 92–100. Springer, 2019.

[41] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[42] Weihao Weng and Xin Zhu. Inet: convolutional networks for biomedical image segmentation. *Ieee Access*, 9:16591–16603, 2021.

[43] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436–444, 2015.

[44] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[45] Simon Jégou, Michal Drozdzal, David Vazquez, Adriana Romero, and Yoshua Bengio. The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 11–19, 2017.

[46] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia. Pyramid scene parsing network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2881–2890, 2017.

[47] Chao Peng, Xiangyu Zhang, Gang Yu, Guiming Luo, and Jian Sun. Large kernel matters–improve semantic segmentation by global convolutional network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4353–4361, 2017.

Chapter 4

Synthesis of semitransparent titania nanotubes formed out of Cu-doped Ti film and study on photoelectrochemical properties

Dujearic-Stephane Kouao

Institute of Fluid Flow Machinery, Polish Academy of Sciences,
Fiszera 14, 80-231 Gdansk, Poland

4.1 Introduction

The increasing interest for high-efficiency and low-cost photoelectrodes dedicated to the photoelectrochemical (PEC) systems has encouraged major research efforts on innovative nanostructures and semiconducting metal oxides [1]. The PEC process is describes as follows; electron-hole pairs are generated at the photoelectrode/ electrolyte interface upon exposure to light, which leads to oxidation or reduction reactions of the electrolyte species [2]. Among the most widely investigated materials, titania nanotube arrays (TNTs) still remain the preferred choice for practical applications in photocatalysis, solar cells, electrochromic devices, and sensors [3]. In order to improve the efficiency of the PEC systems, TNTs directly grown on transparent conducting oxides (TCO) has been explored. In this case, titanium films are first deposited onto TCO via physical coating methods including radio-frequency (RF), direct current (DC) magnetron sputtering, and pulsed laser deposition, and the obtained substrates are anodized to form the nanotubes onto the TCO. The advantage of such photoelectrodes lies to the fact that, TCO allow light to pass through the entire PCE cell structure [4]. Moreover, TCO namely fluorine-doped tin oxide (FTO) and TiO_2 have aligned energy levels of the conduction band edges, thus promoting the interfacial electron transfer from TiO_2 to FTO [5]. FTO layer acts as a collector of electrons photogenerated at TiO_2 and transfer them to the external circuit.

Although TNTs formed out on TCO are considered as promising photoelectrodes for PEC systems, the wide band gap of titania (TiO_2)-based materials (anatase TiO_2, ca. 3.2 eV) leads to poor absorption of light, and need to be tuned to further improve the performance of PEC devices [6, 7]. Doping with transition metal is one of the simplest and efficient method to enhance the light absorption properties of TiO_2 [7, 8]. A variety of dopant atoms such as Ag, Nb, Cu, Fe, Ru, Cr, Zn, Pd, W, Sn and Au have already been investigated to modify the intrinsic optical properties of the undoped semitransparent TNTs counterparts. Two methods are frequently used to modify the semitransparent TNTs structures. The most utilized method consists first, to anodize Ti films coated

onto TCO, followed by a thermal treatment in an inert atmosphere to deposit the metal nanoparticles onto the nanotubes [9–11]. In another method, composite metallic films (Ti-metal film) were deposited onto TCO by simultaneous co-sputtering technique from two targets namely titanium and metal sources, thereafter the obtained films are anodized to achieve semitransparent metal-doped TNTs [12–16]. Recently, Bjelajac et al. [17] synthesized transparent Sn-doped TNTs by annealing the as-anodized semitransparent TNTs formed out onto FTO between 500 and 550°C, which allowed the dynamic diffusion of Sn from FTO to the nanotube walls. Both the anodization procedures and targets used for the deposition of Ti or Ti-metal films onto the TCO are presented in the Tab. 4.1.

The doping atoms generate inter-band levels in the TiO_2 band gap, thus facilitating the excitation of the electrons from the valance band to the conduction band [17]. In the case of semitransparent titania, for example Mor et al. [13] studied the effect of iron doping on TNTs photoelectrochemical properties. They synthesized the photoanode by a three-steps process. First, Ti-Fe films were deposited on FTO by simultaneous co-sputtering technique using both Ti- and Fe targets. Then the obtained substrates were anodized in ethylene glycol (EG) based electrolyte containing 0.3 wt% ammonium fluoride (NH_4F) and 2 vol% deionized water (H_2O). Last, the anodized materials were annealed at 500°C for 2h. The prepared Fe-doped TNTs photoelectrode exhibited a shift of the absorption edge from 380 nm (undoped TNTs) to 570 nm (Fe-doped TNTs), indicating an improvement in the light absorption properties of TNTs by Fe-doping. Moreover, under illumination (100 mW/cm^2) the Fe-doped TNTs electrode achieved a photocurrent density of 2 mA/cm^2 at 0.65 V. Similar results also have been reported by Gui et al. [15], when investigating the impact of tungsten-doping on the optical properties of TNTs. W-Ti films sputter coated onto FTO from two targets composed of Ti and W, were subsequently anodized in EG electrolyte containing 0.3 M NH_4F and 4 vol% H_2O.

Tab. 4.1: Anodization procedures and targets used for the deposition of Ti or Ti-metal films onto the TCO.

Doping atom	**Substrate**	**Targets**	**Electrolyte**	**Reference**
Cu	FTO	Ti/Cu	EG + 0.25 wt% NH_4F + 1 vol% H_2O	[14]
Fe	FTO	Ti/Fe	EG + 0.3 wt% NH_4F + 2 vol% H_2O	[13]
Nb	Glass	Ti/Nb	CH_3COOH/CH_3COONa buffer solution	[16]
W	FTO	Ti/W	EG + 0.3 wt% NH_4F + 4 vol% H_2O	[15]
Ag	ITO	Ti	EG + 0.3 M NH_4F + 3 vol% H_2O	[12]
Au	FTO	Ti	EG + 0.3 M NH_4F + 2 vol% H_2O	[9]
Cr	FTO	Ti	EG + 0.4 wt% NH_4F + 5 wt% H_2O	[11]
Au	FTO	Ti	EG + 0.3 wt% NH_4F + 4 vol% H_2O	[10]
Ru	FTO	Ti	EG + 0.3 wt% NH_4F + 4 vol% H_2O	[10]
ZnPd	FTO	Ti	EG + 0.3 wt% NH_4F + 4 vol% H_2O	[10]
Sn	FTO	Ti	EG + 0.3 wt% NH_4F + 2 vol% H_2O	[17]

The prepared material annealed at 450°C for 3 h exhibits a high light transmittance of 71.5% at wavelength of 632.8 nm. Among all the transitional metals, copper (Cu) is high abundance with low cost availability. Moreover, since the radius of Cu^{+2} is close to that of Ti^{+4}, Cu can be easily incorporated into TiO_2 lattice [8]. In order to understand the influence of copper nanoparticles on titania photoelectrochemical activity, Mor et al. [14] investigated the photoresponse properties of Cu-doped TiO_2 based material. They deposited Cu−Ti films on FTO coated glass substrate by

co-sputtering from titanium and copper targets. During the deposition the substrate temperature was maintained at 300°C, to improve the adhesion of the sputtered films. The deposited films were anodized in EG bath containing 0.25 wt% NH_4F and 1 vol% H_2O at 30 V, and then annealed at 450°C. The photoelectrochemical activities of the prepared Cu-doped TNTs electrodes were studied in 0.1 M disodium hydrogen phosphate (Na_2HPO_4) electrolyte. Under illumination the Cu-doped TNTs based photoanodes demonstrated a photocurrent density of 0.25 mA/cm^2 from the analysis of the chronoamperometry curves, with a photoconversion efficiency of 0.30%. Although the above reported works are encouraging, and present metal-doped TNTs as a promising photoelectrode for photoelectrochemical systems, no work was conducted to investigate the influence of the nanotubes architecture on the photoresponse of metal-doped TNTs based materials.
This chapter presents the electrochemical and photoelectrochemical properties of aligned and spaced TNTs doped with different amounts of copper. Close-packed and loose-packed Cu-doped TNTs were synthesized via anodization of TiCu films sputter coated onto indium tin oxide (ITO).

4.2 Experimental

Ethylene glycol, diethylene glycol, ammonium fluoride, phosphoric acid, hydrofluoric acid were brought from Sigma-Aldrich Chemical Co. All chemical reagents were used without any purification. Ti or Cu-Ti films coated onto the indium tin oxide (ITO) substrates were received from Czech Republic: University of South Bohemia in Ceske Budejovice, Charles University in Prague within the cooperation established for the realization of CEUS-UNISONO project 2020/02/Y/ST8/00030. The anodization of the Ti or Cu-Ti films was carried out in a two-electrode setup. Ti or Cu-Ti films deposited onto ITO act as anode, whereas a platinum mesh served as cathode. Aligned titania nanotubes were obtained by anodizing the sputtered Ti or Cu-doped Ti films (i.e. 0.8, 2.5 and 4.6% of copper content). The deposition of Cu was together with Ti using both Cu and Ti plates as a target. The deposition of those layers was

realized by the scientific partner form Czech Republic via magnetron sputtering. For all samples i.e. pure Ti or Cu doped Ti films the thickness of the deposited Ti or Cu-Ti layer was 1.2 μm. The received Cu-Ti films were anodized in ethylene glycol based electrolyte containing 1M H_3PO_4 and 0.27 M NH_4F and 5 vol% H_2O at 23°C. The electrodes were placed 2 cm apart and the voltage was ramped up to 60 V at a sweep rate of 0.1 V/s, and then held for 3h by using a voltage generator (MPS-600-5 L-2), and the anodized electrodes were labelled aTNT, aTNT-0.8Cu, aTNT-2.5Cu and aTNT-4.6Cu for the 0, 0.8, 2.5 and 4.6% copper content in the Ti layer, respectively. To achieve spaced titania nanotubes, the anodization is carried out in diethylene glycol bath mixed with 0.15 M NH_4F, 0.5 M HF and 7 vol% H_2O at 40°C, the voltage was ramped up to 40 V at a sweep rate of 0.1 V/s, and then held for 30 min, and the obtained electrodes were labelled sTNT and sTNT-0.8Cu for the undoped and 0.8% copper content, respectively.
The prepared materials were annealed at 450°C for 1h in tubular furnace (Nabertherm), and their surface morphology was investigated using field emission scanning electron microscopy, FEI Quanta FEG 250 (FESEM). The geometrical features i. e. average internal diameter, tube-to-tube spacing and the nanotube density (number of nanotubes per unit area) were measured using Image J software. Raman spectra were recorded using a Raman spectrometer (Renishaw), equipped with a confocal microscope with an argon ion laser operating at 514 nm at a power level of 10 mW. The spectra were measured in the wave number range from 100 to 3200 cm^{-1}.
Diffusive reflectance spectra of the annealed samples were registered using Lambda 35 UV-vis spectrophotometer (Perkin Elmer) in a range of $200 - 600$ nm, at a scanning speed of 120 nm/min. The bandgap energy was determined by using Kubelka-Munk function, F(R). The relationship between the reflectance and the absorption coefficient is given as follows, $F(R) = (1\text{-}R)^2/(2R) = k/s$, where R is the reflectance, k is the absorption coefficient and s the scattering coefficient. Replacing absorption coefficient α by F(R) in the Tauc equation $((\alpha h\nu)^{1/n} = B(h\nu\text{-}Eg))$, results in the form: $(F(R)h\nu)^{1/n} = B(h\nu\text{-}Eg)$. Eg is the band gap energy, B is a constant, and h is the Planck constant, ν is the photon frequency, the n factor depends

on the type of the electron transition. From literature it is well known that for a semiconductor with a direct band gap transition n takes the value 0.5 and for an indirect band gap transition, n is equal to 2. Singh et al. [18] demonstrated that titania based materials can have either an direct or indirect band gap structure depending on the temperature at which the materials are synthesized. They showed that titania based materials prepared at temperature around 400°C (anatase) have an indirect band gap, while materials synthesized at a higher temperature of 900°C (rutile), have a direct band gap structure [18, 19]. Therefore, in this study the value of factor n will be chosen equal 0.5 for the determination of the band gaps of the annealed materials at 450°C.
The electrochemical and photoelectrochemical activities were studied in three-electrode setup. Fig. 4.1 describes the photoelectrochemical setup. The measurement setup consists of a platinium (Pt) counter electrode, Ag/AgCl (0.1 M) reference electrode, and the aligned or spaced Cu-doped TNTs photoanode as a working electrode immersed in 0.1 M NaOH. Prior to each measurement, argon gas was bubbled through the solution for 30 min to deaerate the electrolyte and during the experiment argon flow was kept above the electrolyte. For comparison pure aligned and spaced TNTs had also been investigated as working electrode. The photoresponse tests were performed by measuring the photocurrent under chopped light irradiation (light/dark every 5 s) at a fixed external bias of 0.5 V versus Ag/AgCl. Cyclic voltammetry (CV), electrochemical impedance spectroscopy (EIS), linear sweep voltammetry (LSV) measurements were conducted at room temperature and controlled by Autolab PGSTAT302N potentiostat-galvanostat system (Methrom, Autolab).

The CV was performed within potential range of -1 to $+1$ V, at a sweep rate of 50 mV/s. The EIS measurement was carried out either in dark or under UV-vis illumination, by applying AC voltage amplitude of 10 mV in the 0. 1-20 kHz frequency range. In order to perform the Mott-Schottky analysis, electrochemical impedance spectra were recorded at a constant frequency of 1 kHz, in the potential range of + 0.1 to -1.2 V. LSV was carried out by using

Tab. 4.2: Anodization procedures and targets used for the deposition of Ti or Ti-metal films onto the TCO.

Sample label	**Cu content in deposited Ti layer (%)**	**Morphology: spaced or aligned**
sTNT	0	spaced
sTNT−0.8Cu	0.8	spaced
aTNT	0	aligned
aTNT−0.8Cu	0.8	aligned
aTNT−2.5Cu	2.5	aligned
aTNT−4.6Cu	4.6	aligned

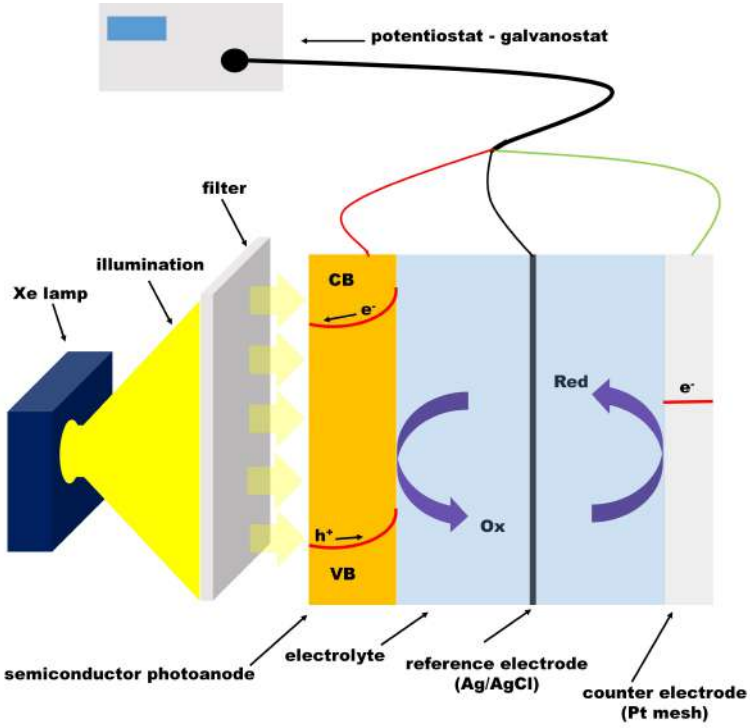


Fig. 4.1: Schematic illustration of the photoelectrochemical measurement setup.

an automatic chopped simulated AM1.5G illumination (light/dark every 5 s) in either UV−vis or visible light irradiation from −0.5 to +1 V at a scan rate of 10 mV/s. For the measurement carried out in visible region, an additional cut-off optical filter (λ>400 nm) was place between the light source and the three-electrode setup, as shown in Fig. 4.1. CA was recorded for 300 seconds under chopped illumination in UV-vis light (light/dark every 5 s). The description of the samples according to their morphology and composition of the anodized TiCu layer is provided in the Tab. 4.2. Depending on the distance between nanotubes, verified on the basis of SEM inspection, the spaced or aligned term is used.

4.3 Results and discussion

4.3.1 Morphology and microstructural properties

For all investigated samples, the morphology consists of highly ordered, top-end-open, and uniform nanotube arrays.

The SEM images of the spaced and aligned-TNTs materials are shown in the Fig. 4.2 and Fig. 4.3. The geometrical characteristics of the materials as prepared are summarized in Tab. 4.3.

Tab. 4.3: Averaged inner diameters, tube-to-tube spacing and nanotube density of as-prepared TNTs.

Sample	**Average inner diameter (nm)**	**Average tube-to-tube spacing (nm)**	**Nanotube density ($/\mu m^2$)**
sTNT	64 ± 7	88 ± 10	80 ± 6
sTNT-0. 8Cu	78 ± 11	177 ± 20	35 ± 4
aTNT	27 ± 3	no spacing	142 ± 11
aTNT-0. 8Cu	38 ± 5	no spacing	140 ± 12
aTNT-2. 5Cu	45 ± 5	no spacing	136 ± 10
aTNT-4. 6Cu	42 ± 5	no spacing	134 ± 10

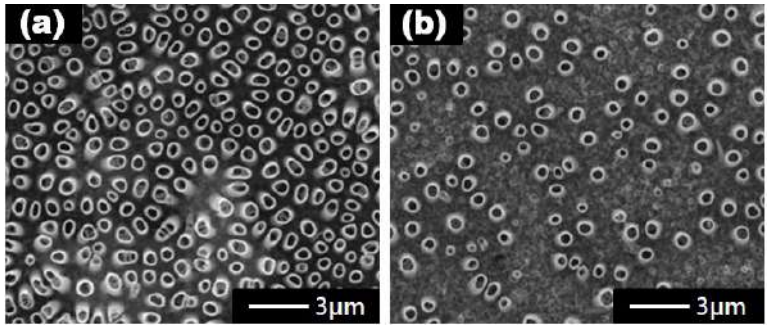


Fig. 4.2: SEM images of (a) sTNT, (b) sTNT-0.8Cu.

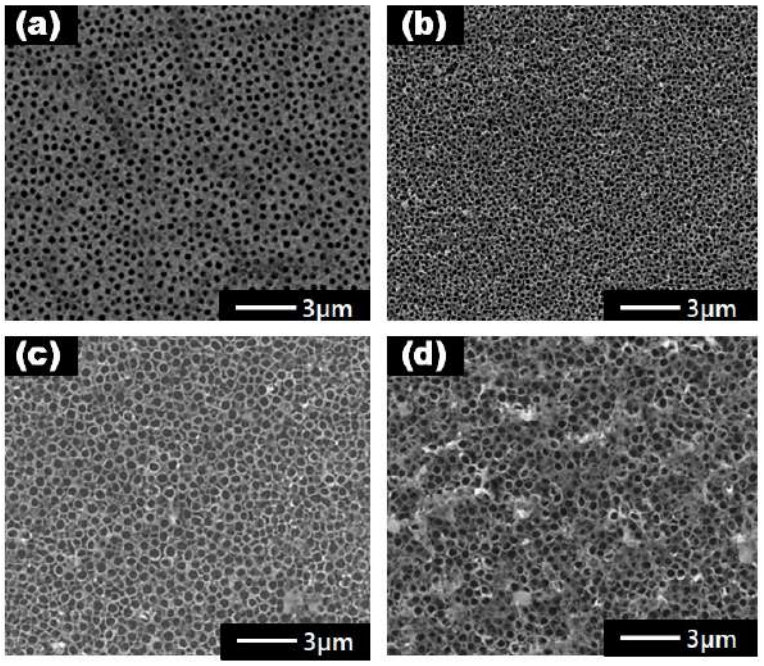


Fig. 4.3: SEM images of (a) aTNT, (b) aTNT-0.8Cu (c) aTNT-2.5Cu, (d) aTNT-4.6Cu.

For the aligned structure, the average inner diameter increased with the increase of copper content, from 27 nm ± 3 nm for aTNT (pure material) to 42 nm ± 5 nm for aTNT-4.6Cu. For both architectures

(aligned or spaced-TNTs) the copper doping has led to decreased the nanotube density (population of nanotubes per unit area) of the materials comparing to the layer when pure Ti has been anodized, as shown in Tab. 4.3. Overall, significant changes in the geometric characteristics of TNTs such as the tube diameter, and the tube-to-tube spacing were observed after Cu-doping. However, no deterioration in the tubular structures was detected with a copper content in the range 0.8 to 4.6%.
The crystalline phase of pure and Cu-doped TNTs was inspected by Raman spectroscopy and the results are shown in Fig. 4.4.

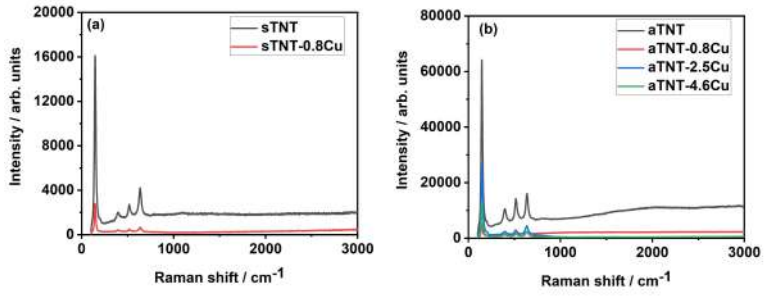


Fig. 4.4: Raman spectra of (a) sTNT and sTNT−0.8Cu, (b) aTNT, aTNT−0.8Cu, aTNT−2.5Cu, aTNT−4.6Cu.

The Raman spectra of pure aligned and spaced TNTs show the characteristic peaks of anatase at 143, 198, 398, 515, and 640 cm^{-1}, which are assigned to the active vibrational modes of Eg, Eg, B1g, A1g, and Eg respectively [20]. There are no additional peaks related to any copper species. It can indicate that so low amount of copper species comparing to titania does not provide any visible signal or the dopant cations are inserted into the substitutional positions of titania crystal lattice [21]. However, a decrease in the intensity of Raman peaks was observed for all Cu-doped TNTs materials presumably due to the decrease in the material crystallinity [22].

4.3.2 Optical properties

To assess the effect of copper doping on the optical properties of both architecture i.e. aligned and spaced TNTs, both the absorption and UV-vis diffuse reflectance spectra of the materials were presented in the Fig. 4.5. In the UV region (below 400 nm), there is only a slight change in the absorbance of both spaced-TNTs based materials, namely 90% for sTNT and 88% for sTNTs-0. 8Cu, whereas an increase of the absorbance from 76% to 87% was observed with the rise of the copper content in the deposited film up to 2.5% for the aligned architecture. In the visible range i.e. above 400 nm, there is no significant change in the absorbance of all TNTs layers. Moreover, one can observe that all investigated samples mainly absorb light in UV region of the spectrum. This indicated that the change in the materials morphology obtained after the Cu-doping did not caused the absorption shift towards the visible region. Moreover, the dopant did not lead to the reduced band gap of both pure materials, as shown in the Tab. 4.4. For the pure aligned and spaced TNTs, the estimated values of the band gap are 2.95 and 2.90 eV respectively. This decrease in the band gap as compared to the anatase TiO_2 (3.20 eV) is due to geometrical characteristics of the prepared titania nanotues [23].

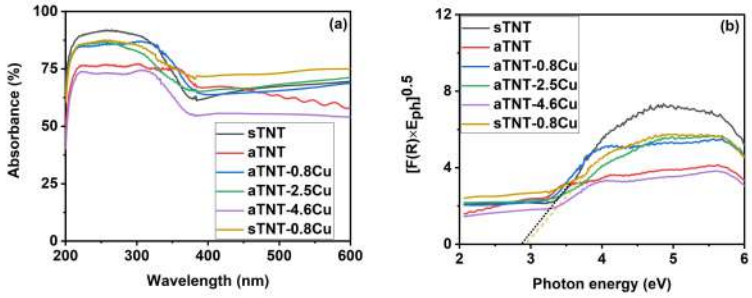


Fig. 4.5: (a) Absorbance spectra, and (b) UV−vis diffuse reflectance spectra of sTNT and sTNT−0.8Cu, aTNT, aTNT−0.8Cu, aTNT−2.5Cu, aTNT−4.6Cu.

Tab. 4.4: Band-gap energy values of the as-prepared materials.

Sample	band gap (eV)
sTNT	2.90
sTNT−0.8Cu	2.94
aTNT	2.95
aTNT−0.8Cu	3.09
aTNT−2.5Cu	3.11
aTNT−406Cu	2.98

4.3.3 Electrochemical and photoelectrochemical activity

In order to verify the stability of the material as an electrode immersed in the solution, some preliminary measurements were conducted. First the cyclic voltammetry technique was used to verify the potential window in which the material is electrochemically stable and if some oxidation/reduction peaks arises. The voltammograms were collected for samples immersed in 0.1 M NaOH at 50mV/s. The prepared materials were studied in a basic electrolyte due to their copper content. Indeed, according to the Pourbaix diagram for copper at 25°C reported by Gambino et al. [24] the copper species are thermodynamically more stable and do not dissolve in the basic electrolyte. For the spaced TNTs materials the current density is significantly reduced in the cathodic part after the Cu-doping, as shown in Fig. 4.6.

The drop in the current density can be ascribed to the nanotube density (number of nanotubes per unit area) by the Cu-doping, as already discussed above. Indeed, it is well known that the nanotubes act as electric channels for electrons transportation [25]. Therefore, the decrease in the nanotube density will drop the current density of the material. In the case of the aligned TNTs, the current density of the sample aTNT−2.5Cu is negligible (Fig. 4.5.b). The voltammogram areas of the samples aTNT, aTNT−0.8Cu, aTNT−4.6Cu are very similar. Furthermore, no reversible redox peaks were detected on these CV curves indicating that the Cu

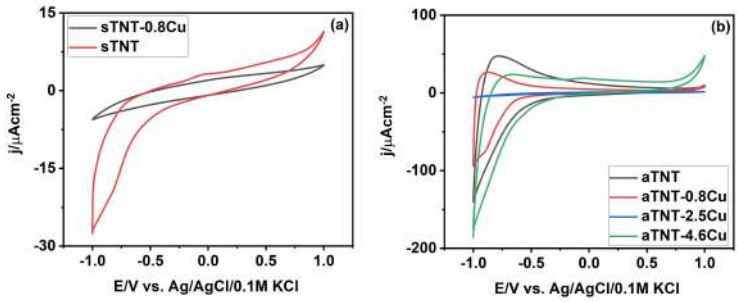


Fig. 4.6: Cyclic voltammetry curves (scan at 50mV/s) of (a) sTNT and sTNT-0.8Cu, (b) aTNT, aTNT-0.8Cu, aTNT-2.5Cu, aTNT-4.6Cu.

doping did not significantly changes the electrochemical activity of the aligned TNTs [26]. However, the cathodic current density increase with the rise of the copper content in deposited Ti layer. This increase of current value is related with the reduction of Ti^{+4} to Ti^{+3} [27].

The electrochemical impedance spectroscopy (EIS) spectra of the materials recorded under UV-vis irradiation are not clearly distinguishable from the spectra obtained in the dark, which can be seen in Fig. 4.7. This indicates that the materials have a poor photoresponse. The spectra of spaced TNTs shifted to higher real resistance values after Cu doping. This may be due to both the drop in electrical conductivity of the material by the copper species and the high electrode/electrolyte interfacial resistance.

Fig. 4.7.b shows the impedance spectra of the aligned TNTs. Two parts can be distinguished in the Nyquist plots, the semicircle impedance loops obtained at high frequency region, showed in the insert in the Fig. 4.7.b, is resulting from the charge-transfer resistance at the electrode/electrolyte interface and since, no redox peaks were observed in the cyclic voltammograms of the prepared material the sloping lines in the lower frequency range may be resulted from the charges transport through the nanotube walls associted to the Warburg open-circuit element (Wo) [28]. Both two photoelec-

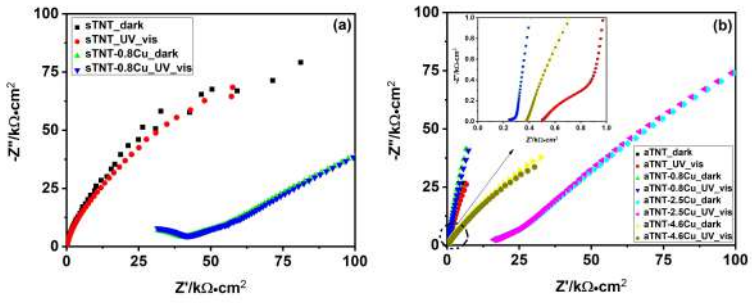


Fig. 4.7: Nyquist plots of EIS measurements of (a) sTNT and sTNT−0.8Cu, (b) aTNT, aTNT−0.8Cu, aTNT−2.5Cu, aTNT−4.6Cu in a frequency range of 0. 1−20 kHz.

trodes i.e. 0.8 and 4.6 % of copper contents have smaller semicircles as compared to the undoped aligned TNTs material. This indicates that Cu-doping promote the charge transfer at electrode/electrolyte interface, by reducing the interfacial resistance. Contrary to the observations made in the case of spaced TNTs materials, a Cu content of 0.8 or 4.6% in the aligned TNTs matrix shifts the impedance spectra to lower real resistance values, suggesting an improvement in the conductivity of the material. However, no improvement was observed in the electrical properties of the aligned TNTs material doped with 2.5% of copper. According to Ma et al. [29] titania crystalline structures are very sensitive to copper concentration. Hence, the drop in the electrical conductivity observed for the Cu-doped aligned TNTs (2.5 %) can be due to a significant lattice distortion in the aligned TNTs microstructure. The good electrical properties were obtained with the highest value of copper content in the Ti:Cu film (4.6 %). This corroborates with the results reported by Guo et al. [7]. They also study the effect of Cu content on the photoelectric properties of TNTs, and they showed that the electrical properties of TNTs can be improved by increasing the copper content.

Fig. 4.8 shows linear sweep voltammetry (LSV) curves of the prepared materials under chopped light irradiation (light/dark every 5 s). In general, the current density increases under light illumina-

tion and returned to the background value as the light turned off. The spaced TNTs material demonstrated good photoresponse under UV-vis light. However, its photoelectrochemecal activity dropped dramatically under visible light. When doped with a copper content of 0.8%, the material loses its photosensitivity upon exposure to both UV-vis light irradiation. The aligned TNTs doped with 4.6% of Cu exhibited the best photoresponse under UV-vis light. However, similarly to the spaced TNTs materials, the photoresponses of the undoped and Cu-doped aligned TNTs (0.8, 2.5 and 4.6%) are almost negligible under visible light as shown in the Fig. 4.8.d. The presence of the dopant metal do not provide improved photoresponse of aligned and spaced TNTs in the visible region.

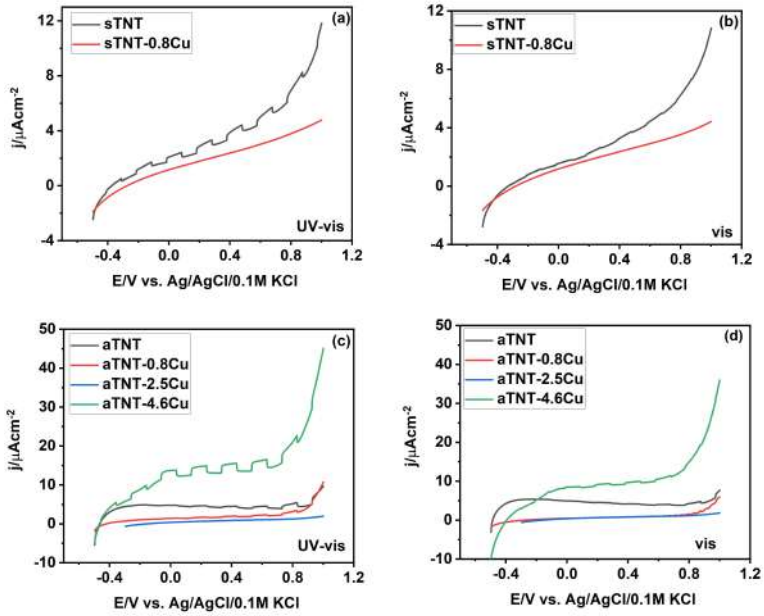


Fig. 4.8: Linear sweep voltammograms of sTNT and sTNT-0 8Cu under chopped (a) UV-vis illumination, (b) visible illumination and aTNT-0.8Cu, aTNT-2.5Cu, aTNT-4.6Cu under chopped (c) UV-vis illumination, (d) visible illumination, (light/dark every 5 s).

Chronoamperometry was carried out to investigate the reproducibility and stability of the prepared photoelectrodes. The photocurrent was recorded as a function of time by switching on and off the light simulator (light/dark every 5 s). For the undoped space TNTs photoelectrode the photocurrent density was increased to 1.85 μA/cm^2 from about 0.9 μA/cm^2 when the light was turned on. The current density was stable during the measurement, as shown on the Fig. 4.9.a. For the aligned TNTs materials the improve-

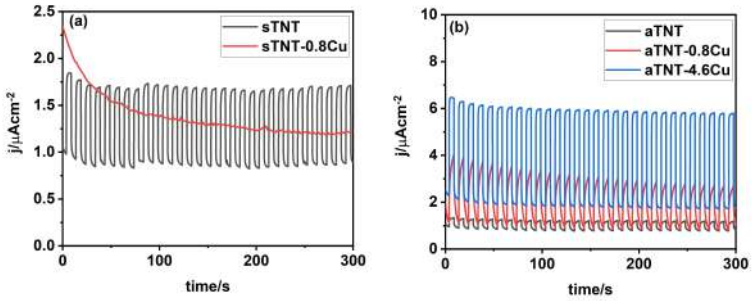


Fig. 4.9: Transient photocurrent curves (light/dark every 5 s) of (a) sTNT and sTNT−0 8Cu, (b) aTNT, aTNT−0.8Cu, aTNT−2.5Cu, aTNT−4.6Cu.

ment in the current density (Δj= light current density - dark current density) was strongly dependent on the amount of the dopant, Δj = 4.3, 3.05 and 0.5 μA/cm^2 for 4.6, 0.8% copper content and the undoped aligned TNTs respectively. For the aligned TNTs the photoresponse increased linearly with the rise of copper content. According to Mohajernia et al. [30] the enhancement of the beneficial effect due to the copper presence may be justified by the light-induced oxidation of Cu species, namely Cu^+ to Cu^{2+} and Cu^{2+} to Cu^{3+} in the titanium oxide matrix. Although the undoped and 4.6% copper doped aligned TNTs materials exhibited very good stability, the photocurrent density of the sample doped with 0.8% Cu was sharply decreased from Δj = 3.05 μA/cm^2 to 1.69 μA/cm^2 after 300 s, indicating a poor stability of the material. Mott−Schottky plots analysis were provided to conduct a qualitative comparison of donor density present within the bandgap of

the synthesis materials. Mott-Schottky relation for n-type semiconductor is given as: $C^{-2} = (2/q\varepsilon\varepsilon_0 N_D)(E\text{-}E_{fb}\text{-}kT/q)$, C is the space charge (depletion layer) capacitance, ε is permittivity of the anatase titania ($\varepsilon = 38$) [31], ε_0 (8. 85×10^{-14} F/cm) is the permittivity of free space, q is the electron charge, N_D is the donor density, E is the applied potential, E_{fb} is the flat band potential, k is the Boltzmann constant, and T is the temperature. The capacitance of the space charge layer can be calculated using the following equation: $C = -1/2\pi f Z_{im}$, f is the frequency of the AC signal, and Z_{im} is the imaginary part of impedance. The flat band potential values (E_{fb}) are determined by the extrapolation of the linear regions of the curve C^{-2} vs. E, as shown in the Fig. 4.10. The donor density N_D can be derived by the slope of the Mott−Schottky plot and is calculated via the equation: $N_D = (2/q\varepsilon\varepsilon_0)(dE/d(C^{-2}))$ [12].

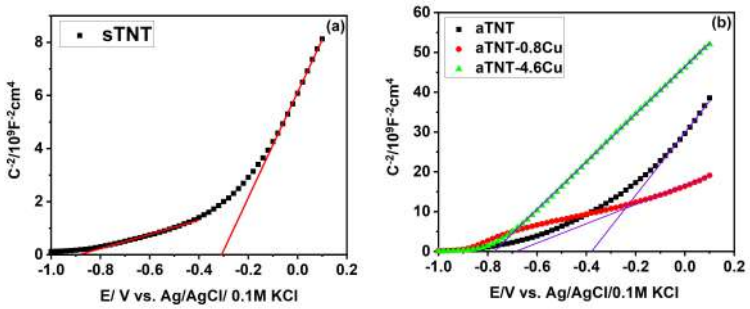


Fig. 4.10: Mott−Schottky curves of (a) sTNT, (b) aTNT, aTNT−0.8Cu, aTNT−4.6Cu.

The positive slopes obtained for all samples in the linear regions of the Mott-Schottky plots indicate that all investigated materials exhibit n-type semiconductor behavior (Fig. 4.10). The calculated values of the flat band potential at the electrode/electrolyte junction and donor density are summarized in the Tab. 4.5.

Basically, the flat band potential is a qualitative measure of the depth of the Fermi level in a band gap [32]. At this potential, band bending due to electron depletion in the semiconductor vanishes. The two linear regions observed on the Mott-Schottky plot

Tab. 4.5: Flat band potential, donor concentration of the samples.

Sample	$\mathbf{E}_{fb}$**(V vs. Ag/AgCl /0. 1M KCl)**	$\mathbf{N}_D$**(/cm^3)**
sTNT	−0.35 and −0.90	1.9×10^{20} and $1.\ 3\times10^{21}$
aTNT	−0.39	4.6×10^{19}
aTNT−0.8Cu	−0.70	1.6×10^{20}
aTNT−4.6Cu	−0.80	6.2×10^{19}

of sample sTNT indicate the presence of two donor states present within the band gap [33, 34]. One can observe in the Fig. 4.10.b that the increase of the copper amount leaded to a negative shift of the flat band potential of the aligned TNTs material. In general, the origin of this shift is not only due to the presence of dopant atoms, but also the structural defect in the materials i.e. oxygen vacancies in the semiconductor [35]. According to Ikram et al. [32] the shift of the flat band potential towards more negative potential promotes the electron transfer at the electrode/electrolyte interface due the decrease in the free energy change. The results obtained is this work is in agreement with the observation of Ikram et al. [32] Indeed, for the aligned TNTs architecture the sample aTNT−4.6Cu has the more negative flat band potential, with the best photoresponse (according to the chronoamperometry analysis). In addition, the value of donor densities increased from $4.\ 6\times10^{19}$/cm^3 to $1.\ 6\times10^{20}$/cm^3 for 0.8% copper content and from 4 6×10^{19}/cm^3 to 6.2×10^{19} /cm^3 for 4.6% copper content in the initial Ti layer. Similar increment in the donor density after the introduction of copper species in TNTs matrix has also been reported previously by Wu et al. [36]. At lower Cu content (0.8%) the increase of the donor concentration as compared to the undoped material aTNTs can be justified by the fact that the introduction of dopant atoms in the titania lattice generates oxide ion vacancies that are regarded as electron donors [37, 38]. However, it is well known that since the valence of Cu ions (Cu^{2+} or Cu^{+}) is lower than Ti^{4+}, the Cu dopants incorporating Ti sites into titania are considered as point defects exhibiting acceptor nature [29]. Thus the decrease of the

donor concentration observed at the relatively high copper content 4.6% maybe due to fact that vacancies have being occupied by copper acceptors [37].

4.4 Conclusion

Aligned and spaced TNTs architectures were synthesized through anodization of copper doped Ti films. The influence of copper content in metallic films on the final microstructure of the prepared materials was investigated using field emission scanning electron microscopy. The results showed significant change in the nanotubes geometry with the increase of the copper content in the materials. For the aligned-TNTs materials, the inner diameter was increased from 27 nm ± 3 nm for aTNT to 42 nm ± 5 nm for aTNT-4. 6Cu. In case of the spaced-TNTs, the average intertube spacing was increased from 88 nm ± 10 nm for sTNT to 177 nm ± 20 nm for 0.8% doped TNTs. For both architectures: aligned and spaced-TNTs, the number of nanotubes per unit area decreased with the increase of copper content. However, no distortion in the synthesized TNTs morphological features was detected with the rise of copper content up to 4.6%. Moreover, Raman study has revealed that despite Cu is present in titania layer, anatase phase is preserved after calcination process. However, the optical properties of Cu-doped TNTs must be further improved to broaden the light absorption window of the materials towards the visible region. The investigation of the photoelectrochemical activity showed that sTNT-0.8Cu sample is not photoactive under both UV and visible irradiation. For the aligned-TNTs, the photocurrent density increased linearly with the rise of copper content in initial Ti:Cu film from $\Delta j = 0.5\ \mu A/cm^2$ for pure aTNTs, to $\Delta j = 4.3\ \mu A/cm^2$ for 4.6% copper content. From the chronoamperometry curve one can point out that a copper content of 4.6% in deposited Ti film can affect the improvement the photosensitivity of the aligned TNTs architecture. However, the performance of the fabricated material need to be further developed for the application as an highly efficient photoelectrode.

Acknowledgment

This work received financial support from the Polish National Science Centre: CEUS-UNISONO 2020/02/Y/ST8/00030.

Bibliography

[1] Hong Zhu, Jie Tao, and Xiang Dong. Preparation and photoelectrochemical activity of Cr-doped TiO_2 nanorods with nanocavities. *The Journal of Physical Chemistry C*, 114(7):2873–2879, 2010.

[2] Bard Allen. Photoelectrochemistry. *Science*, 207(4427):139–144, 1980.

[3] Fabrication and characterization of copper doped Ti.

[4] Wenrui Dai, Ying Tao, Hangjun Zou, Shuning Xiao, Guisheng Li, Dieqing Zhang, and Hexing Li. Gas-phase photoelectrocatalytic oxidation of NO via TiO_2 nanorod array/FTO photoanodes. *Environmental Science & Technology*, 54(9):5902–5912, 2020.

[5] Nibedita Paul, Debarun Dhar Purkayastha, M Ghanashyam Krishna, and sang Talinung. TiO_2/SnO_2 and SnO_2/TiO_2 heterostructures as photocatalysts for degradation of stearic acid and methylene blue under uv irradiation. *Superlattices and Microstructures*, 129:105–114, 2019.

[6] Salih Veziroglu, Jin Hwang, Jonas Drewes, Igor Barg, Shondo Josiah, Strunskus Thomas, Oleksandr Polonskyi, Franz Faupel, and Aktas Oral Cenka. PdO nanoparticles decorated TiO_2 film with enhanced photocatalytic and self-cleaning properties. *Materials Today Chemistry*, 16:100251, 2020.

[7] Fei Guo, Jiameng Liu, Wanggang Zhang, Zhuobin Yu, Yiming Liu, and Wei Liang. Synthesis of Cu, N-doped TiO_2 nanotube by a novel magnetron sputtering method and its photoelectric property. *Vacuum*, 165:223–231, 2019.

[8] Mohamed Mahmoud, Enas Ahmed, Emad Abdelghani, and Nasser Barakat. Influence of Mn, Cu, and Cd -doping for titanium oxide nanotubes on the photocatalytic activity toward water splitting under visible light irradiation. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 554:100–109, 2018.

[9] Pingping Li, Jinshu Wang, Guannan Zu, Peng Jiao, Shaojing Liu, Yunfei Yang, Shuqun Chen, and Hongyi Li. Synergetic catalytic properties of gold nanoparticles planted on transparent titanium dioxide nanotube array bed. *Materials Chemistry and Physics*, 217:437–444, 2018.

[10] Piyush Kar, Samira Farsinezhad, Najia Mahdi, Yun Zhang, Uchenna Obuekwe, Himani Sharma, Jing Shen, Natalia Semagina, and Karthik Shankar. Enhanced CH_4 yield by photocatalytic CO_2 reduction using TiO_2 nanotube arrays grafted with Au, Ru, and ZnPd nanoparticles. *Nano Research*, 9(11):3478–3493, 2016.

[11] Burcu Bozkurt Çırak. Fabrication and characterization of transparent Cr-decorated TiO_2 nanoporous electrode for enhanced photo-electrocatalytic performance. *Vacuum*, 177:109375, 2020.

[12] Shiqi Wang, Zeling Zhang, Wenyi Huo, Kui Zhu, Xuhai Zhang, Xuefeng Zhou, Feng Fang, Zonghan Xie, and Jianqing Jiang. Preferentially oriented Ag-TiO_2 nanotube array film: An efficient visible-light-driven photocatalyst. *Journal of Hazardous Materials*, 399:123016, 2020.

[13] Gopal K Mor, Haripriya E Prakasam, Oomman K Varghese, Karthik Shankar, and Craig A Grimes. Vertically oriented Ti-Fe-O nanotube array films: toward a useful material architec-

ture for solar spectrum water photoelectrolysis. *Nano letters*, 7(8):2356–2364, 2007.

[14] Gopal K Mor, Oomman K Varghese, Rudeger HT Wilke, Sanjeev Sharma, Karthik Shankar, Thomas J Latempa, Kyoung-Shin Choi, and Craig A Grimes. p-type Cu- Ti-O nanotube arrays and their use in self-biased heterojunction photoelectrochemical diodes for hydrogen generation. *Nano Letters*, 8(7):1906–1911, 2008.

[15] Yang Gui and Daniel Blackwood. Honey-comb structured WO_3/ TiO_2 thin films with improved electrochromic properties. *Journal of The Electrochemical Society*, 162(10):E205, 2015.

[16] Petra Bleckenwegner, Cezarina Cela Mardare, Christoph Cobet, Jan Philipp Kollender, Achim Walter Hassel, and Andrei Ionut Mardare. Compositionally dependent nonlinear optical bandgap behavior of mixed anodic oxides in niobium-titanium system. *ACS Combinatorial Science*, 19(2):121–129, 2017.

[17] Andjelika Bjelajac, Rada Petrović, Jelena Vujancevic, Katerina Veltruska, Vladimir Matolin, Zdravko Siketic, George Provatas, Milko Jaksic, George E Stan, Gabriel Socol, et al. Sn-doped TiO_2 nanotubular thin film for photocatalytic degradation of methyl orange dye. *Journal of Physics and Chemistry of Solids*, 147:109609, 2020.

[18] Mrityunjay Kumar Singh and Mohan Singh Mehata. Phase-dependent optical and photocatalytic performance of synthesized titanium dioxide (TiO_2) nanoparticles. *Optik*, 193:163011, 2019.

[19] Andrea Welte, Christoph Waldauf, Christoph Brabec, and Peter J Wellmann. Application of optical absorbance for the investigation of electronic and structural properties of sol-gel processed TiO_2 films. *Thin Solid Films*, 516(20):7256–7259, 2008.

[20] Kissoum Yacine, Mekki, Mohamed Bououdina, Elfahem Sakher, and Stefano Bellucci. Dependence of Fe doping and milling on TiO_2 phase transformation: optical and magnetic studies. *Journal of Superconductivity and Novel Magnetism*, 33(2):427–440, 2020.

[21] Nazli Turkten, Zekiye Cinar, Ayse Tomruk, and Miray Bekbolet. Copper-doped TiO_2 photocatalysts: application to drinking water by humic matter degradation. *Environmental Science and Pollution Research*, 26(36):36096–36106, 2019.

[22] Harry Charalambous, Shikhar Krishn Jha, Han Wang, Xin Li Phuah, Haiyan Wang, and Thomas Tsakalakos. Inhomogeneous reduction and its relation to grain growth of titania during flash sintering. *Scripta Materialia*, 155:37–40, 2018.

[23] Freitas Renato, Santanna MA, and Ernesto Pereira. Dependence of TiO_2 nanotube microstructural and electronic properties on water splitting. *Journal of Power Sources*, 251:178–186, 2014.

[24] Gambino J, Robbins J, Rutkowski T, Johnson C, DeVries K, Rath D, Vereecken P, Walton E, Porth B, Wenner M, McDevitt T, Chapple-Sokol J, and Luce and S. Etching of copper in deionized water rinse. In *2008 15th International Symposium on the Physical and Failure Analysis of Integrated Circuits*, pages 1–4. IEEE, 2008.

[25] Tsung-Chi Pan, Sheng-Hung Wang, Yu-Sheng Lai, Jih-Mirn Jehng, and Shi-Jin Huang. Study of the silver modified TiO_2 nanotube array applied to hydrogen evolution. *Applied surface science*, 296:189–194, 2014.

[26] Nurul Asma Samsudin, Zulkarnain Zainal, Hong Ngee Lim, Yusran Sulaiman, Sook-Keng Chang, Ying-Chin Lim, and Wardatun Nadrah Mohd Amin. Enhancement of capacitive performance in titania nanotubes modified by an electrochemical reduction method. *Journal of Nanomaterials*, 2018, 2018.

[27] Hana Pelouchova, Pavel Janda, Jan Weber, and Ladislav Kavan. Charge transfer reductive doping of single crystal TiO_2 anatase. *Journal of Electroanalytical Chemistry*, 566(1):73–83, 2004.

[28] Katarzyna Siuzdak, Mariusz Szkoda, Mirosław Sawczak, Anna Lisowska-Oleksiak, Jakub Karczewski, and Jacek Ryl. Enhanced photoelectrochemical and photocatalytic performance of iodine-doped titania nanotube arrays. *Rsc Advances*, 5(62):50379–50391, 2015.

[29] Ma Qing, Liu Shaojun, Weng Luqian, and Bin Liu. Growth, structure and photocatalytic properties of hierarchical Cu-Ti-O nanotube arrays by anodization. *Journal of alloys and compounds*, 501(2):333–338, 2010.

[30] Shiva Mohajernia, Seyedsina Hejazi, Pavlina Andryskova, Giorgio Zoppellaro, Ondrej Tomanec, Radek Zboril, and Patrik Schmuki. Conductive Cu-doped TiO_2 nanotubes for enhanced photoelectrochemical methanol oxidation and concomitant hydrogen generation. *ChemElectroChem*, 6(4):1244–1249, 2019.

[31] Bae Ho Park, Lin Song Li, Brady Gibbons, Jung Huang, and Jia Qin. Photovoltaic response and dielectric properties of epitaxial anatase-TiO_2 films grown on conductive $La_{0.5}Sr_{0.5}CoO_3$ electrodes. *Applied Physics Letters*, 79(17):2797–2799, 2001.

[32] Ashi Ikram, Sonal Sahai, Snigdha Rai, Sahab Dass, Rohit Shrivastav, and Vibha R Satsangi. Improved charge transportation at PbSQDs/TiO_2 interface for efficient pec hydrogen generation. *Physical Chemistry Chemical Physics*, 18(23):15815–15821, 2016.

[33] Nir Baram and Yair Ein-Eli. Electrochemical impedance spectroscopy of porous TiO_2 for photocatalytic applications. *The Journal of Physical Chemistry C*, 114(21):9781–9790, 2010.

[34] Antonio Gonzãlez Muñoz, Qizhi Chen, and Patrik Schmuki. Interfacial properties of self-organized TiO_2 nanotubes stud-

ied by impedance spectroscopy. *Journal of Solid State Electrochemistry*, 11(8):1077–1084, 2007.

[35] Corrado Garlisi, Chia-Yun Lai, Leslie George, Matteo Chiesa, and Giovanni Palmisano. Relating photoelectrochemistry and wettability of sputtered Cu-and N-doped TiO_2 thin films via an integrated approach. *The Journal of Physical Chemistry C*, 122(23):12369–12376, 2018.

[36] Mengyao Wu, Tigang Duan, Ye Chen, Qing Wen, Yuyang Wang, and HoFngmei Xin. Surface modification of TiO_2 nanotube arrays with metal copper particle for high efficient photocatalytic reduction of Cr(VI). *Desalination and Water Treatment*, 57(23):10790–10801, 2016.

[37] Ilie Hanzu, Thierry Djenizian, and Philippe Knauth. Electrical and point defect properties of TiO_2 nanotubes fabricated by electrochemical anodization. *The Journal of Physical Chemistry C*, 115(13):5989–5996, 2011.

[38] Mariusz Szkoda, Anna Lisowska-Oleksiak, and Katarzyna Siuzdak. Optimization of boron-doping process of titania nanotubes via electrochemical method toward enhanced photoactivity. *Journal of Solid State Electrochemistry*, 20(6):1765–1774, 2016.

Chapter 5

The propagation of elastic waves generated in non-contact way in carbon and glass fibre-reinforced polymers

Damian Mindykowski

Institute of Fluid Flow Machinery, Polish Academy of Sciences,
Fiszera 14, 80-231 Gdansk, Poland

5.1 Introduction

The investigation contained in this chapter is related to the analysis of elastic waves propagation in two anisotropic materials – carbon fibre- and glass fibre-reinforced polymer (CFRP and GFRP). These waves are initially generated by air-coupled transducer (ACT) in non-contact way. The analysed specimens are plate-like structures.

Previous work of the author of this chapter [1] was referred to the analysis of ACT-based elastic waves generation in CFRP plate-like structure. However, most of the measurements was performed with the ACT slope angle steps of 5°. One of the recommendations for further work was to investigate the problem of selective A_0/S_0 Lamb waves mode generation assuming denser measurement grid in terms of the slope angle steps. The analysis of elastic waves interactions with defects also was desired.

The relevant measurements were thus performed in this chapter with the angle steps of 1°. Additionally, elastic waves propagation was considered while detecting additional mass-type defects in the same (in terms of the material) but smaller CFRP plate. Beside of CFRP specimens, elastic waves propagation analysis was performed in GFRP plate-like structure containing four delaminations of different shapes. The performance of different ACTs arrays was investigated and compared. The summary with conclusions and recommendations for further work is at the end of the chapter.

One of the aims of research in frame of which the current chapter is written is the analysis of elastic waves propagation in CFRP plate (undamaged/damaged state) and GFRP plate (damaged state). The main task is to increase the current state of knowledge related to the conversion phenomenon of acoustic waves propagating in the air into elastic waves propagating in the composite specimen.

5.2 Elastic waves generated in non-contact way

Non-contact generation of elastic waves often relies on the usage of air-coupled transducer (ACT). The waves are initially generated

in the air (in fact - inside the ACT casing) in form of the pressure waves, with alternative denser and looser regions of the air particles. After generation, these acoustic waves propagate through the air medium in direction of the solid specimen. When they reach the specimen surface they are converted into elastic waves during interactions with the solid. Obviously, the conversion is performed with the presence of energy losses – some of the acoustic waves are totally reflected back in the air and some of them are refracted in the solid (part of the waves energy is lost). The important factor influencing the efficiency of elastic waves generation is optimal slope angle of the transducer (the efficient waves generation means low acoustic waves energy losses and generation of the selected Lamb waves modes with high signal amplitudes). Hence, the operation of properly sloped ACT causes that bigger part of acoustic waves energy is transferred into selected waves mode/modes. The situation is schematically shown in the Fig. 5.1. Each of the optimal slope angles ($\theta_{\mathrm{opt},1}$, $\theta_{\mathrm{opt},2}$, $\theta_{\mathrm{opt},3}$) refers to the efficient generation of different Lamb waves mode. Hence, few different modes may be generated simultaneously, however with different shares of total acoustic waves energy. However, if ACT is sloped, it decreases the coverage area of specimen (coverage area means this part of the whole investigated area which is covered by elastic waves propagating with relatively high signal amplitudes).

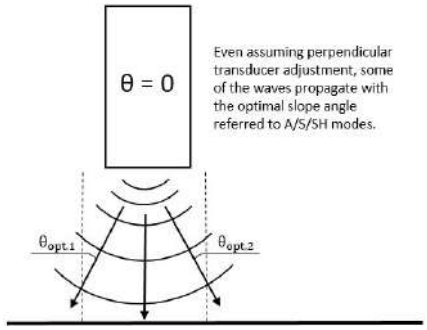
As the meaning of optimal ACT slope angle is important, it should be underlined how to calculate it. Optimal value of the slope angle can thus be determined using the Snell law:

$$\theta_{\mathrm{opt}} = \arcsin\left(\frac{\mathrm{c_{air}}}{\mathrm{c_p}}\right), \tag{5.1}$$

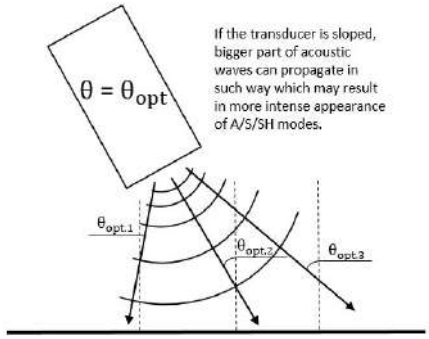
where: $\mathrm{c_{air}}$ – acoustic waves velocity in the air, assumed as 343 $\frac{\mathrm{m}}{\mathrm{s}}$ in so-called normal atmospheric conditions (air pressure of 1 atm, air temperature of 20°C); $\mathrm{c_{ph}}$ – phase velocity of elastic waves mode, calculated from the following formula:

$$\mathrm{c_{ph}} = \lambda \cdot \mathrm{f} = \frac{1}{\mathrm{k}} \cdot \mathrm{f}\,, \tag{5.2}$$

where: λ – wavelength of the selected waves mode [m], f – centre excitation frequency of the transducer [Hz], k – wavenumber corresponding to the selected mode $[\frac{1}{\mathrm{m}}]$.



(a) Non-sloped ACT.



(b) Sloped ACT.

Fig. 5.1: Scheme of the process of acoustic waves conversion into elastic waves due to interactions with solid specimen.

Wavenumbers corresponding to the selected waves modes can be calculated numerically or estimated based on dispersion curves, included in dispersion maps. These curves depict dependency between the waves mode velocity values (or wavenumber values) and excitation frequency values multiplied by the thickness of specimen. Dispersive character of elastic waves means that they can propagate with different velocity, depended of the excitation frequency. Hence, wavenumber value of the chosen waves mode for the se-

lected frequency and specimen thickness can be read from dispersion map. Its outlook also depends on the relation with material fibres orientation. Hence, if the structure is not uniform, more than one dispersion map should be prepared for that. It turns out that elastic waves properties (signal amplitude, wavenumber) may differ, dependent on the relation with the fibres. Example dispersion graph determined in Dispersion Calculator program for the orthotropic epoxy-based plate-like structure with the thickness of 2 mm is shown in the Fig. 5.2.

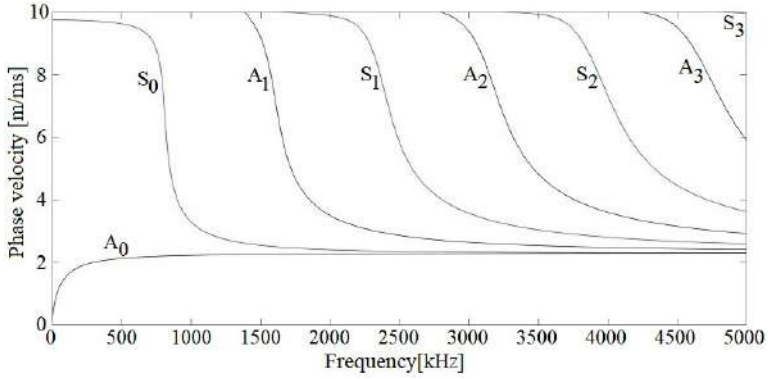
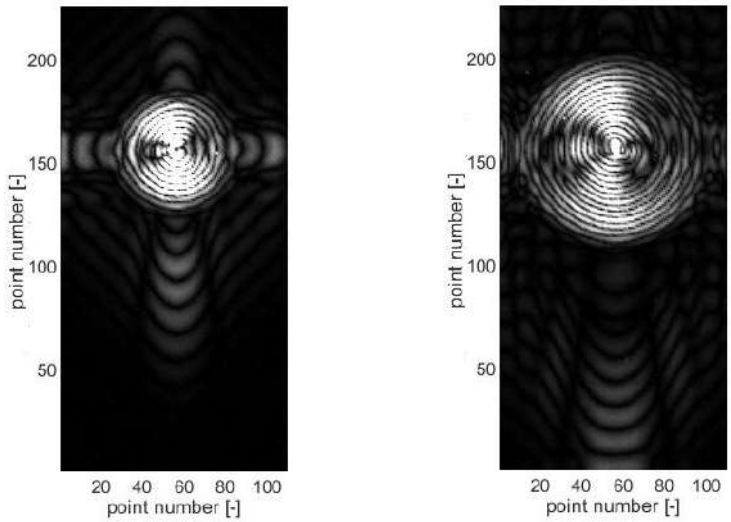


Fig. 5.2: Dispersion graph for 2 mm thick composite plate-like structure, prepared in Dispersion Calculator.

Analysing Fig. 9.2, it can be found that various Lamb waves modes can propagate in the specimen within different excitation frequencies. Antisymmetric and symmetric modes of different orders may propagate and generally high order modes propagate within relatively high excitation frequencies. For the lower frequency values (below ∼1200 kHz) it can be found that only zero order modes propagate in the specimen. Symmetric S_0 mode propagates faster than antisymmetric A_0 mode, with simultaneously having longer wavelength. Such situation can be observed while watching example results of signal processing of the measurement data. The animation consisted of frames, presenting Lamb waves propagation in plate-like structure (CFRP) in different time steps can thus be considered. Piezoelectric transducer (PZT) was used to generate

elastic waves. The excitation frequency equalled 150 kHz. Selected frames from the aforementioned animation are shown in the Fig. 5.3. It is clearly visible that S_0 mode is characterized by longer wavelength and higher velocity.



(a) Time step no. 100. (b) Time step no. 150.

Fig. 5.3: Selected frames from the animation depicting A_0 and S_0 Lamb waves propagation in example CFRP structure.

In addition to the Lamb waves modes of different orders, shear horizontal (SH) waves modes also can propagate in plate-like structures. During SH waves propagation solid particles oscillate perpendicularly to their propagation direction, forming horizontal lines. More details about guided waves can be found in [2].

5.3 Literature overview

Nowadays, composite structures are getting still rising meaning – especially in few industry branches, including aerospace, aircraft, automotive and civil engineering. It is mainly caused by their good mechanical and physical properties while having relatively low weight [3] compared to metal-based materials (steel, aluminium and

their alloys). Along with the continuous development of more and more advanced structures, there is still growing demand on non-destructive testing (NDT) methods referred to these structures, in order to prevent accidents and catastrophes. The promising technology offering fast, accurate and low cost inspections is related to the guided waves-based methods. The technology is developing since 1970 [4]. Guided waves are elastic waves which may propagate in solids limited by parallel surfaces, causing vibrations of the material particles while they propagate. Hence, the differences between guided waves-based methods rely mainly on the usage of different equipment, resulting in different mechanisms of the guided waves generation and detection processes. The authors of [5] used 40 kHz-based ACT for the elastic waves generation and scanning laser Doppler vibrometer (SLDV) Polytec PSV-400 for their detection in 3.9 mm CFRP plate-like structure case. They performed numerical and experimental analysis referred to determination of the optimal transducer slope angles, in order to generate different zero order Lamb waves modes (A_0, S_0) and SH_0 mode efficiently. Within the numerical research they prepared the model of considered phenomena using semi analytical spectral element method (SASE). Another work (however only experimental) using the similar equipment was done by Xiao et al. [6]. They used different ACTs with higher excitation frequencies (350 kHz and 510 kHz), generating the waves propagating in 7.5 mm thick K9 glass plate. The research was aimed at analysis of the air-coupled signals. The sensitivity of each ACT was examined, based on measuring of the radiated acoustic power. Slightly another equipment (PZT for guided waves generation and SLDV for the waves detection) was used by Hu et al. [7]. Similarly as in [5] they conducted both numerical and experimental analysis. It was related to the guided waves mode conversion between A_0 mode and quasi-Scholte (QS) waves. Scientists considered partially immersed 2 mm thick steel plate (the plate was immersed vertically). They used Abaqus program for the numerical investigation. In case of the experimental research, they initially generated A_0 mode on the dry side of the specimen. Then, the waves propagated in direction of the other (wet) side of the specimen and there they converted into QS waves. It was found out that GW mode

may shift its frequency if phase velocity of the incident A_0 mode is higher than the sound speed in water. It was observed that the resulted quasi-Scholte waves energy decreases while the incident A_0 excitation frequency increases. Another similar analysis (in terms of the used equipment and software) was performed by Bustamante et al. [8]. They considered GW propagation in 4 mm thick CFRP plate, using Matlab and Abaqus in numerical part of the research. They used PZT to excite elastic waves and three scanning heads of SLDV to detect them. Using this equipment, they analysed distortions in the registered signals, trying to connect them with defects existed in the specimen. In turn, Rao and Duan [9] used Macro Fibre Composite (MFC) actuators for GW generation and Fibre Bragg Grating (FBG) sensors for GW detection. They performed measurements referred to the GW propagation in the frequency range of 20-140 kHz. They distinguished A_0 and S_0 waves modes based on A-scan images.

The modern guided waves-based approach is related to their non-contact excitation using ACT [5, 7, 8] and to their non-contact detection using SLDV [10–12]. Other popular approach is referred to the usage of PZTs to generate elastic waves and SLDV for their detection [13–15]. Many popular works correlated to the aforementioned equipment are concerned on the analysis of GW propagation from viewpoint of detection of different defects. Some of the researchers are focused on detection of delaminations [14–16], disbonds [11] or other defects [10, 17]. Tian et al. [13] focused on the detection of two delamination-type defects, correlated with the matrix cracking in CFRP plate. The considered specimen was in the form of plate-like structure having the thickness of 2.8 mm, in opposite to the authors of [14] who considered two different GFRP plates with the thickness of 1.5 mm, simultaneously investigating four circular delaminations in the form of thin Teflon film inserts and three impact-type damages. They analysed elastic waves mode conversion from A_0 to S_0 mode during the waves interactions with delaminations. The mode conversion index (MCI) was proposed. Sha et al. [15] also considered Teflon-type delaminations in 2 mm thick GFRP plate. During the analysis they generated toneburst signal with the central frequency of 50 kHz, consisted of 5 sine cycles

modulated by Hann window. Quattrocchi et al. [16] detected two delamination-type defects and regions with local excess and lack of the epoxy resin in two GFRP plates with the thickness of 9 mm each. Scientists concluded that within the regions with excess of the resin low signal amplitudes of the waves were noted, meanwhile within the regions with lack of the resin high signal amplitudes of the waves were registered. Besides, in terms of the excitation frequency of the generated signal different values were assumed within researches. For instance, in opposite to [15] it equalled 120 kHz [13], 140 kHz, 200 kHz and 500 kHz [16], 100-200 kHz [7] or 200 kHz [18].

Beside of some experimental works [8, 10, 19] some of the researchers also done numerical researches with different aims [17, 18, 20]. Agrawal et al. [20] performed numerical investigation of non-linear interactions of dual-frequency toneburst signal in delamination zone. In other words, they implemented so-called contact acoustic nonlinearity (CAN) approach to model interactions of sub-laminates in delaminated region. It was proved that resulted waves frequencies (higher harmonics) are related to delamination location and width. Mardanshahi et al. [17] done the modelling process in Abaqus (finite element (FE) simulations) with the usage of input data such as damping coefficients of the investigated material. The Lamb waves simulations were performed in aim to detect matrix cracking in GFRP specimen. Instead of [17] Liu et al. [18] performed numerical investigation based on virtual time reverse (VTR) algorithm. In case of this algorithm the time reverse is achieved by proper signal processing of the results. VTR algorithm is the modified version of the well-known reconstruction algorithm for probabilistic inspection of damage (RAPID). Luca et al. [21] prepared FE model for GW propagation in complex blended winglet structure, made of GFRP. They used 1 PZT for elastic waves generation and 11 PZTs for the waves detection. The considered defect was in the form of low velocity impact (LVI) damage, made during drop test of the mass weighted 0.8 kg. For the waves actuation the signal with base excitation frequency of 100 kHz was generated. Instead of LVI damage, Zhang et al. [22] analysed the process of fatigue damage evolution. The damage was in the form of transverse crack as the response for the micro fatigue damage accumulation. During

the experimental research they used pulsed laser to generate elastic waves and acoustic emission sensor for the waves detection. Later, numerical research referred to the dependency between Lamb waves mode conversion effect and matrix crack density was executed, with the preparation of FE model in Abaqus. Another research related to the identification of elastic constants of the material of the considered specimen was done by Kudela et al. [23]. Elastic constants were identified based on GW-based dispersion curves and genetic algorithm for improved simulations. They developed spectral element (SE) model and later semi-analytical spectral element (SASE) model to perform numerical research in the relevant software. Within the experimental part of research chirp and toneburst 5 sine cycles-based signal in the excitation frequency range of 0-500 kHz was generated by scientists. Median filtering was applied to reduce noise in spatial images. 3D Fast Fourier Transform (FFT) also was used in the research.

5.4 Measurements

As it was aforementioned, measurements which will be described in this chapter are related to CFRP and GFRP plate-like structures. Different types of measurements were done in aim to analyse elastic waves propagation in the considered specimens, focusing on the use of non-contact waves excitation method. Furthermore, the detection process of different defects in the specimens was performed.

5.4.1 The measurement setup

The measurement setup was similar in case of all investigated specimens. Each plate-like specimen was considered separately. The specimen was mounted vertically. From the back side of the plate the transducer (elastic waves generator) was placed – either PZT (Noliac NCE51, bonded to the middle point of the plate surface) or ACT (adjusted perpendicularly/sloped to the plate, mounted using holder which was ∼24.5 mm distant from the plate middle point). From the plate front side scanning laser Doppler vibrometer (SLDV) was located (distance of ∼2 m) and used for elastic

waves sensing. In terms of the mutual location of the measurement table and the plate – they were adjusted in similar way as in the setup B described in [1]. Overall scheme of the measurement setup is shown in the Fig. 5.4. The used SLDV is Polytec PSV-400.

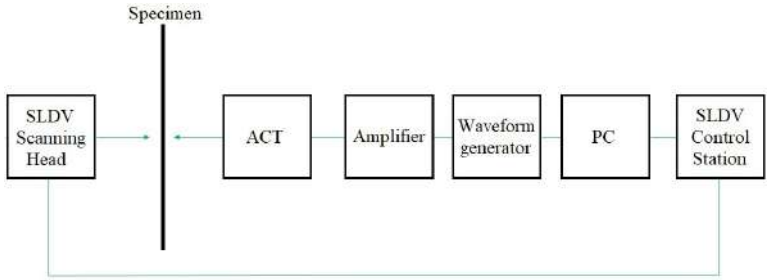


Fig. 5.4: Outlook of the measurement setup.

In terms of the used ACT, the transducer with base excitation frequency of 40 kHz was engaged. Its external diameter equalled 16 mm and total length – 12 mm. The ACT is made of metal membrane, metal disc (tuned for desired excitation frequency), piezoceramic/piezoelectric disc (active element) and two-partial plastic foundation, placed inside metal casing. Inside the casing, the membrane is coupled with metal disc by acrylic glue and the other parts are coupled by silicone glue). The used ACT is shown in the Fig. 5.5a,b. Moreover, the used SLDV scanning head is depicted in the Fig. 5.5c.

The electric signal supplied to the ACT consisted of the package of 10 sine cycles, modulated by the Hann window. All experiments were conducted while maintaining the lab air temperature on constant level of 20±1°C. To process the measurement data, Polytec laser software was used. For the data post processing, Matlab program was used.

5.4.2 CFRP specimens

Two CFRP specimens are considered in the section. Larger specimen has dimensions of 1200 x 1200 x 3.9 mm^3 and smaller specimen has dimensions of 500 x 500 x 3.9 mm^3. Material properties in both

(a) Photo of the ACT external outlook.

(b) Photo of the ACT inner construction.

(c) Photo of the SLDV scanning head.

Fig. 5.5: Outlook of the used elastic waves generator and detector.

structures are the same. The specimens consist of 40 prepreg uniaxial TDS-75g/m^2 layers and the IMP503Z epoxy resin. The plates have unidirectional structure in terms of the carbon fibres orientation (i.e. the fibres are parallel to one of the plate sides).

As it was aforementioned, point-wise measurements, measurements along line (5°-steps) and full wavefield measurements were done in larger CFRP plate case and the results were described in the previous work [1]. Afterwards, the measurements along line were additionally conducted with 1°-steps in terms of the transducer slope angle, in order to check the correctness of the results of previous measurements. The results of 1°-step measurements are therefore depicted in this section along with their comparison with 5°-step

measurements. The results of detection process of additional mass-type defects in smaller CFRP plate case are described in further part of the section.

Measurements along line

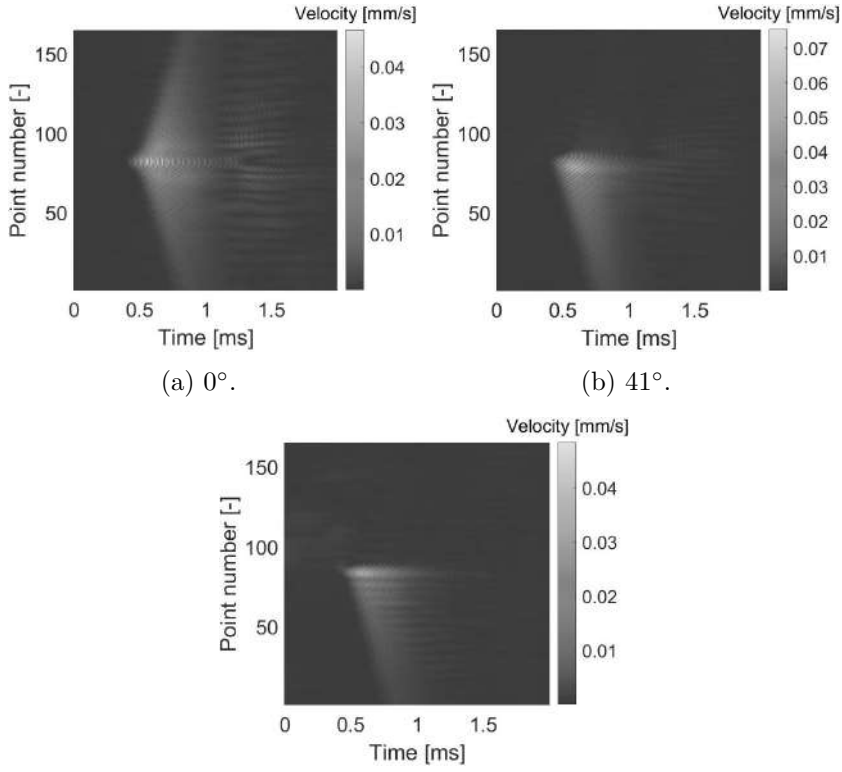
The manner of the measurements along line relies on gathering of the measurement values (i.e. vibration velocity values of material particles) along the line of measurement points which can be located somewhere in the analysed specimen area. Measurements along line can also be named as one dimensional (1D) full wavefield technique.

Elastic waves measurements in CFRP plate were conducted in 165 measurement points located along line with the length of ∼42 cm. The measurement line was defined in the middle zone of the plate. The point of the elastic waves generation was located in the middle of this line - similarly as in 5°-step measurements [1]. For the comparison, 127 measurement points were located along line during 5°-step measurements.

5.4.3 Signal amplitude measurements for various ACT slope angles – across-fibres case

Vibration velocity values of the material particles were thus measured for different ACT slope angles in the range of 0-70°. In the Fig. 5.6 time-space domain data in the form of waterfall maps are presented for the selected probe slope angles.

Obtained results are generally similar to these obtained from the measurements performed with the probe slope angle steps of 5°, what confirms the certainty of done measurements. Elastic waves propagation become almost completely unidirectional since the slope angle is ∼40°. Within lower slope angles the waves patterns visible on the right side of the graphs seem to be irregularly shaped. Their regularity is slightly improved within higher slope angles, though their visibility is more indistinct (at least after ∼1 ms). It may suggest that waves energy trapped in the structure vanishes faster within higher slope angles. Besides, irregular strips visible in 0°



(a) 0°.

(b) 41°.

(c) 70°.

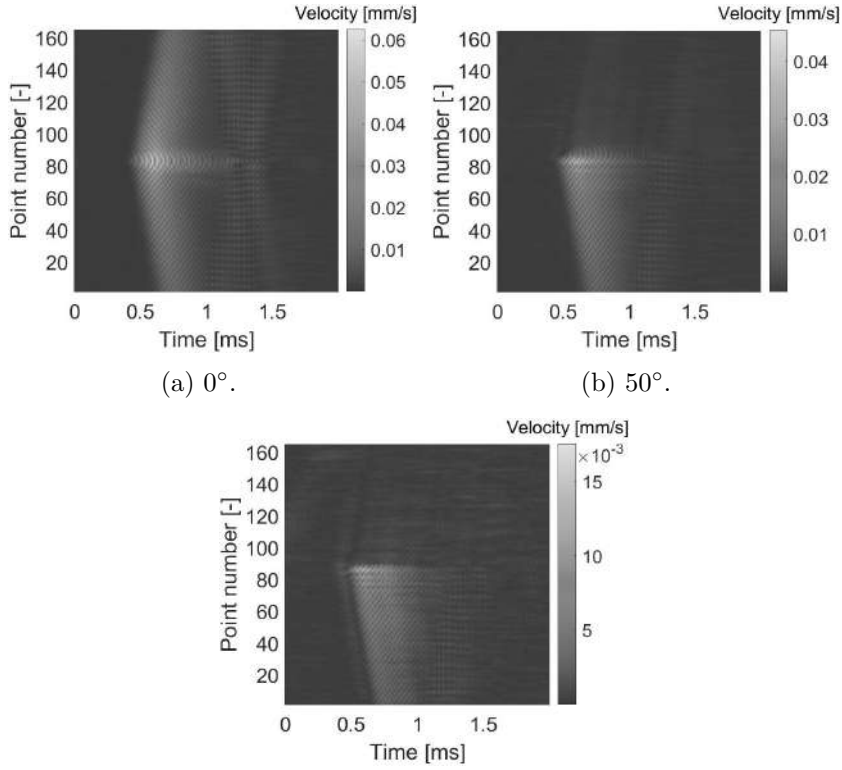
Fig. 5.6: Waterfall graphs comparison in across-fibres case for different, selected ACT slope angles.

case (Fig. 5.6a) are now narrower than in 5°-step-case, probably due to higher amount of the measurement points.

5.4.4 Signal amplitude measurements for various ACT slope angles - along-fibres case

The results in along-fibres case are shown in the Fig. 5.7. Analyzing Fig. 5.7a, one may notice more equal signal amplitude distribution among different measurement points in case of the elastic waves propagation in comparison with the equivalent map in across-fibres case. The similarity between the maps in Fig. 5.7 is visible on the

right side of the maps - the waves patterns are relatively regularly shaped, though they are narrower than in across-fibres case. Within the slope angles of ∼50° it can be clearly seen that some elastic waves (although with lower amplitude) still propagate in direction opposite to the ACT slope (similarly as in 5°-step measurements). For the slope angle of 70° one may notice significant decrease in signal amplitudes of the waves propagating in the plate. Differences in the amplitude values between the slope angles of 0°, 50° and 70° are slightly bigger comparing to the similar measurements done with 5°-steps.



(a) 0°.

(b) 50°.

(c) 70°.

Fig. 5.7: Waterfall graphs comparison in along-fibres case for different, selected ACT slope angles.

5.4.5 Comparison of waterfall maps in across-fibres and along – fibres case

Analysing the maps in across-fibres and along-fibres case, one may notice that in along-fibres case the waves have higher velocity in the material - they reach border measurement points faster. In both along-fibres and across-fibres case the dependency between distance from the generation point and time of the waves propagation is approximately linear (this linearity can be explained by the constant group velocity of Lamb waves modes propagating along the certain direction).

5.4.6 RMS graphs in across-fibres and along-fibres case - different ACT slope angles

Similarly as for the measurements conducted with 5°-steps, root mean square (RMS) graphs representing the signal energy in different measurement points for the selected probe slope angles in both across-fibres and along-fibres case are shown in the Fig. 5.8. Signal energy values were calculated using the following formula:

$$E_{RMS} = \sqrt{\frac{\sum_i^n c_{v,i}^2}{n}}, \tag{5.3}$$

where: $c_{v,i}$ – vibration velocity value of the material particles [V], n – amount of time steps during the measurement time period (1024).

One may conclude that:

- for the slope angle of 0° the highest signal energies were registered in points located in the nearby of the point no. 82 (the middle part of the line made of 165 points). Furthermore, higher signal energy values for this angle were registered in along-fibres case. In turn, the highest signal energy taking all considered slope angles into account can be seen in across-fibres case (for the slope of 40°). Similarly, for higher transducer slope angles (e.g. 70°) signals registered within the middle points have ∼2-3 times higher energy compared

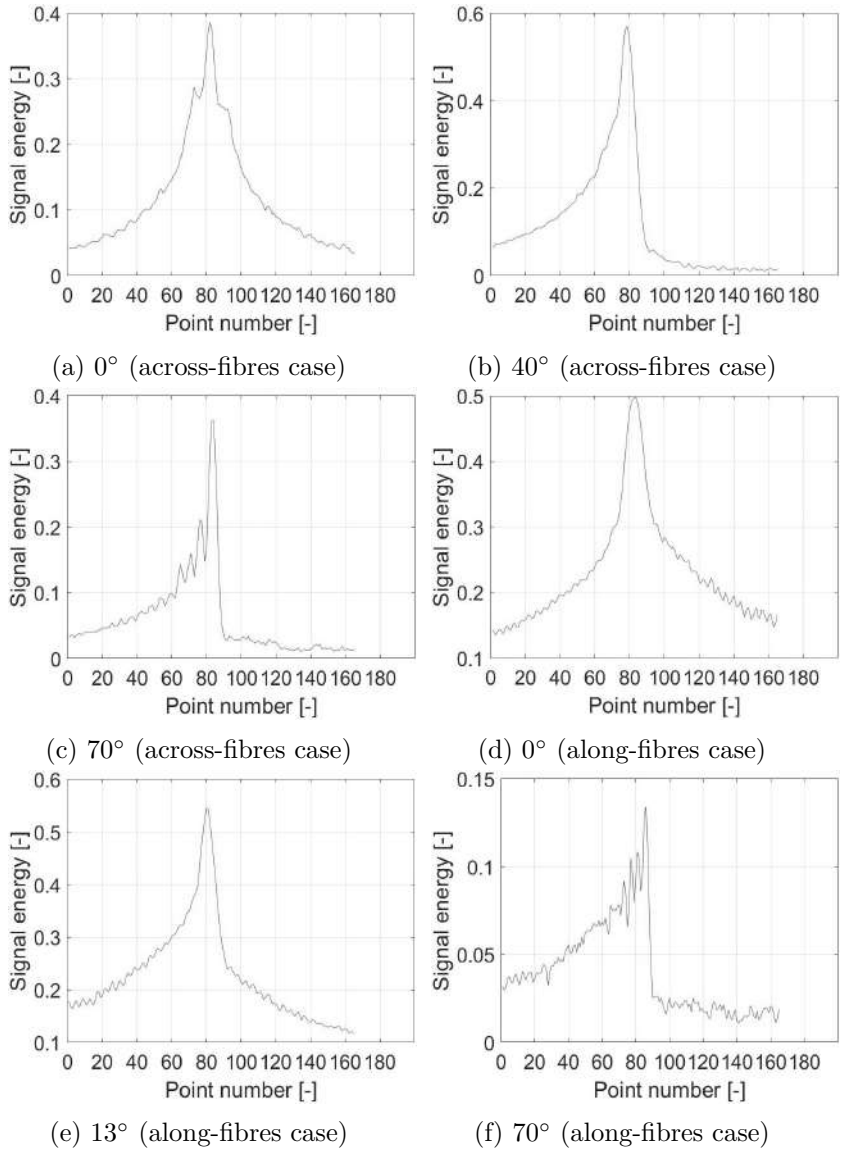
to along-fibres case,

- measurement points located in direction of the probe slope are characterized by higher signal energies than points located in the opposite direction. In some slope angle cases registered signal energy values in the nearby points vary less (e.g. points no. 1-60 in Fig. 5.8e - the difference is ∼70%) and sometimes these values vary more (e.g. points 1-60 in Fig. 5.8a – the difference is over 200%). Such situation results from higher elastic waves attenuation in the material when the waves propagate across fibres,
- along with the higher slope angles, more significant differences in signal energy values between left and right part of measurement points are visible, what is in line with expectations.

5.4.7 RMS graphs in across-fibres and along-fibres case - different measurement points

RMS graphs representing signal energy values calculated based on registered vibration velocity values in selected measurement points assuming various ACT slope angles are presented in the Fig. 5.9. It is analysed in order to compare the elastic waves-based signal amplitude profile between different measurement points, representing different locations of the specimen in relation to the waves generation point (i.e. middle point of the measurement line). Analysing the graphs in Fig. 5.9, the following issues can be found:

- the highest signal energies were registered in the point no. 79 (across-fibres case) and in the point no. 81 (along-fibres case). These points are located in the nearby of the middle point among 165 measurement points, what result from the fact that together with the slope angle increase the main point of elastic waves excitation moves closer to one of the border points. Summarizing, the highest signal energy noted in these points occurred for the slope of 40° in across-fibres case and for 13° in along-fibres case,
- in the middle point case - slightly higher signal energies can be observed in across-fibres case, though for less amount of



(a) 0° (across-fibres case)

(b) 40° (across-fibres case)

(c) 70° (across-fibres case)

(d) 0° (along-fibres case)

(e) 13° (along-fibres case)

(f) 70° (along-fibres case)

Fig. 5.8: Signal energy RMS graphs compared for different material fibres orientation and for different ACT slope angles.

slope angles than in along-fibres case,

- in the border points cases – higher signal energies can be found in along-fibres case, though course of the curves was generally the same.

5.4.8 Comparison between 5°-step measurements and 1°-step measurements

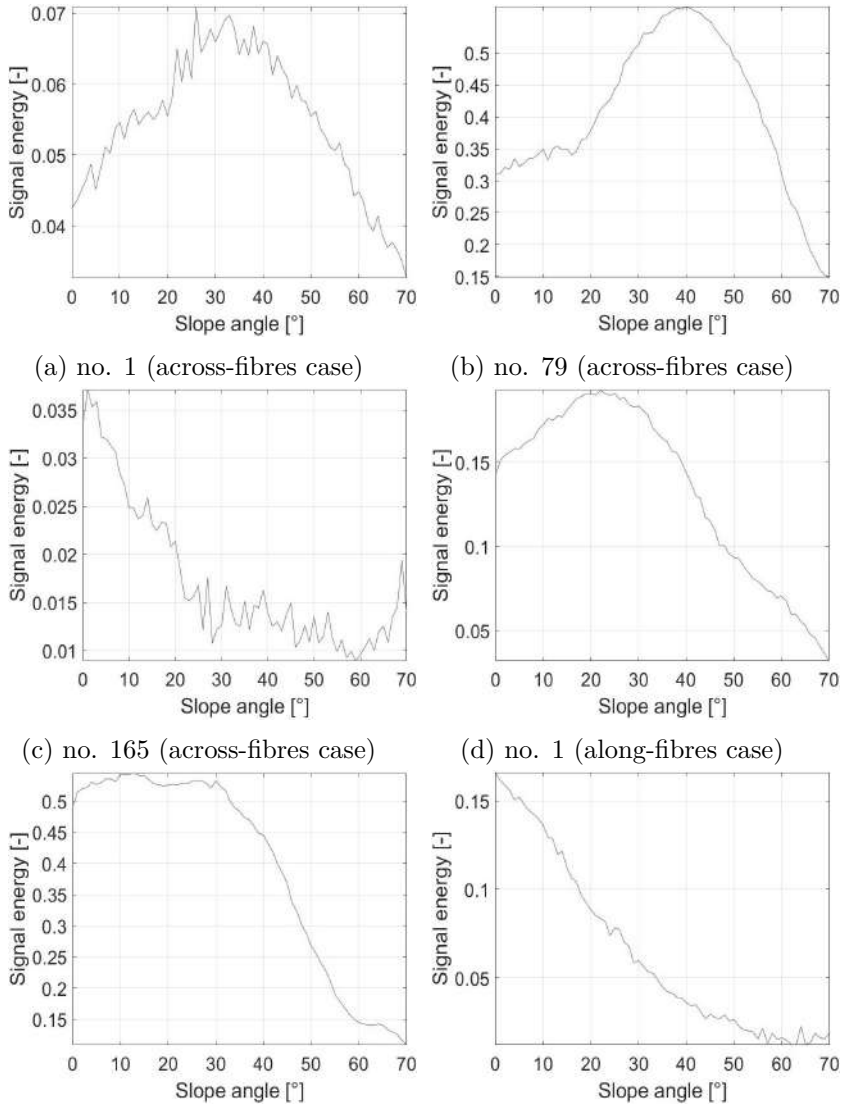
Although the analysis of data collected during both 5°-step measurements [1] and 1°-step measurements gave generally similar results, there are some differences between them if to take a closer insight. These differences may result from the following facts:

- length of the measurement line and amount of the measurement points are not the same,
- higher accuracy of measurements in 1°-step measurements cause more precise and thus different results,
- adjustment of the measurement equipment can vary a little, as each measurement series was performed in different time period.

Waterfall maps

Collating waterfall graphs prepared for the 5°-step case and 1°-step case, it can be found that:

- registered signal values are similar in both measurement series in across-fibres case. However, taking the along-fibres case into account, the signal values visible in the graphs prepared for selected probe slope angles are smaller for the slope of 0° and higher for the slope of 70° in 5°-step case compared to 1°-step case. For the slope angles of 40°/41° and 50° (i.e. the angle values over which the waves propagation is aimed almost in one direction) the registered values are similar,



(a) no. 1 (across-fibres case)

(b) no. 79 (across-fibres case)

(c) no. 165 (across-fibres case)

(d) no. 1 (along-fibres case)

(e) no. 81 (along-fibres case)

(f) no. 165 (along-fibres case)

Fig. 5.9: Signal energy RMS graphs compared for different material fibres orientation and for different measurement points.

- in along-fibres case higher equality in terms of the signal values registered in points located along the measurement line can be observed in 5°-steps case than in 1°-steps case. Comparing the graphs in across-fibres case - similar relation can be seen,
- similar shapes of the front waves patterns are observed.

RMS graphs for different ACT slope angles

Comparing the 5°-step case graphs with 1°-step case graphs plotted for different measurement points (each time assuming selected probe slope angle) one may notice that:

- shapes of gained results (i.e. the curves) in 1°-step measurements (Fig. 5.8) are more or less in line with the expectations. Comparing them with the graphs plotted for 5°-step measurements [1] it can be concluded that higher measurement accuracy causes less smooth course of the curves,
- graphs prepared for various, selected slope angles have similar outlook, except in 0° slope angle cases in along-fibres case (Fig. 5.8d), [1]). It is suspected that such situation occurred due to different location of the material fibres in relation to the ACT middle axis (elastic waves propagate differently within the plate fibres and between them), as two compared measurement series took place in different time. Although the effort was made to mount the plate perfectly horizontally and similarly in both series a small error could happen, as the mounting was manual. In case of the registered signal values - they are similar,
- in both 5°-step and 1°-step case the highest signal values were observed for similar slope angles, i.e. for 35° (5°-step case, across-fibres case [1]), 40° (1°-step case, across-fibres case, Fig. 5.8b), 15° (5°-step case, along-fibres case [1]), and for 13° (1°-step case, along-fibres case, Fig. 5.8e).

RMS graphs for different measurement points

Analysing the graphs in both 5°-step case [1] and 1°-step case (Fig. 5.9) for different ACT slope angles (each time assuming selected measurement point), the following issues can be found:

- courses of the curves plotted for 5°-step measurements and 1°-step measurements are similar (in terms of the shape),
- the registered energy values have similar scale of magnitude. Comparing the results for the grid middle points (Fig. 5.9b,e) and Fig. **??**a,b [1]), signal energy values are similar – differences are below ∼15%.

Point-wise and full wavefield measurements

Point-wise measurements are the measurements performed at one measurement point, comparing the registered values for different measurement setup settings. In turn, full wavefield technique is two dimensional (2D). The measurements are gathered for dense mesh of measurement points spanned over the surface of the investigated sample. It allows to i.a. display animations presenting phenomenon of elastic waves propagation and to perform the analysis of their propagation.

As it was aforementioned, point-wise and full wavefield measurement results were previously presented [1], together with depiction of the measurements specification. It was decided to remind the main issues from that measurements.

5.4.9 Point-wise measurements for different excitation frequency values

Measurements done in one measurement point for different excitation frequencies were performed in aim to check the impact of the frequency on the magnitude of generated signal amplitude (and RMS-based signal energy). As the considered ACT has the base frequency of 40 kHz, it was decided to check its performance in the frequency range of 35-45 kHz with 100 Hz step (the Fig. 5.5 in [1]). It turned out that:

- the highest signal amplitude was registered within the frequency of 38.6 kHz,
- the highest signal energy was observed for the frequency of 38.3 kHz.

Differences between the maxima values mentioned above and these observed within the frequency of 40 kHz are equal to ∼5%, so they are relatively small. However (although the performance of the ACT is better for the frequencies below 40 kHz), to let the measurement results be comparable with other works it was decided to assume 40 kHz in all experiments. ACTs with the base frequency of 40 kHz are more common in the probe market than 38.6 kHz or 38.3 kHz.

5.4.10 Point-wise measurements for various distances between ACT and plate

Measurements done in one measurement point for different distances between the transducer and the plate were performed in aim to check the impact of the distance on the registered signals. Additionally, this impact was checked for two different measurement setups – A and B, however the setup B was finally assumed. The investigation was conducted for both material fibres orientations (across the fibres and along the fibres). The current subsection is thus based on the Figs. 1.7 and 1.8 [1]. Signal processing of the measurement data was performed in such manner to depict the following dependencies:

- mean vibration velocity values in function of the distance and time,
- maximum and mean vibration velocity values for different distance values (for each distance 1024 velocity values were registered during each measurement in the form of 1024 time steps). Distance values taken into account for the setup B equalled 0-3 mm with 1 mm steps and 5-125 mm with 2 mm steps.

Based on the results, it can be claimed that the highest signal amplitude values were noted for the distance values below 10 mm

(shorter the distance – higher the amplitude). Based on waterfall maps, it can be confirmed that beside of the first acoustic waves reaching the plate there are next acoustic waves packages which reach the plate after their first reflections from the plate and the ACT probe casing. In order to generate the second acoustic waves package with relatively high (visible in comparison to the first package) vibration velocity values it was noted that the aforementioned distance should not be longer than ∼20 mm.

5.4.11 Full wavefield measurements

Areal measurements were done for the considered specimen assuming different measurement setups, ACT slope angles and specimen fibres orientations. The aim of these measurements was the investigation of elastic waves generation in the plate in terms of registered signal amplitudes and the coverage area. After initial considerations, the setup B was finally assumed for the rest of measurements. The main reason of the setup B selection is the willingness to obtain accurate results, undisturbed by the measurement table presence, etc. The other reason to choose setup B is to compare full wavefield results with other results of this research (most of the measurements were done assuming setup B). Two specimen fibres orientations in relation to the ACT slope (across fibres, along fibres) were then taken into account and the measurements were done for different transducer slope angles of 0-70° with 10°-steps. Based on the results, the following issues can be specified:

- the highest signal energy values (calculated using the RMS approach - formula (3)) were noted for the slope angle of 30° in case of both fibres orientations,
- the highest coverage area (using one transducer) existed for the slope of 0°, although within certain range of higher slope angles (e.g. for 30°) one half of the whole considered area is better covered by elastic waves than in 0° case. However, considering the whole coverage area – it decreases while the slope angle increases,
- elastic waves “prefer” to propagate along the fibres than across

them. The word "prefer" means the efficient propagation, i.e. the propagation with higher velocities and lower energy losses.

Detection of the additional mass-type defects

As it was aforementioned, detection of mass-type defects was performed in case of the smaller CFRP plate. In total, three steel nuts of the same size (diameter of 20 mm and height of 5 mm) were coupled to the plate surface using cyanoacrylate glue. The nuts are destined to simulate real damages. Overall outlook of the plate is depicted in the Fig. 5.10. The process of mass-type defects detection was described in [5], hence in this chapter it will be just reminded and the relevant issues will be depicted. Also, the results analysis will be performed in slightly different way as in [5].

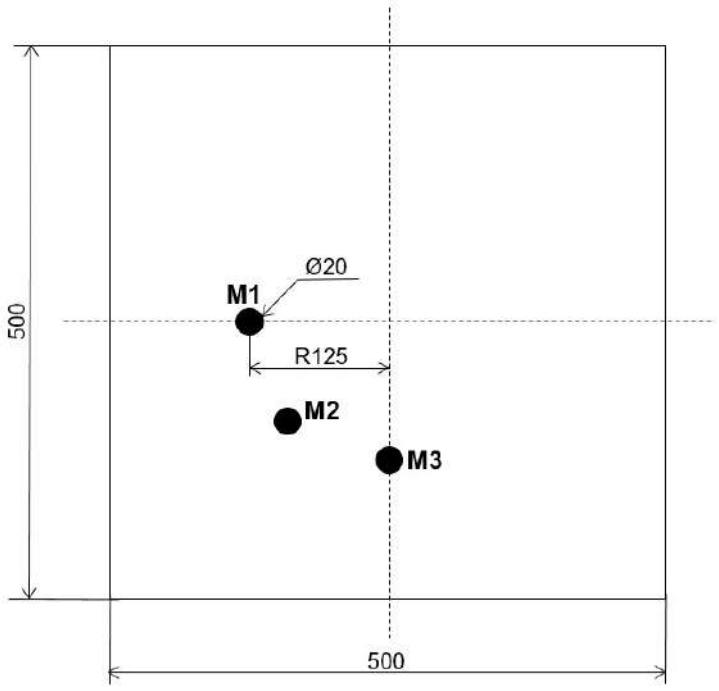


Fig. 5.10: Scheme of the CFRP plate with marked mass-type defects locations (dimensions in mm).

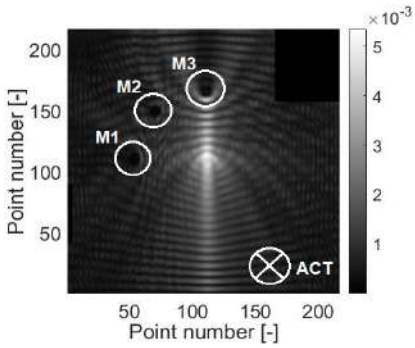
Obviously, to generate elastic waves and detect the defects one ACT was used - the same one as in previous measurements (base excitation frequency of 40 kHz). Measurements were done in full wavefield manner for different adjustments of the ACT:

- ACT located perpendicularly to the plate,
- ACT sloped in direction of the defects (across-fibres case),
- ACT sloped in direction opposite to the defects (across-fibres case),
- ACT sloped in direction of the defects (along-fibres case),
- ACT sloped in direction opposite to the defects (along-fibres case).

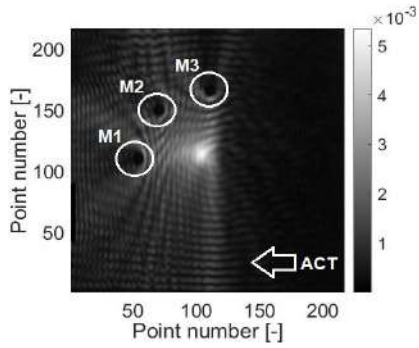
Based on currently done research [1, 2] the transducer slope (in relation to the line normal to the specimen surface) was set to 34° in across-fibres case and 16° in along-fibres case – to generate A_0 Lamb waves mode efficiently. Initially, the cases with ACT sloped in direction of the defects were compared with the perpendicular ACT setting case. The ACT was located in front of the specimen centre in each considered case. The results are shown in the Fig. 5.11.

Most of the elastic waves energy propagate along the material fibres, what results in accurate detection of the defect M3 (Fig. 5.11a, Fig. 5.11c). Analysing the results, one may notice that considered mass-type defects can be detected in the most accurate way while the ACT is sloped in the defects direction in across-fibres case (Fig. 5.11b) - across-fibres measurement manner helps in the efficient detection of M3 defect and slope of the ACT helps to detect M1 and M2 defects. Besides, considering elastic waves propagation patterns the waves reflected from the defects can be noted. The correlated energy strips are especially visible in Fig. 5.11a and Fig. 5.11c. Similar values in terms of the waves energy existed in all cases in Fig. 5.11. Furthermore, during the preparation of Fig. 5.11a measurement errors appeared within right upper part of the map. Elastic waves energy in that region was thus minimized, to avoid the errors impact on the rest of the results.

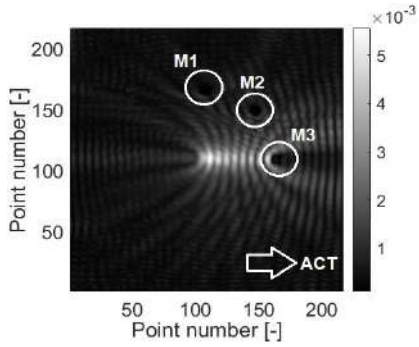
Other, similar issue is the analysis of full wavefield results con-



(a) ACT set perpendicularly (fibres oriented vertically).



(b) ACT sloped in direction of the defects (across-fibres case).



(c) ACT sloped in direction of the defects (along-fibres case)

Fig. 5.11: Detection of damages in smaller CFRP plate.

cerning across-fibres and along-fibres case with the ACT sloped in direction opposite to the defects locations. Further information can be found in [5], however in both cases only one mass-type defect was detected (mainly in terms of the location, not in terms of shape and size).

5.4.12 GFRP specimen

One GFRP specimen is considered in the section. It is plate-like structure with external dimensions of 500 x 500 x 1.5 mm^3 and

cross-type (0°/90°) orientation of the glass fibres. The structure consists of 12 prepreg cross-type layers (VV192T 202 g/m^2) and the epoxy resin (IMP503). The structure area investigated by the lab equipment (e.g. SLDV) has dimensions of 500 x 500 mm^3. The specimen contains four artificially-made defects of different shapes, which can simulate real delaminations (circle with diameter of 20 mm, square 20 x 20 mm^2, rectangular 20 x 30 mm^2 and ellipse 20 x 30 mm^2. Delaminations are thus simulated by thin Teflon film inserts embedded on the same level, between selected material layers (∼0.5 mm down from the upper specimen surface). Overall outlook of the plate can be found in the Fig. 5.12.

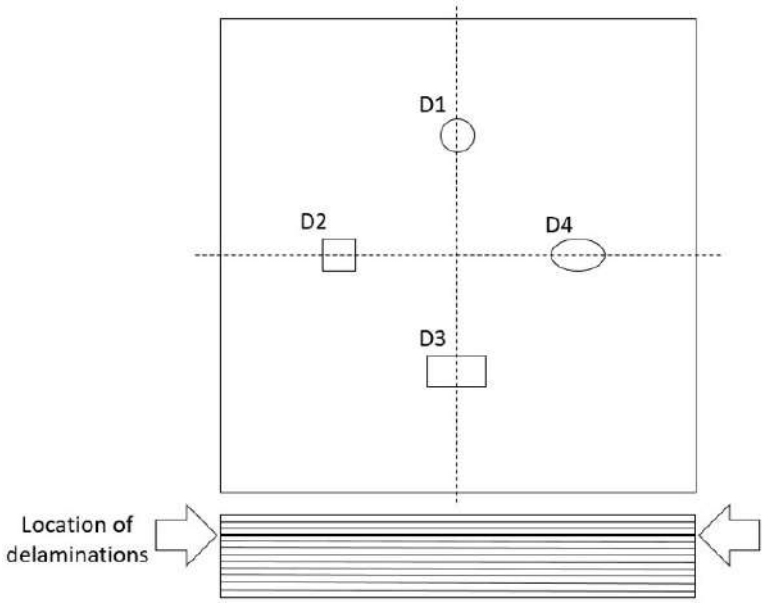
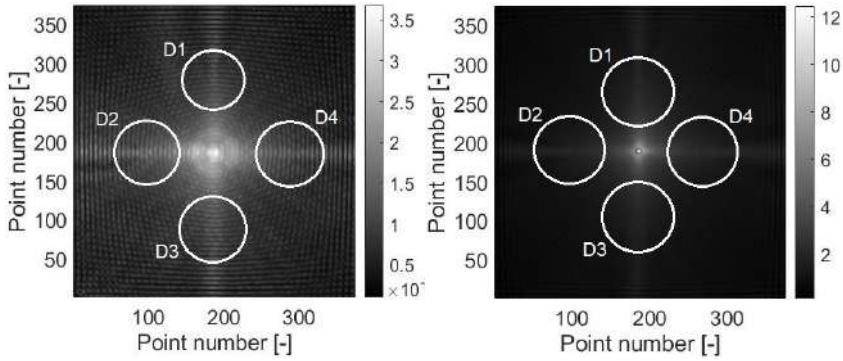


Fig. 5.12: The scheme of considered GFRP plate-like structure.

Comparison of the ACT and PZT performance

In order to compare the results, it was decided to check the considered defects detection ability using the ACT and the PZT. However, it is worth to note that the accuracy of detection of these defects is also referred to the advancement of signal processing ap-

proach. At the beginning, signal energy maps were prepared based on the RMS approach (formula (3)). The results in case of the usage of 1 ACT adjusted perpendicularly to the plate and 1 PZT bonded to the plate in its middle point are presented in the Fig. 5.13. After the measurement grid implementation, the area is related to the matrix of 375x375 points (sometimes the matrix is smaller, as some data rows or columns of the matrix are cut before RMS due to errors caused by the measuring device).



(a) ACT. (b) PZT.

Fig. 5.13: RMS energy maps.

As it can be noticed in Fig. 5.13, regions with delaminations are visible in the RMS maps in the form of disturbances of energy distribution but relatively low signal energy can be observed there compared to the middle plate zone (elastic waves excitation zone). Delaminations are therefore barely visible, except their average locations in the RMS maps. The defects are more noticeable in the ACT detection case.

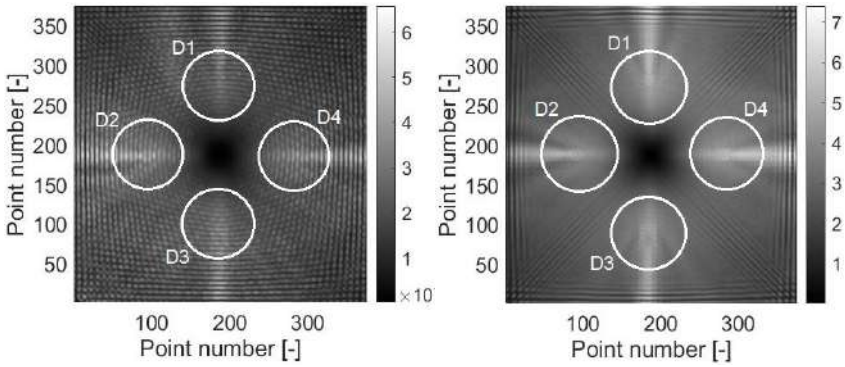
To ease delaminations detection, the RMS maps should be further processed, thus allowing for the recognition of their shapes, accurate size and location. In aim to do that, the Gaussian root mean square (GRMS) approach is proposed. Its principle rely on combination of the Gaussian distribution function with the usual RMS signal energy calculation in each measurement point of the grid. The Gaussian distribution function (G) can be defined by the following

equation:

$$G = e^{-\frac{(x-\mu)^{b_1}}{\sigma^{b_2}}}, \tag{5.4}$$

where: x – vector of the numbers of measurement points, related to the grid, μ – number of the middle point of x, σ – standard deviation, b_1, b_2 – parameters of the Gaussian distribution function.

Before the multiplication of the function G with the signal energy in each grid point of RMS map, the function G was inversed. Due to the fact that considered RMS maps are two-dimensional, the multiplication took place two times (one for each dimension). The Fig. 5.14 presents the results of the GRMS application in GFRP plate case. Again (to compare the results), performance of the ACT was compared to the PZT.



(a) ACT. (b) PZT.

Fig. 5.14: GRMS maps.

As it is visible in Fig. 5.14a, GRMS gave desired results - all four delaminations are now finely detected in terms of their location, though not such finely in case of their size and shape. The usage of PZT (Fig. 5.14b) resulted in more accurate detection of delaminations. It can be found within Fig. 5.14 that contact elastic waves excitation ensures higher signal amplitudes than their non-contact generation. Such situation results in sharper overview of the damaged areas, helping in the accurate estimation of shape and size of

the defects.

Evaluation of different ACTs arrays

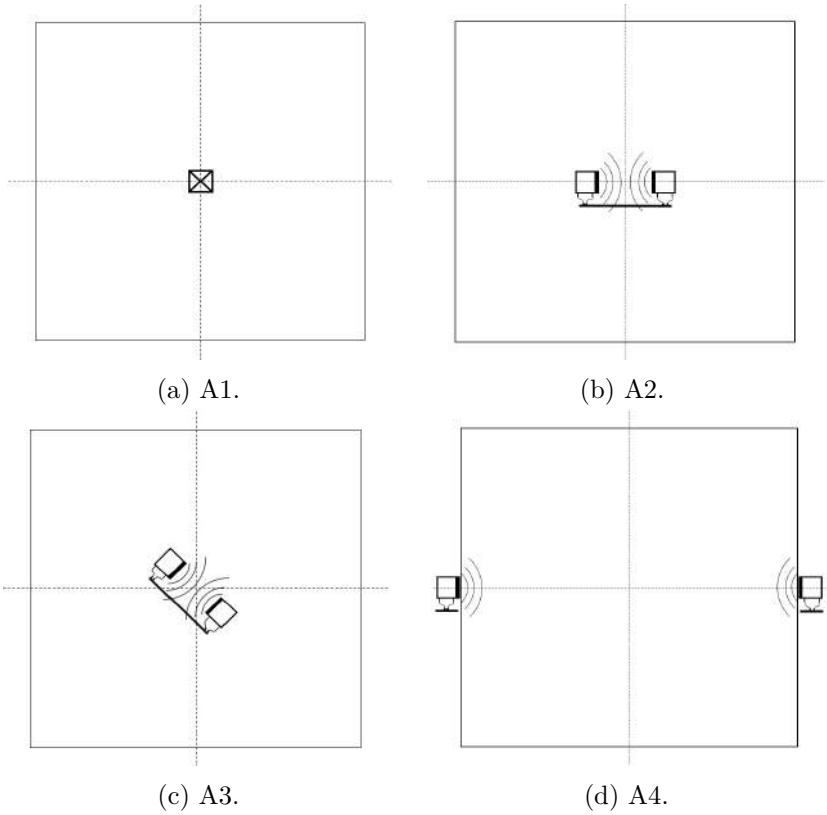
After the results comparison of the usage of single ACT and PZT, it was decided to analyse the performance of different ACTs arrays from viewpoint of delamination defects detection. ACTs used in the arrays are the same. The schemes of arrays adjustment in relation to the specimen are depicted in the Fig. 5.15. The following ACTs arrays (A) are investigated:

- A1: 1 ACT mounted in front of the middle zone of the plate, set perpendicularly (the slope of 0°) to its surface,
- A2: 2 ACTs coupled in the array via putting them on the cartoon stand, mounted horizontally in front of the middle plate region and aimed in the middle plate region direction (figure **??**a),
- A3: 2 ACTs coupled in the array via putting them on the cartoon stand, mounted diagonally in relation to the plate fibres in front of the middle plate region and aimed in the middle plate region direction,
- A4: 2 ACTs bonded to the wooden pieces, coupled to the considered specimen within the middle zones of the left and right plate borders using copolymer thermoplastic glue (Fig **??**b).

The arrays will be compared, taking the following criteria into account:

- the coverage area (0-10 scale) - based on RMS results,
- maximum signal energy value (unitless) - based on RMS results,
- delaminations detection ability (description) - based on GRMS results.

The grade of the coverage area criterion is referred to the analysis of histograms – the graphs which contain information about the elastic waves energy distribution within the measurement grid. The grade 0 means that 0% of the considered area is finely covered by elastic waves. In turn, the grade 10 means that the whole area is



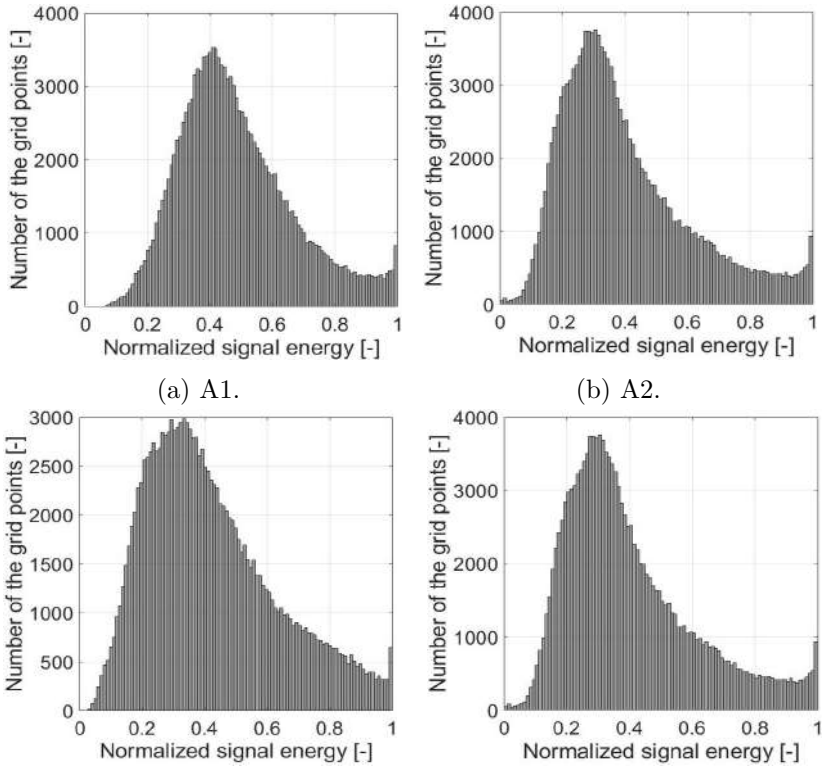
(a) A1. (b) A2.

(c) A3. (d) A4.

Fig. 5.15: ACTs arrays considered during the research.

finely covered. Before the evaluation of this criterion, the values of signal amplitudes were normalized separately for the each ACT arrangement (scale 0-1), in order to compare the performance of different arrangements. In each case, the normalization was done by referring signal energy value noted in each grid point to maximum signal energy value registered in considered measurement series. Histograms for the arrays are shown in the Fig. 5.16.

Higher profile in Fig. 5.16 means that elastic waves with relatively high signal amplitudes reach the bigger measurement grid part. Hence, it can be claimed that the coverage area is similar in case of A2 and A4 arrays, though the ACTs locations differ.



(a) A1. (b) A2.

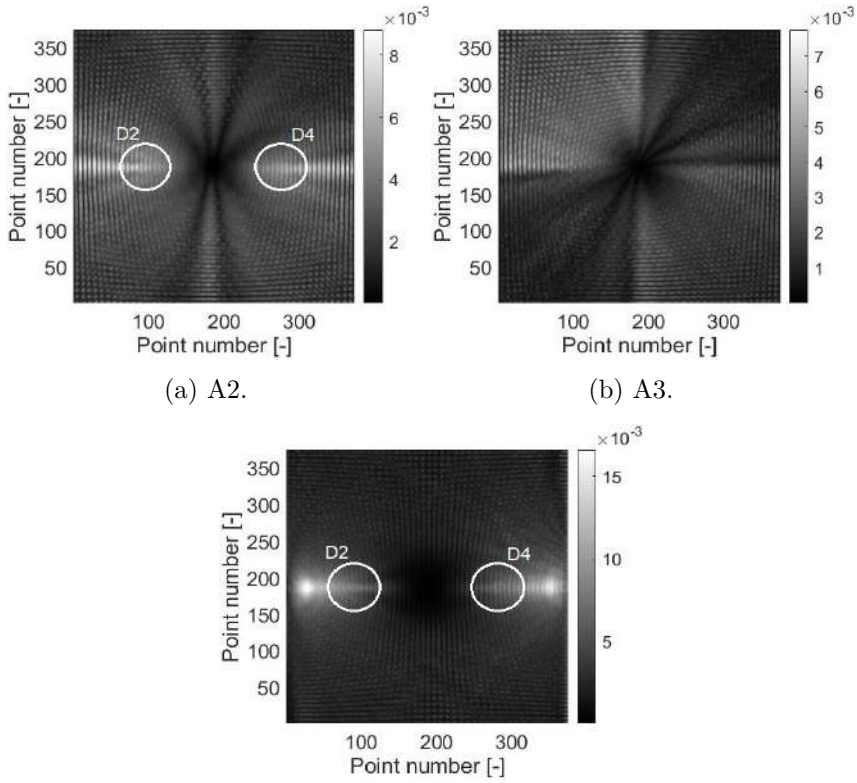
(c) A3. (d) A4.

Fig. 5.16: Histograms related to the coverage area achieved by ACTs arrays.

Another criterion which is considered while comparing the arrays performance is delaminations detection ability. This ability is described for the each array separately, based on GRMS results. These results are presented for the considered arrays in Fig. 5.14a (A1) and in the Fig. 5.17 (A2, A3, A4).

As it was aforementioned, analysed ACTs arrays are desired to be compared. The comparison is thus presented in the Tab. 5.1.

Comparing the arrays achievements, it can be found that the array A1 reached the highest coverage area, despite the fact that it con-



(a) A2.

(b) A3.

(c) A4.

Fig. 5.17: GRMS results for considered ACTs arrays.

tains only one ACT. It turns out that amount of the used ACTs does not strictly influence the coverage area – elastic waves generated by different ACTs may experience the negative interference, making the signal amplitudes lower. The array A1 is also the best in terms of delaminations detection. However, higher maximum signal energy was achieved by the arrays A2, A3 and A4 (some of the waves generated by different ACTs interfered positively).

Tab. 5.1: Comparison of the ACTs arrays performance.

ACTs array	**Coverage area**	**Maximum signal energy**	**Delaminations detection ability**
A1	6	0.0037	D1, D2, D4 – location, estimated size D3 – estimated location, estimated size
A2	4	0.0049	D2, D4 – location, estimated size
A3	5	0.0069	No delaminations detected
A4	4	0.0056	D2, D4 – estimated location

5.5 Summary

Elastic waves propagation analysis was performed in CFRP and GFRP specimens in the current research. The focus was put on the waves which are generated in non-contact way using ACT. In addition, the interactions of guided waves with different defects were considered.

Considering the bigger CFRP sample, the measurements along line were performed with the bigger accuracy compared to the previous investigation (ACT slope angle steps equalled 1° instead of 5°[1]), in aim to check the correctness of the gained results. In general, the results are similar to these obtained previously, though there exist small differences (they were precisely specified in the chapter). For instance, courses of the curves depicting signal energy values calculated for different measurement points (or for different ACT slope angles) are smoother and in some cases they are even slightly different. But, it was claimed that such situation is caused by i.a. slightly disparate adjustment of the specimen in relation to the ACT in different measurement series, though big effort was put to mount it in the same way. The ACT slope angles for which the highest signal energy values were determined (i.e. the optimal slope angles) differ a little between the measurement series, however these differences result from the fact that the measurement grid density (the angle steps) varied. In case of signal energy values determined in different measurement points, they also varied among the measurement series. It can be caused by different amount of the points located along the measurement line (165 points in

1°-step case instead of 127 in 5°-step case, assuming the same step length). Elastic waves lose more energy if they propagate to more distant points.

In terms of the smaller CFRP sample, three additional mass-type defects simulating the real damage were detected by the ACT in sloped/non-sloped manner. The most accurate detection was noticed while using the sloped ACT (with the optimal angle) in across-fibres case. Obviously, it does not mean that this adjustment of the ACT in relation to the plate always results in the most efficient defects detection in CFRP plate-like structures. The adjustment offering the most efficient defects detection differ, dependent on the defects locations and types. What is worth to be mentioned – when the ACT was sloped in direction opposite to the defects locations, it was still possible to detect the defects.

The elastic waves propagation in GFRP specimen differed from their propagation in CFRP samples, mainly due to different orientation of the fibres. It was found out that in terms of the fibres the factor which strongly influences the elastic waves propagation is their orientation. Energy losses of elastic waves during their propagation are smaller if the waves propagate along the fibres, not across them. Due to cross-type fibres orientation in GFRP specimen, elastic waves generated by the non-sloped ACT (i.e. the ACT set perpendicularly to the plate) propagated in horizontal and vertical direction with similar energy. However, the delamination defects detection is significantly more difficult than the mass-type defects detection in CFRP specimen. It is caused by the fact that considered delaminations cause local changes in the specimen structure in smaller magnitude compared with the changes introduced by the mass-type defects. Hence, it was decided to implement more advanced signal processing approach than RMS - GRMS. It improved the situation, however to detect delaminations more effectively it would be necessary to use either another, more advanced signal processing algorithm or more powerful ACT which is able to generate elastic waves with higher signal amplitudes. The usefulness of the second way of improvement (more powerful ACT) is visible while comparing performance of the ACT with the PZT. Due to

contact manner of guided waves generation in the PZT case, the generated waves have higher signal amplitudes. It resulted in much more accurate delamination defects detection. Besides, comparing selected ACTs arrays from viewpoint of the effective elastic waves propagation and delaminations detection, it turned out that the best overall performance was achieved when just one ACT is used. On the first fleeting glance it was surprising, as one would expect better performance while using more ACTs. The important factor influencing the performance of each ACTs array is the magnitude of positive/negative interference of the waves generated by ACTs.

In terms of the future recommendations it is desired to check the performance of more powerful ACT/ACTs in the considered CFRP and GFRP plate-like structures. The performance of ACTs with different base excitation frequencies from viewpoint of the generated elastic waves also could be checked. Moreover, more advanced signal processing algorithm than GRMS is desired to be developed and tested. All of these issues may help in more efficient detection of different defects. It would also be challenging to perform the whole analysis in case of the structures with the shapes other than plate-like shape, with the defects different than mass-type defects and Teflon-simulated delaminations. For instance, it would be valuable to focus on detection of real delaminations.

Acknowledgements

The author would like to gratefully acknowledge the support given by the National Science Centre, Poland under the grant agreement no. 2018/31/B/ST8/01298 in the frame of OPUS project, entitled: "Phenomenon of conversion of the acoustic waves to elastic waves due to interactions with solids".

Bibliography

[1] Damian Mindykowski. *Analysis of non-contact elastic waves generation using air-coupled transducer.* Chapter in Wybrane zagadnienia inżynierii mechanicznej II, edited by M. Mieloszyk and T. Ochrymiuk. Monography, Wydawnictwo Instytutu Maszyn Przepływowych PAN, 2021.

[2] Damian Mindykowski. *Overview of non-destructive testing methods.* Chapter in Wybrane zagadnienia inżynierii mechanicznej, edited by M. Mieloszyk and T. Ochrymiuk. Monography, Wydawnictwo Instytutu Maszyn Przepływowych PAN, 2020.

[3] Jian Cai, Lei Qiu, Shenfang Yuan, Lihua Shi, PeiPei Liu, and Dong Liang. Structural health monitoring for composite materials. *Composites and Their Applications*, 2012.

[4] Ling En Hong and Ruzairi Hj Abdul Rahim. A review on ultrasonic guided wave technology. *Australian Journal of Mechanical Engineering*, 2017.

[5] Tomasz Wandowski, Damian Mindykowski, Pawel Kudela, and Maciej Radzienski. Analysis of air-coupled transducer-based elastic waves generation in cfrp plates. *Sensors*, 21, 2021.

[6] Wu Qiao, Zhou Yuxuan, Lian Guoxuan, and Wang Xiaomin. Air-coupled ultrasound acoustic field measurements using laser doppler vibrometers. *State Key Laboratory of Acoustics, Institute of Acoustics, Chinese Academy of Sciences*, 2016.

[7] Xianwen Hu, Ching-Tai Ng, and Andrei Kotousov. Numerical and experimental investigations on mode conversion of guided waves in partially immersed plates. *Measurement*, 190, 2022.

[8] Leslie Bustamante, Nandu Jeyaprakash, and Che-Hua Yang. Hybrid laser and air-coupled ultrasonic defect detection of aluminium and cfrp plates by means of lamb mode. *Physics*, 19, 2020.

[9] Chunfang Rao and Lingze Duan. Bidirectional, bimodal ultrasonic lamb wave sensing in a composite plate using a polarization-maintaining fiber bragg grating. *Sensors*, 19, 2019.

[10] Homin Song and Yongchao Yang. Accelerated noncontact guided wave array imaging via sparse array data reconstruction. *Ultrasonics*, 121, 2022.

[11] Jakub Spytek, Lukasz Ambrozinski, and Lukasz Pieczonka. Evaluation of disbonds in adhesively bonded multilayer plates through local wavenumber estimation. *Journal of Sound and Vibration*, 520, 2022.

[12] Lv Ming Zeng, Baoding Wang, Xu Liu, Junwei Wu, Lijun Deng, Maodan Yuan, Yan Chen, Yu Deng, Yongkang Zhang, and Xuanrong Ji. High-resolution air-coupled laser ultrasound imaging of microstructure and defects in braided cfrp. *Composites Communications*, 28, 2021.

[13] Zhenhua Tian, Lingyu Yu, and Cara Leckey. Delamination detection using guided wave phased arrays. In *Proceedings*. AIP Conference, 2016.

[14] Tomasz Wandowski, Pawel Malinowski, and Wieslaw Ostachowicz. Analysis of s0/a0 elastic wave mode conversion phenomenon in glass fibre reinforced polymers. In *Proceedings*, volume 188. MATEC Web of Conferences, 2018.

[15] Ganggang Sha, Maciej Radzienski, Rohan Soman, Tomasz Wandowski, Maosen Cao, and Wieslaw Ostachowicz. Delamination imaging in laminated composite plates using 2d wavelet

analysis of guided wavefields. *Smart Materials and Structures*, 30, 2021.

[16] Antonino Quattrocchi, Fabrizio Freni, and Roberto Montanini. Air-coupled ultrasonic testing to estimate internal defects in composite panels used for boats and luxury yachts. *International Journal on Interactive Design and Manufacturing*, 14, 2020.

[17] Ali Mardanshahi, Mahmood Shokrieh, and Siavash Kazemirad. Simulated lamb wave propagation method for nondestructive monitoring of matrix cracking in laminated composites. *Structural Health Monitoring*, 2021.

[18] Zenghua Liu, Hongtao Yu, Junwei Fan, Yanan Hu, Cunfu He, and Bin Wu. Baseline-free delamination inspection in composite plates by synthesizing non-contact air-coupled lamb wave scan method and virtual time reversal algorithm. *Smart Materials and Structures*, 24, 2015.

[19] Tomasz Wandowski and Pawel Malinowski. Full wavefield method for damage detection in stiffened cfrp panel. In *Proceedings*, volume 11593. SPIE Smart Structures + Nondestructive Evaluation, 2021.

[20] Yamnesh Agrawal, Akhilendra Singh, and Dhanashri Joglekar. Mixing of lamb waves at delamination defect in a unidirectional gfrp laminate. *Advances in Non-destructive Evaluation*, 2021.

[21] Alessandro De Luca, Donato Perfetto, Angelo De Fenza, Giuseppe Petrone, and F. Caputo. Guided wave shm system for damage detection in complex composite structure. *Theoretical and Applied Fracture Mechanics*, 105, 2020.

[22] Chao Zhang, Zhenyu Zhang, Hongli Ji, Jinhao Qiu, and Chongcong Tao. Mode conversion behavior of guided wave in glass fiber reinforced polymer with fatigue damage accumulation. *Composites Science and Technology*, 192, 2020.

[23] Pawel Kudela, Maciej Radzienski, Piotr Fiborek, and Tomasz Wandowski. Elastic constants identification of fibre-reinforced composites by using guided wave dispersion curves and genetic algorithm for improved simulations. *Composite Structures*, 272, 2021.

Chapter 6

Thermal loading process on additive manufactured composites

Isyna Izzal Muna

Institute of Fluid Flow Machinery, Polish Academy of Sciences,
Fiszera 14, 80-231 Gdansk, Poland

AM Additive Manufacturing

CCF continuous carbon fiber

CFRP carbon fiber reinforcement polymer

FBG fiber Bragg grating

FDM Fused Deposition Modeling

FEM Finite Element Method

PLA PolyLactic Acid

MD multidirectional

SCF short carbon fiber

Tg glass transition temperature

Tm melting temperature

UD unidirectional

6.1 Introduction

The technological advancements in the area of Additive Manufacturing (AM) over the last decade have been a major drive for the development of composite materials. It is widely known that AM technologies specifically Fused Deposition Modeling (FDM) method has been utilized in many sectors owing to its flexibility and simplicity with small amounts of waste material. The numerous benefits of various AM techniques indeed open up endless possibilities in engineering and industrial fields. However, there are still some inevitable manufacturing constraints in AM such as porosity, shrinkage, anisotropic material properties, thermal distortion, and internal cracks [1–5]. These defects are affected by various printing process and different printing parameters which can have a significant impact on the final part's quality and safety. A massive research studies have been conducted in attempt to overcome printing process issues and improved mechanical properties on printed structures by incorporating various printer properties, processing parameters, types of reinforcement, and the composite structures [2, 6–14]. These experiments are typically performed in pre-processing or

during processing stage. In normal operational conditions, the printed composite materials are more likely exposed to various environmental conditions, specifically related to temperatures. Exposure to different temperature magnitudes and time period will affect its mechanical strength and morphological structure differently. Researchers have not explored the thermal degradation of additive manufactured polymer composite comprehensively yet it is an important phenomenon that will be useful in the structural health prediction of lightweight material. According to studies, the thermal degradation process is affected by the material, the environment, and the heat source [15]. The damage caused by at least one of subjected temperature conditions such as continuous exposure to high temperature, cyclic temperature, and local-global temperature on polymer composite material is referred to as thermal damage. Over time, an irreversible thermodynamic process caused by heat damage on printed composites will lead to thermal degradation which may result in structural deformation and/or chemical changes of these composite systems [16]. In an attempt to observe thermal damage on printed polymeric composite, many research works have been conducted in defect detection methods using advanced technologies since it is fundamental to observe thermal failures on printed samples and their mechanical qualities at early stage [17–19]. Understanding defects in FDM printed materials and their causes indicate the first step to inspect and eliminate them.

Among the afore-mentioned pre-treatment methods for strength improvement, various post-processing treatments through thermal annealing (heat-treatment) are also becoming widely applied in 3D printed composite materials [20–24]. Following the thermal treatment for improving the mechanical properties of composites, it is highly possible that materials also undergo thermal degradation. Yet, so far the majority of studies have focused only on improving strength by annealing and/or studying the thermal degradation of composite materials without the intent to enhance the properties. There is still a limited investigation into thermal degradation after the annealing process. Understanding thermal degradation not only on untreated materials but also on treated ones will open up a new horizon in the life and failure prediction of composites.

This chapter describes a variety of temperature loadings subjected to AM polymer composites to understand the difference in each temperature condition process. Also, the influence of temperature treatment on mechanical qualities and internal structures of printed composites employing testing and detection technologies will be studied to anticipate composite failures. By the end of this chapter, brief analyses of work in progress about the mechanical strength of additive manufactured samples with unidirectional (UD) alignments of the fibre reinforcement. In future, two groups of samples will be analysed: intact (after AM process) and damaged (with introduced damage (e.g. delamination, notch, crack)). The samples will be exposed to thermal loading with different parameters: magnitude of temperature (negative, positive), exposition time (continuous, elevated, cycle), and sample area (global/local).

6.2 Thermal loading characterization

In previous book chapter, thermal degradation of printed polymeric composites has been discussed [25]. Understanding various environmental factors depending on its exposure type, will also help to characterise the material damage. In case of solid materials such as polymeric composites, the damage creation is influenced by various variables/factors related to temperature such as selected exposure of sample area (local and global), temperature exposure time or duration of subjected temperature (elevated, stable continuous, cyclic), thermal magnitude (positive, negative), heating or cooling rate, and the application heat mode (conduction, convection, radiation).

Nowadays, as the most promising material, PolyLactic Acid (PLA) is widely used in FDM 3D printing technology. Based on a variety of applications of 3D-printed parts, polymeric composited may have been exposed to varying temperatures during their service life. The mechanical performance of thermoplastic materials can be improved by changing polymer's molecular structure via thermal annealing [26]. Thermal annealing is a process that involves heating the polymer to a temperature above the glass transition temperature and then slowly cooling it back to room temperature [22].

There has been a numerous number of research related to post-processing techniques for reducing defects, enhancing the cross-linking state and crystallinity of molecular chains in AM polymeric composites as well as to improve the mechanical properties of manufactured specimens [20–24, 26]. This annealing treatment enhances the mechanical properties of semi-crystalline polymers, particularly those with low crystallinity, such as carbon fiber PLA. Crystallinity is a significant component impacting the strength of a semicrystalline material; this is due to intrinsic polymer properties and its thermal history, such as cooling kinetics in thermal processing [27]. The ideal temperature for annealing process must be above the glass transition temperature and below the melting point to obtain changes. The deformation of specimens can be analyzed after annealing or heat-treatment experiments. Technological heating devices which are utilized by majority of scientist for the heating treatment of composites are electrical oven, microwave, temperature/environmental chamber, and hot plates [18, 22, 28]. The post-processing of polymeric composites using microwave energy has been less explored [28]. Heating mechanisms in microwave processing depends on the type of target material, for e.g., nonmagnetic, or magnetic. Most polymeric materials have a low dielectric loss factor and are not affected by microwaves. This is a significant challenge in microwave processing of polymeric systems, as it requires the addition of suitable phases, such as fibers as reinforcement or susceptors. The polymer matrix composites are cured or moulded in the low-temperature area (less than 500°C) during microwave process. Uniform heating distribution is expected to occur during microwave processing of polymer-matrix composites (PMCs) as compared to the conventional heating techniques due to its volumetric heating characteristics. To ensuring uniform microwave heating, temperature control in curing process is necessary. The major challenge in application of microwave curing technology is the uneven temperature distribution on the composite surface due to uneven EM field distribution inside the applicator [28].

In polymeric composites, thermal loads cause the alternating strains at the different scales (see Fig.6.1) [29].

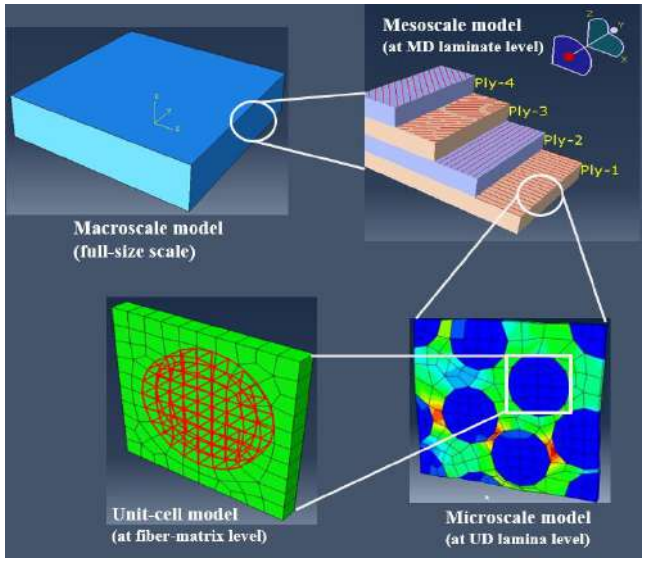


Fig. 6.1: Multiscale of composite laminate.

There are two possible causes of damage in a thermally loaded carbon fiber reinforced composites: (1) thermal-induced stresses at the microscale level and (2) thermal-induced stresses at the macro level. Thermal expansion of the materials 'carbon fibre' and 'polymer matrix' differs at the fiber-matrix-scale (microscale level) and causes strains in the fiber, the matrix, and the interface between them. The anisotropic thermal expansion behavior of the UD composite ply is caused by the various thermal expansion rates of the composite elements. In a UD laminate, only the micro level effect is present, but both effects are stacked in a multidirectional (MD) laminate. In MD laminate (macro level) consisting of several UD plies with different fibre orientation, this results in thermally induced stresses in the individual plies and at the interfaces, as the different plies aim for different thermal expansion in the different material directions. At the fibre-matrix-scale (micro level), stresses are induced in the fibres, in the matrix and in the fibre-matrix-interface due to the different thermal expansion of the materials 'carbon fibre' and 'polymer matrix'. The different thermal expan-

sion of the composite constituents leads to an anisotropic thermal expansion behaviour of the UD composite ply.

In the following subsections, various type of temperature loadings which can be applied on printed polymer matrix composites will be characterized. Damage induced by temperature loading must be described in terms of many elements, such as material properties, environmental impact, degradation process, and damage characteristics [15].

6.2.1 Continuous heating

Continuous heating is defined as temperature exposition process onto materials at a stable temperature for a certain period of time, ideally for a few hours. Typically, specimens are placed at preheated chamber in which desired temperature is hold constantly for a certain period of time.

Global heating

In global heating, polymeric composites are conditioned inside the environmental instruments to receive temperature exposure at all area of material. The typical annealing temperature is between above the glass transition temperature (Tg) and below the melting temperature (Tm) of polymer. In case of PLA materials after printing process at the applied conditions reached only a partial attainable crystallinity degree, therefore in order to increase the degree of crystallinity post-processing annealing at elevated temperature, above Tg and below Tm, was employed [22, 27]. The ideal duration for thermal annealing is at least 30 minutes to attain a certain degree of crystallinity [30].

The experimental investigation by Zhou et al. revealed that thermal annealing by means of microwave in unidirectional carbon fiber reinforced polymer (CFRP) laminates with 0° ply angles, fibers aligned parallel to electric field do not absorb microwave energy and behave as a reflective material [31]. Whereas for 90° ply angle, maximum microwave energy is absorbed by the perpendicularly aligned carbon fibers to electric field and it mainly comes from conduction loss in the carbon fibers. Based on some studies, microwaves do

not affect the curing reaction, but rather enhance the rate of reaction by volumetric heating [28]. Gonzalez et al. proposed post-processing temperature optimization on CFRP coupons via compaction method in order to provide the solution for the typical drawbacks of FDM printing technique [18]. Sample coupons were post-processed in the hot plates at 8 different temperatures ranging from 70°C to 270°C during 15 min (heating speed of 10°C/min). The post-processing temperature effects on microstructure, thermal stability and interlaminar properties of the printed CFRP were analysed. When taking into consideration the dimensional precision, the microstructure, the thermal properties, and the mechanical properties of the final products, the ideal temperature for post-processing is 150°C.

A number of research works have investigated the effect of heat treatment of 3D printed polymeric composites using electrical oven. Nassar et al. studied the comparison between annealed and non-annealed specimens and it was revealed that a noticeable fusion of filaments results in eminent contact between layers [21]. With annealing, polymer filaments are reheated such that the crosslinking of filament surfaces are improved. Another investigation stated that mechanical properties of polymer composites is improved by perfecting the inter-facial bonding since it will lead to an increase in the density [12].

According to the study finding by Wach et al., only the annealed FDM-PLA samples have a significant crystalline part and show diffraction patterns of crystallites, whereas the untreated samples have a wide amorphous halo without crystal peaks [27]. Once the heat treatment is done, samples are being evaluated with three-point bending (flexural test) and tensile test. The degree of crystallinity of the treated polymer at 215°C is increased by subjecting it to a thermal annealing process that occurs at a temperature higher than its glass transition temperature. The flexural stress of the samples is increased by 11-17% as a result of an increase in the degree of crystallinity of FDM-PLA. As a post-treatment to improve the specimen's mechanical properties, thermal annealing was performed by Arjun et al. at 65°C, 95°C, 125°C, and 155°C for 30, 60,

120, and 240 min [22]. An excellent optimized annealing temperature and duration were achieved for 95°C and 120 min respectively which resulted in increased tensile strength by 14%.

The influence of temperature on different structure and material reinforcement type such as sandwich composite and short carbon fiber reinforced composite have also been reported [20, 32]. Ivey et al. studied the impact of short carbon fiber fillers on the mechanical characteristics of 3D-printed PLA and the effects of the annealing process on polymer crystallinity and mechanical properties were examined [20]. It was shown that annealing increased the crystallinity of both sample groups, but had no statistically significant effect on the mechanical properties. According to the tensile properties of PLA and PLA/carbon fiber(CF) filaments, the addition of carbon fibers to the PLA filament significantly increased the samples' elastic modulus. However, according to study by Jiayi Liu et al. for sandwich structure with pyramidal truss core at higher temperature, especially when the temperature is above the glass transition temperature, the matrix gets softened and cracked, and the fiber-matrix interface becomes weakened [32]. The weakened interface makes the fibers debond easier from the matrix. The bending failure load decreased as temperature increased, which was caused by the degradation of the matrix properties and fiber-matrix interface properties at high temperature. With an increase in temperature, the bending failure load reduced, and it decreased by 92.27% when the temperature went from 20°C to 200°C, mostly because the matrix characteristics and fiber-matrix interface properties degraded.

Local heating

Uniform heat damage of polymeric composites have been studied by a large number of researchers yet the exploration of thermal damage in a localised area is still very limited. Typically, heat damage can occur as a result of a fire or during service operations where they are exposed to hot gasses, lightning strikes, jet engine flux during landing, and jet engine exhaust. The majority of these degradation occur in small locations and may go unnoticed. Thus, understanding and characterizing such damage is critical in order to avoid catastrophic collapse [15].

Local heating or also known as short/pulse heating can be applied on the material surface at specific area. Recently, Addepali et al. investigated the degrading process that occurs during the exposure of a localized area to extreme temperatures, as well as the laminates' response to such situations [15]. Non destructive testing was employed to measure thermal diffusivity of CFRP composites when heating process was exposed in a localised area. A hot air gun was utilized as the heat source. Both pulsed thermography and ultrasonic testing (UT) were able to identify the presence of physical and surface breaking damage. Although it is possible to evaluate the resistance of the laminate to the effects of physical damage, it is an extremely difficult task to spot discoloration or the beginning stages of damage on a large structure. Yao Ma et al. carried out an experimental work in locally rising temperature of braided CFRP and its damage morphology using long pulse laser with different energy densities and pulse widths [33]. The investigation revealed that under the same pulse width, the time required for the braided CFRP surface to reach maximum temperature decreases with laser energy density, rapidly in the low energy density region but slowly in the high energy density range. With a wider pulse width, the braided CFRP surface heats up more slowly. When laser pulse width and energy density are determined, the time for the CFRP surface to attain maximum temperature increases with target thickness.

6.2.2 Cyclic modulated heating

The application of polymeric composite parts in aviation and automotive industries more likely undergo temperature fluctuations between -50°C and 130°C [34]. Structural materials exposed to this extreme environment may degrade over time, and cyclic temperature influence on mechanical behaviour of polymeric composites should be investigated. Thermal cycling can cause fatigue failure, which is a gradual kind of local damage [35]. The process of thermal cycling can lead to the onset of fatigue failure, which is a sort of gradual local damage. As a result of the fiber and matrix shrinking at different rates, thermal cycling produces thermal re-

sidual stresses. In composite laminates, residual stresses can lead to delamination, warping, and fatigue failures. Low heating and cooling rates are required to prevent the material from being subjected to thermal shock loading; nevertheless, this results in longer cycle times [29].

Parameters like temperature limits, dwell time at both temperatures and change rate between the limits are critical in thermal cycling processes [36]. Temperature limits are essential for the acceleration testing level. The larger the disparity between the boundaries, the greater the tensions they induce. However, if the limitations are excessively stringent, there is a significant chance that over-stress failures will occur, resulting to premature failures that would never occur under actual use conditions. The glass transition temperature, Tg, of polymer materials is an example of a typical critical limit. A temperature limit exceeding Tg may result in an easily observable catastrophic failure, but it may also alter the failure processes to be unrealistic. In addition, it is advantageous to set the temperature limit as high as possible since, if the difference between the limits is insufficient, the test has a very low acceleration factor and the test time is quite long.

Dwelling time means how long of an exposure period is appropriate at each temperature limit. If a prolonged dwell time is employed, the duration of the test will lengthen until the number of cycles is decreased. Then again, In some instances, a long dwell period can potentially speed up the testing process if it generates structural changes that increase the stresses during temperature fluctuations. At high temperatures, polymer materials, for instance, relax or creep - the polymer chains within the material move to alleviate the stresses imposed by the high temperature. When the temperature is decreased, these alterations may result in a large increase in the stresses imposed on buildings. However, sufficient dwell time is necessary for these modifications to take place.

As thermal cycling is very expensive and time-consuming, typically only a few (1-100) cycles may be realized [37, 38] and certain study groups achieved 200-5000 cycles [39–42]. The illustration of one cycle in heating process is shown in Fig.6.5. Ghasemi et al. studied

the effects of temperature and the number of thermal cycling cycles on residual stresses of polymeric composites at several stacking sequences during thermal cycling. One thermal cycle is a 70°C to 100°C then back to 70°C. The heating rate is 3-5°C every minute; dwell time at both extremes of temperature is 15 min. It was reported that the condition of thermal cycling leads to a decrease in the residual stresses and an increase in the failure index. The failure index results amount can be used to determine whether or not a material will fail under loading conditions. The material fails if the failure index is equal to or greater than 1. Guigon et al. studied mechanisms of degradation in carbon-epoxy three-dimensional woven composites subjected to temperature cycling [43]. It was found that iso-thermal aging revealed thermo-impact oxidation's on matrix elasticity. In addition, thermal cycling ageing in neutral (nitrogen) and oxidative (air and oxygen) atmospheres caused micro-cracks in the matrix. Zhang et al. investigated thermal cycling-induced between 55°C and 120°C in carbon/epoxy braided composites experimentally and numerically. It was reported that composites and resins had less than 1% mass and volume change after 160 cycles. However, micro-cracks have not been found in acoustic emission and X-ray CT tests. T700s/3502 acoustic emission (AE) events have a logarithmic time-versus-temperature curve and occur periodically at low temperatures. Most events occur below 0°C. The six-layer samples showed micro-cracking on X-ray CT. All three fiber directions had split fiber bundles [44].

6.3 Thermal loadings on AM CFRP Composites

As it has been discussed in previous section, there are various types of thermal exposure which might be subjected to additive manufactured parts specifically continuous carbon fiber reinforced polymer composites. A global heating of unidirectional (UD) continuous carbon fiber reinforced polymer (CFRP) composite has been experimentally and numerically investigated using non destructive testing by means of fiber Bragg grating (FBG) sensor by Muna et al [45].

The coupled thermal-displacement analysis was used during the numerical calculations. The strain in the sample due to its exposition on elevated temperature was measured using fiber Bragg grating (FBG) sensors. Furthermore, the author of this chapter has also performed an experimental work of global continuous and cyclic thermal treatment on CFRP composites as part of research internship program in Kaunas University of Technology, Lithuania. The experimental work and numerical simulation of thermal exposure of these composites will be briefly described in following subsection.

6.3.1 Experimental method

Additive manufacturing process and materials

The CFRP composite specimens were additively manufactured with fused deposition modeling (FDM) method with a modified 3D printer by Kaunas University of Technology [46]. The matrix agent was PLA 3D850 filaments from Natureworks with a tensile strength of 51 MPa, Young's modulus 2315 MPa, and density 1.24 g/cm^3. For the reinforcement material, continuous carbon fiber T300B-1000 from Toray was chosen with tensile strength of 3530 MPa, Young's modulus of 230 GPa, and density of 1.76 g/cm^3. The 3D printer and materials can be seen in Fig. 6.2. For the specimen subjected to global elevated (increased) temperature, the printing machine parameters are as follow: nozzle diameter is of 1.5 mm, printing speed is 3 mm/s, bed temperature is 70°C, extruder temperature is 200°C, extrusion multiplier is 0.6, and extrusion width is 1.2 mm. The dimension of unidirectional (UD) composite laminate is 92x95x2 mm with 6 layers of same fiber alignment in all layers. Another set of CFRP samples was printed for different thermal exposures (global continuous and cyclic temperature) with dimension of 150x13x2 mm. The printing parameters are shown in Table 6.1. There were 9 groups of printed CFRP samples for various thermal treatments as summarized in Table 6.2. In total, the number of specimens printed was 45 samples in which 5 samples for each group of thermal treatment were required in accordance with the D3039 ASTM standard used for tensile testing. After the manufacturing stage is completed, the composite samples then were measured in

weight and size dimension using digital scale and caliper, respectively. The weight and dimension measurement is necessary to be performed in order to compare the changes in size and mass before and after the thermal loading. Furthermore, it is important for the calculation of mechanical properties after the mechanical testing.

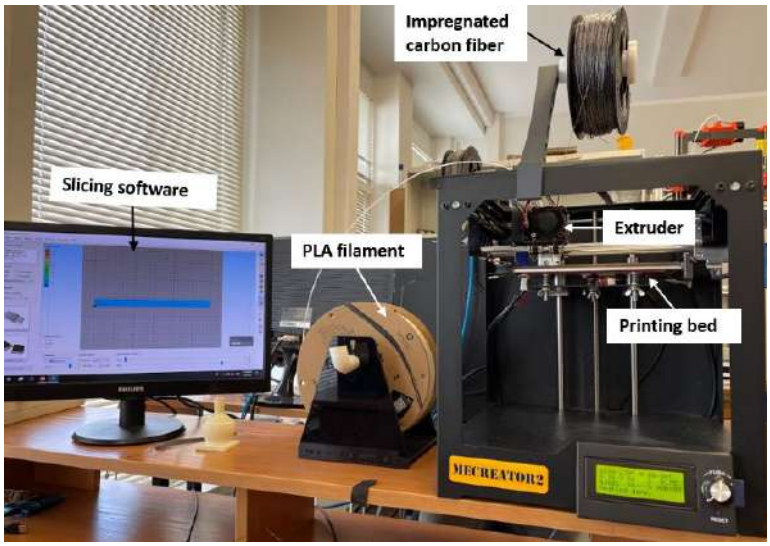


Fig. 6.2: 3D printer device of FDM technique with its matrix and reinforcement material.

Tab. 6.1: Printing parameters

Parameter	**Value**
Nozzle diameter	1.5 mm
Extrusion multiplier	0.7
Primary layer height	0.5 mm
Interior infill	100%
Infill pattern	rectilinear
Nozzle temperature	220°C
Bed temperature	90°C
Printing speed	3 mm/s

Tab. 6.2: Sample groups

Group name	Value
Intact	untreated samples
HS-A	hot stable at 65 °C for 6 hours
HS-B	hot stable at 145 °C for 6 hours
HC-A	hot cyclic between 50-70 °C with 6 cycles
HC-B	hot cyclic between 140-150 °C with 6 cycles
CS-A	cold stable at 0 °C with 12 cycles
CS-B	cold stable at -20 degC, 12 cycles
CC-A	cold cyclic between 0 to 5 °C for 6 hours
CC-B	cold cyclic between -20 to -15 °C for 6 hours

Global elevated temperature

One sample with specific geometrical and material properties was selected to be placed inside the environmental chamber. The chosen approach allowed a more detailed description of the processes that occur in FDM printed material. Each sample (even manufactured under the same printing conditions) can show some material specification differences, therefore in order to avoid the uncertainty and repeatability error that can occur due to manufacturing process. The sample temperature was set at elevated temperature started from 10 to 50°C with 5°C step. The uneven distribution of polymeric matrix and carbon fiber in the sample volume was caused by a difficulty during the AM process caused by the tiny thickness of the carbon fiber bundles. The fibers are also slightly rotated. It is well visible in the sample surface photograph in the part marked by the red ellipse. The sample is consisted of 6 layers and in the middle of the sample (between the 3^{rd} and the 4^{th} layer) fiber Bragg grating (FBG) sensor was embedded in order to examine the strain resulted during the thermal testing. The temperature of the sample

was measured using a FBG temperature probe and the FBG was used acquiring and processing signals from FBG sensors in operation the temperature. The photograph of the sample is presented in Fig. 6.3.

Fig. 6.3: Surface micrograph of the sample. C – carbon fibre bundle, M – polymeric matrix.

Cyclic hot and cold temperature

In attempt to investigate the thermal cyclic effect on 3D printed CFRP composite, the author also performed an experimental work of cyclic thermal treatment at positive and negative temperature as part of research internship program in Kaunas University of Technology, Lithuania. There were 2 sets of samples which were subjected to the cyclic positive temperature using a Universal Oven Memmert UNB-3000 as presented in Fig. 6.4. One group was set to have temperature range between 50 and 70°C (group HC-A) whereas another sample group has temperature range of 140 and 150°C (group HC-B). The number of cycles for each group is and dwelling time for the two extreme temperature was 10 minutes. The schematic figure for thermal cyclic at positive temperature is illustrated in Fig. 6.5.

For HC-A group the heating and cooling rate were 1°C/min and 2.5°C/min, respectively. The heating and cooling rates for the HC-B group were 0.8°C/min and 2.2°C/min, respectively. It must be noted that the difference in heating and cooling rate is due to the operational mechanism of the air-circulated oven. The oven ventilation uses natural convection type of heat transfer where the air circulation control is supplied by a continuous adjustment of fresh air in certain percentage of ambient air.

For the sub-zero (negative) thermal cycling treatment, an automated environmental chamber MyDiscovery DM600C was utilized for 2 sets of samples in which one group was set to have temperature range between 0 and -5°C and another group were set in a range between -15°C and -20°C with 12 cycles for each group. The cooling rate for treated samples in negative temperature was 0.8°C/min. The chamber device is shown in Fig. 6.6.

Fig. 6.4: An air-circulated oven for hot temperature treatment.

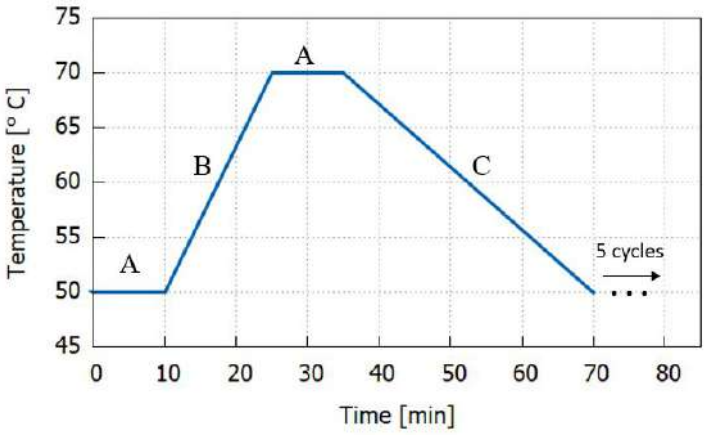


Fig. 6.5: Thermal cycling profile. A – dwell time; B – heating rate; C – cooling rate.

Fig. 6.6: An automated environmental chamber for cold temperature treatment.

Global continuous hot and cold temperature

Recently, the author has also performed an experimental work of stable continuous heating treatment at global area of CFRP composites. There were 2 sets of samples subjected to the prolonged positive temperatures, one group (HS-A) was exposed to 65°C and another group (HS-B) was exposed to 140°C. These temperatures were chosen in a such a way that it will be above the glass transition temperature (Tg) and below the melting temperature (Tm) of the matrix material used in the specimens. As for the continuous thermal treatment at zero and sub-zero temperature, two groups of samples were subjected to 0°C and -20°C, respectively; and both groups were kept for 6 hours in the chamber.

6.3.2 Morphological analyses

The micro-morphological analyses on the samples' surface were performed before and after performing the cyclic and stable thermal loadings by utilizing an optical microscope device (Nikon Eclipse LV100ND) equipped with a high-definition color camera (Nikon DS-Ri2). The imaging software (NIS Elements 4.5.1.00) was used to prepare and process the data at 5× magnification. The maximum sample size observed with the optical microscope was 150 × 150 mm. A scanning electron microscope device (FE-SEM SU5000) was employed to investigate the micro-structure damage of different specimen groups after tensile testing. The maximum specimen size observed with the SEM was 200 mm in diameter and 80 mm in height. A digital microscope (Levenhuk) was utilized to capture the macroscopic images of the tensile-fractured samples. The maximum specimen size observed with the digital microscope was about 150 × 100 mm.

It can be seen that there is no noticeable difference on structural surface before and after the thermal loading at lower temperature (Fig. 6.7). It appeared that the heat is not enough to change the visible surface structure of matrix material. However, a further investigation is needed to observe its internal structures such as degree of crystallinity and cross linking of polymer PLA. In order

to achieve the information of these analyses, differential scanning calorimeter (DSC) and thermogravic analyses (TGA) could be performed. As for the higher temperature, it can be seen from Fig. 6.8 that the morphological surface undergo some structural changes where the polymer PLA as matrix material has smoother and finer appearance after the heating treatment. This is due to the high temperature subjected to the CFRP samples which is close to Tm of PLA. Moreover, the specimens which undergo high temperature also appeared to be slightly deformed i.e. it shows some wrinkle shape along the longitudinal direction.

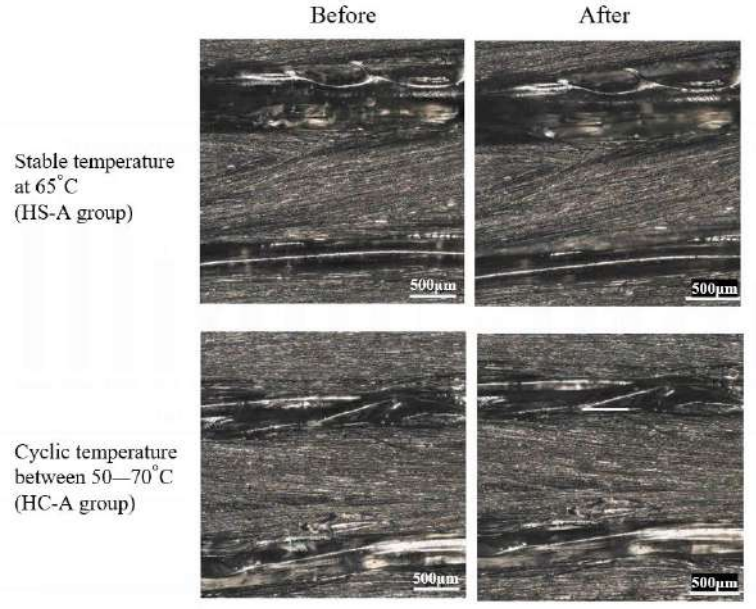


Fig. 6.7: Optical micrographs of sample group before and after stable temperature at 65°C and cyclic temperature between 50-70°C.

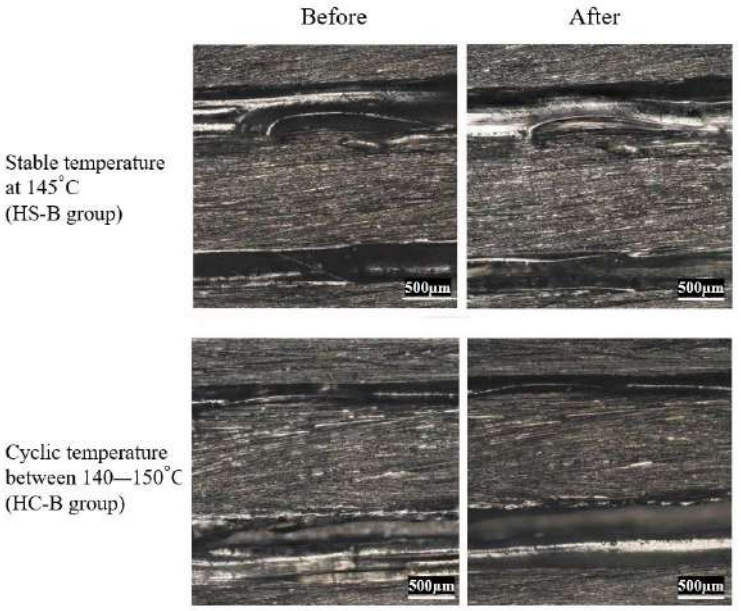


Fig. 6.8: Optical micrographs of sample group before and after stable temperature at 145°C and cyclic temperature between 140-150°C.

6.3.3 Numerical modeling

The numerical study was conducted to model the behavior of polymeric composites in relation to the effect of global elevated temperature on the 3D printed CFRP samples. From a macromechanical approach, the specimen was modeled as stacked solid elements at the laminate level and treated as a homogenous equivalent material. The laminate is composed of six-ply unidirectional (UD) lamina with stacking sequence (LSS) at 0 degree, and the constitutive material behavior of the laminate was treated as if it were a single orthotropic material. As symmetry planes parallel and perpendicular to the fibers, the array of carbon fibers in polymer matrix is assumed to be square. Macro-mechanical modeling was utilized to characterize the deformation and mechanical behavior of a material.

A finite element method (FEM) model was utilized to produce numerical results for a carbon fiber reinforcing composite system. The geometry of the composite model was constructed as a deformable homogenous solid using the computer-aided design software ABAQUS. The model geometry has a dimensional area of 92 x 95 mm and a thickness of 2 mm. The specimen model was discretized into an eight-node trilinear heat transfer brick hexahedral mesh with C3D8T element type. There were 4416 finite elements, 19266 nodes, and 249750 degrees of freedom in the FEM model. The simulated load scenario was divided into two stages: starting step and loading step, with temperatures ranging from 10°C to 55°C with a 5°C step. The relative humidity inside the chamber was expected to be steady at 20%. During the increased heating, the numerical modeling of CFRP composite employing the coupled temperature-displacement approach will be explored within thermo-mechanical strain analysis.

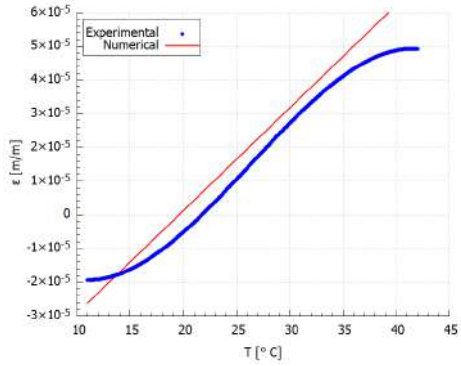


Fig. 6.9: Comparison curves determined experimentally and numerically. Taken with permission from [45].

Comparing the experimental and numerical values presented in Fig. 6.9, the percentage difference between them is calculated as ca. 18%, while the average strain difference is equal to 5.7x10^{-6} m/m. The disparity between the experimental and numerical curves is due to the variability of the AM sample structure, which is clearly seen in Fig. 6.3 (b). The difference between those values are due to some defects which the specimen possess. Some presented voids and

structure irregularity in the sample contributed to the inconsistency during the experimental measurement [14, 47]. Furthermore, the incorporation of reinforcements during FDM printing may increase the strength of the material system, but this benefit is outweighed by a poor reinforcement/matrix interaction, non-uniform reinforcement distribution, and inappropriate impregnation. These variables will result in the formation of new voids [14].

6.4 The mechanical loading on AM CFRP composites

6.4.1 Tensile testing

The tensile testing was performed subsequent to thermal treatment in order to examine the mechanical strength and Young's modulus of the treated specimens as well as untreated (intact) samples.

Fig. 6.10: Tensile testing instrument with infrared camera device.

Tilnius Olsen H25KT (capacity 25 kN) universal testing machine was utilized to perform tensile testing as shown in Fig. 6.10 with the standard head displacement rate of 2 mm/min. The ASTM standard of D3039 is used for the tensile testing which requires each tested specimen to have four tabs in gripping position at top and bottom. Prior to the tensile testing, PLA tabs having dimensions 50 x 12.5 x 2 mm were printed separately and four points were marked 15 mm from the center of the specimens in order to measure the elastic strain. The printed tabs then were adhered to specimens with universal structural bonder (adhesive glue) by applying pressure via clamps in order to hold and secure the parts tightly. The tabs and clamping process are shown in Fig. 6.11.

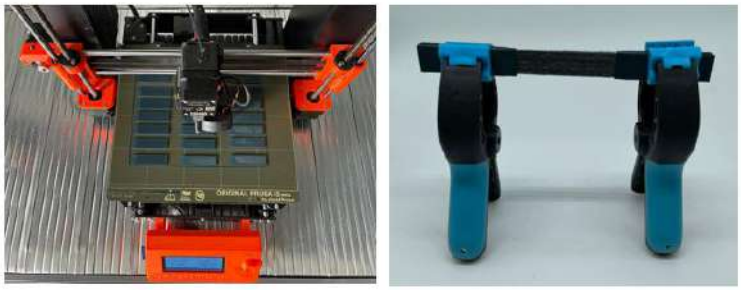


Fig. 6.11: Printed tabs made of PLA and the clamping process.

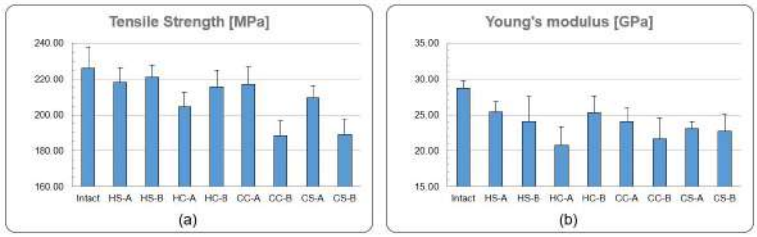


Fig. 6.12: The bar graphs of mechanical properties of additive manufactured CFRP composites. (a) mechanical strength; and (b) Young's modulus. For interpretation of the samples' notation, the reader is suggested to see Table 6.2.

The tensile strength of the 3D printed specimens of each group is shown in Fig. 6.12. From the bar graph plot, it can be seen that the intact samples attained the highest average tensile strength value of 226.14 MPa. The strength of CFRP was then followed by sample group CC-A, CS-A, HC-A, and HS-A with value of 216.8 MPa, 204.41 MPa, 207.19 MPa, and 217.99 MPa respectively. The reduced strength values were assumed due to poor reproducibility of 3D printing technique which caused the deviation in specimen features such as sample dimension.

For both hot continuous and cyclic thermal treatment, the group of HS-B and HC-B with higher temperature have lower ultimate strength of 221.21 MPa and 215.49 MPa respectively compared to specimen group HS-A and HC-A. The opposite trend occurred to cold continuous and cyclic thermal cases where the specimen groups with lower temperature namely CS-B and CC-B have lower ultimate strength of 188.93 MPa and 188.32 MPa respectively compared to CS-A and CC-A group. Furthermore, the similar trend occurred to the elastic modulus (Young's modulus) of each specimen group as in tensile strength. For both hot continuous and cyclic thermal treatment, the specimen group of HS-B and HC-B with higher temperature have lower elastic modulus compared to specimen group HS-A and HC-A. While for cold continuous and cyclic thermal cases, the specimen groups with lower temperature namely CS-B and CC-B have lower elastic modulus compared to CS-A and CC-A group.

The reason for this strength and elastic modulus deterioration could be the poor reproducibility of 3D printing technique which lead to the deviation in specimen features like material dimension and density. Additionally, some porosity amount contained in each specimen group could also lead to decrease values of the strength and modulus of thermally treated samples. The higher porosity which might be existed in high temperature group (HS-B and HC-B) formed lower inter-facial bonding between matrix and fiber component. High temperature has lower moisture which make the polymer matrix dried up more such that its weight fraction reduced since the water content or wettability is decreased. In the case of samples subjected to negative temperature, it is assumed that lower temperat-

ure group (CC-B and CS-B) obtained reduced strength and elastic modulus due to smaller degree of crystallinity of polymer matrix PLA. When the cross-linking of polymer is not strongly occurred, the polymer chain could make a better bonding so it will lead to the decrease in mechanical properties of CFRP composite.

6.4.2 Numerical modeling of tensile and flexural loading

Numerical modelling of tensile and flexural (three-point bending) testing of 3D-printed CFRP specimens as virtual tool using ABAQUS have been performed by Muna et al [48]. A 3D unidirectional composite model at the macro-scale has been established to assess the numerical parameters of samples such as its mechanical strength and strain at failure. The investigated structure is comprised of a combination of matrix and fibres. In this work, the material properties of composite were calculated using rule of mixture (ROM), theoretical model by Voight. The individual volume fractions for matrix PLA, continuous carbon fiber (CCF), and short carbon fiber (SCF) were obtained from literature by [49].

6.4.3 Finite element modeling

As a deformable shell, the composite model geometry was built using the computer-aided design (CAD) software ABAQUS. The generated model is conditioned in accordance with Maqsood's experiment [49]. Calculating the material properties measured on printed specimens from experimental tensile and flexural tests that will be utilized as input data is necessary. The force-displacement curve can be derived using material data and finite element analysis. The geometry of the tensile test specimen is a unidirectional carbon fiber laminate with a rectangular cross-section, 150 mm in length, 13mm in width, and 3 mm in thickness, and 45mm on either end for grasping. The dimensions of the flexural test specimen are 127 x 12.7 x 3.2 mm, and two tabs are positioned 37 mm from each end. The geometry of the test specimen is intended to resemble the actual geometry.

Fig. 6.13: Model specimen of tensile test with boundary and loading conditions. Taken with permission from [48].

A finite element model of a composite specimen subjected to static load was created, and a stress analysis was performed. To reproduce the experimental results, the boundary conditions and loads are applied exactly as they are in the real tensile test. In the lower grip, the boundary conditions for tensile test simulation are fixed or clamped (encastred) in all directions, whereas the upper grip is also fixed in all directions but left free in the direction of the applied force (unconstrained in longitudinal direction). These parameters ensure that the tensile test simulation is as accurate as possible, with no rotations or bending. The top-side grasp was loaded with 50 MPa of uniformly distributed general traction surface traction. Fig. 6.13 depicts a model of a tensile test specimen with boundary and loading conditions. When a simple structure, such as a composite plate, breaks due to applied forces, it fails fast because the load increases as the structure's load carrying capacity falls. Displacement controlled loading minimizes the structure's weight as it

fails, slowing failure. The model was discretized into linear quadrilateral elements of a four-node general-purpose shell, decreased integration with hourglass control, and finite membrane strains. The FEM model comprised of 1950 finite elements and 2114 nodes. Reduced integration denotes a lower order of integration than complete integration.

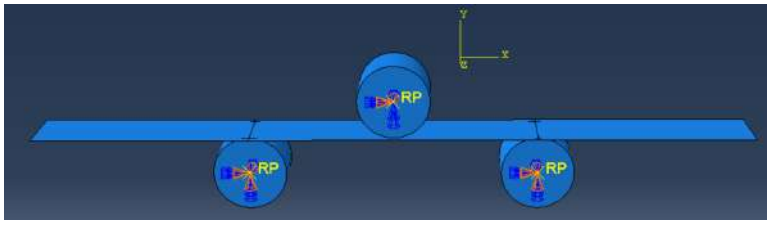


Fig. 6.14: Model specimen of flexural test with boundary and loading conditions. Taken with permission from [48].

The model of three-point bending set-up was created for the flexural test: two supports and one midway load nose with 10 mm diameter of each. In order to replicate the real flexural test, three cylinders were created with diameter of 10 mm and reference point was assigned to each cylinder at its center point since the vertex in model shape is not available. These cylinders then positioned onto sample model with the reference point (one on top and two below) at approximately same coordinates as in experiment. General contact was applied to these shapes in order to introduce the interaction between its surfaces with very few restrictions on the types of surfaces involved. Boundary conditions for two cylinders at bottom were set to be encastred (fixed) while the top cylinder was kept free in pushing direction where load is also applied. Model specimen of tensile test with boundary and loading conditions is presented in Fig. 6.14. Similarly with specimen model in tensile test, the model in flexural test was also discretized into linear quadrilateral elements of four-node general-purpose shell, reduced integration with hourglass control, finite membrane strains within stress/displacement S4R type. The FEM model consisted of 1736 finite elements, and 1599 nodes.

6.4.4 Tensile values

The maximum displacement occurred at the free end and the reaction force was found to be maximum at the clamped end for both models. The results of Young's Modulus for each set of groups is shown in Table 6.3. It was found that the numerical result is in pretty good agreement with the obtained result from the experiment by [49] for the specimens of PLA-CCF and PLA-SCF with an average Young's Modulus value of 27.93 GPa, 2.88 GPa and 2.571 GPa, respectively. The simulation and experimental results for both mentioned constituents are in close relationship less than 5% deviation ratio. The results indicated that by increasing the content of fibers in the matrix will increase the Young's Modulus value. However, the PLA-SCF-CCF numerical simulation found a mean Young's Modulus of 29.07 GPa, which is slightly higher than the experimental result. This value disparity is likely caused by printed samples' larger voids or pores. Since PLA-SCF-CCF has more carbon fiber content than that of other sample types (PLA-SCF and PLA-CCF), printing voids reduce material strength compared to numerical results where voids are considered to be absent.

Tab. 6.3: Tensile strength comparison

Constituent	Young's Modulus [GPa]	
	Experimental	Numerical
PLA-SCF-CCF	27.93	29.07
PLA-CCF	25.94	27.93
PLA-SCF	4.79	2.88

6.4.5 Flexural values

Flexural modulus performance recorded via three-point bend testing and computationally-obtained using ROM technique. A comparison of flexural strength determined experimentally and numerically is presented in Table 6.4. From three-point bend tests, the stress-strain curves were generated and flexural modulus of the

CFRP sample was obtained. Because of the higher modulus of CCF and SCF, CFRP composites constructed of PLA polymer coupled with CCF and SCF have the highest flexural modulus. While the ROM model provides a relatively precise flexural modulus, it overestimates the same when compared to experimental flexural modulus. This finding is consistent with the widely held belief that ROM is an upper-bound estimate of the composite modulus because it does not account for composite defects such as bonding or pores.

Tab. 6.4: Flexural strength comparison

Constituent	Flexural modulus [GPa]	
	Experimental	Numerical
PLA-SCF-CCF	10.85	9.98
PLA-CCF	10.63	9.8
PLA-SCF	4.52	4.03

These simulations can produce results that are comparable to those obtained from actual tensile and flexural testing, which is how specimen mechanical properties are characterized. The projected results from numerical models at the macro-scale level were compared to the available experimental data, and a fairly good match was found between the two sets of data.

6.5 Conclusion

In this chapter, a variety of temperature loadings subjected to additive manufactured carbon fiber reinforced polymer (CFRP) composites were presented and discussed. Next, various thermal treatments at positive and negative temperature on printed composites were studied experimentally to understand the mechanical properties and internal structures of composites. We investigated the effect of stable continuous and cyclic temperature on the mechanical properties of unidirectional 3D CFRP composites. Static tensile testing was performed to obtain the mechanical strength, Young's

modulus, and failure damage of these composites. The results indicated that unidirectional CFRP composites undergo thermal exposures, revealing a material degradation both in mechanical strength and Young's modulus under stable and cyclic temperature. Untreated (intact) samples attained the highest average tensile strength value of 226.14 MPa and Young's modulus of 28.65 GPa.

Previously, a global heat exposure at elevated temperature on a unidirectional (UD) continuous CFRP composite sample has been experimentally investigated using non destructive testing by means of fiber Bragg grating (FBG) sensor and numerically simulated using finite element modeling (FEM) analysis software, ABAQUS. The numerical model's accuracy has been validated using convergence analysis and mesh resolution (grid independent). The convergence was achieved by placing seeds at a distance of 0.25 on each edge. The numerical data result was compared to the experimental result, and a good correlation was found. The mean difference between numerical and experimental values is about 18%, which is most likely due to sample structure heterogeneity. The increasing strain differences at temperatures above 38°C can also be attributed to matrix material processes and the weakness of bondings between fibres and matrix in the composite structure.

Numerical modelling of tensile and flexural (three-point bending) testing of 3D-printed CFRP specimens as virtual tool using ABAQUS have been. The tensile testing simulation results were compared to experimental values obtained and a good agreement with less than 5% deviation ratio was achieved for the PLA-SCF and PLA-CCF samples. The results indicated that by increasing the content of fibers in the matrix will increase the Young's Modulus value. However, the PLA-SCF-CCF numerical simulation found a mean Young's Modulus is far higher than that of the experimental result. This value disparity is likely caused by printed samples' larger voids or pores. Since PLA-SCF-CCF has more carbon fiber content than that of other sample types (PLA-SCF and PLA-CCF), printing voids reduce material strength compared to numerical results where voids are considered to be absent.

The findings presented in this paper suggest future research into the

degradation of 3D composites under various temperature conditions in order to study the structure's integrity and the damage mechanism. Investigating the effect of different thermal cycling parameters such as heating and cooling rates and dwell time on the properties of materials in future work would be beneficial for industrial applications. Thermal degradation of the polymer material used in the additive manufacturing of a composite should be investigated using rheological and thermogravimetric measurements to understand the material's thermal behavior during heating conditions. Finally, comprehending the thermal influence at various conditions on the morphological structure of FDM printed materials and their mechanical properties can demonstrate the first step toward a solution.

Acknowledgement

The research was financed by National Science Centre (NCN), Poland, within the project entitled Thermal degradation processes of additive manufactured structures (2019/35/O/ST8/00757), and the Polish National Agency for Academic Exchange (NAWA) within the Foreign doctoral fellowship program (PPN/STA/2021/1/00006).

Bibliography

[1] Wadea Ameen, Abdulrahman Al-Ahmari, and Osama Abdulhamid. Design for metal additive manufacturing: An investigation of key design application on electron beam melting. *Intl Journal of Mechanical, Aerospace, Industrial, Mechatronic and Manufacturing Engineering*, 13:264–269, 2019.

[2] Xin Wang, Man Jiang, Zuowan Zhou, Jihua Gou, and David Hui. 3d printing of polymer matrix composites: A review and prospective. *Composites Part B: Engineering*, 110:442–458, 2017.

[3] Charles Dapogny, Rafael Estevez, Alexis Faure, and Georgios Michailidis. Shape and topology optimization considering anisotropic features induced by additive manufacturing processes. *Computer Methods in Applied Mechanics and Engineering*, 344:626–665, 2019.

[4] Hao Peng, Morteza Ghasri-Khouzani, Shan Gong, Ross Attardo, Pierre Ostiguy, Ronald B Rogge, Bernice Aboud Gatrell, Joseph Budzinski, Charles Tomonto, Joel Neidig, et al. Fast prediction of thermal distortion in metal powder bed fusion additive manufacturing: Part 2, a quasi-static thermomechanical model. *Additive Manufacturing*, 22:869–882, 2018.

[5] Luke N Carter, Christopher Martin, Philip J Withers, and Moataz M Attallah. The influence of the laser scan strategy on grain structure and cracking behaviour in slm powder-bed fabricated nickel superalloy. *Journal of Alloys and Compounds*, 615:338–347, 2014.

[6] Shukantu Dev Nath and Sabrina Nilufar. An overview of additive manufacturing of polymers and associated composites. *Polymers*, 12(11):2719, 2020.

[7] Ahmed El Moumen, Mostapha Tarfaoui, and Khalid Lafdi. Additive manufacturing of polymer composites: Processing and modeling approaches. *Composites Part B: Engineering*, 171:166–182, 2019.

[8] Fuda Ning, Weilong Cong, Jingjing Qiu, Junhua Wei, and Shiren Wang. Additive manufacturing of carbon fiber reinforced thermoplastic composites using fused deposition modeling (fdm). *Composites Part B: Engineering*, 80:369–378, 2015.

[9] Halil L Tekinalp, Vlastimil Kunc, Gregorio M Velez-Garcia, Chad E Duty, Lonnie J Love, Amit K Naskar, Craig A Blue, and Soydan Ozcan. Highly oriented carbon fiber–polymer composites via additive manufacturing. *Composites Science and Technology*, 105:144–150, 2014.

[10] Rafael Thiago Luiz Ferreira, Igor Cardoso Amatte, Thiago Assis Dutra, and Daniel Burger. Experimental characterization and micrography of 3d printed pla & pla reinforced with short carbon fibers. *Composites Part B: Engineering*, 124:88–100, 2017.

[11] Mohammad Heidari-Rarani, AM Rafiee, and AM Zahedi. Mechanical characterization of fdm 3d printing of continuous carbon fiber reinforced pla composites. *Composites Part B: Engineering*, 175:107147, 2019.

[12] Behnam Akhoundi, Amir Hossein Behravesh, and Arvin Bagheri Saed. Improving mechanical properties of continuous fiber reinforced thermoplastic composites produced by fdm 3d printer. *Journal of Reinforced Plastics and Composites*, 38(3):99–116, 2019.

[13] Mohammad Reza Khosravani, Ali Zolfagharian, Matt Jennings, and Tamara Reinicke. Structural performance of 3d-

printed composites under various loads and environmental conditions. *Polymer testing*, 91:106770, 2020.

[14] Pavan Kumar Penumakala, Jose Santo, and Alen Thomas. A critical review on the fused deposition modeling of thermoplastic polymer composites. *Composites Part B: Engineering*, 201:108336, 2020.

[15] Sri Addepalli, Yifan Zhao, Rajkumar Roy, Wathsala Galhenege, Marine Colle, Jingjing Yu, and Aziz Ucur. Non-destructive evaluation of localised heat damage occurring in carbon composites using thermography and thermal diffusivity measurement. *Measurement*, 131:706–713, 2019.

[16] A Haridas, C Song, K Chan, and VM Murukeshan. Nondestructive characterization of thermal damages and its interactions in carbon fibre composite panels. *Fatigue & Fracture of Engineering Materials & Structures*, 40(10):1562–1580, 2017.

[17] Victor Munoz, Benjamin Valès, Marianne Perrin, Marie-Laetitia Pastor, Hélène Welemane, Arthur Cantarel, and Moussa Karama. Damage detection in cfrp by coupling acoustic emission and infrared thermography. *Composites Part B: Engineering*, 85:68–75, 2016.

[18] C Pascual-González, P San Martín, I Lizarralde, A Fernández, A León, CS Lopes, and JP Fernández-Blázquez. Post-processing effects on microstructure, interlaminar and thermal properties of 3d printed continuous carbon fibre composites. *Composites Part B: Engineering*, 210:108652, 2021.

[19] Yahya Abderrafai, Mohammad Hadi Mahdavi, Facundo Sosa-Rey, Chloé Hérard, Ivonne Otero Navas, Nicola Piccirelli, Martin Lévesque, and Daniel Therriault. Additive manufacturing of short carbon fiber-reinforced polyamide composites by fused filament fabrication: Formulation, manufacturing and characterization. *Materials & Design*, 214:110358, 2022.

[20] Marcus Ivey, Garrett W Melenka, Jason P Carey, and Cagri Ayranci. Characterizing short-fiber-reinforced composites produced using additive manufacturing. *Advanced Manufacturing: Polymer & Composites Science*, 3(3):81–91, 2017.

[21] Amal Nassar, Mona Younis, Mohamed Elzareef, and Eman Nassar. Effects of heat-treatment on tensile behavior and dimension stability of 3d printed carbon fiber reinforced composites. *Polymers*, 13(24):4305, 2021.

[22] P Arjun, VK Bidhun, UK Lenin, VP Amritha, Ribin Varghese Pazhamannil, and P Govindan. Effects of process parameters and annealing on the tensile strength of 3d printed carbon fiber reinforced polylactic acid. *Materials Today: Proceedings*, 2022.

[23] Anouar El Magri, Khalil El Mabrouk, Sébastien Vaudreuil, and Mohamed Ebn Touhami. Mechanical properties of cf-reinforced pla parts manufactured by fused deposition modeling. *Journal of Thermoplastic Composite Materials*, 34(5):581–595, 2021.

[24] Claire Benwood, Andrew Anstey, Jacek Andrzejewski, Manjusri Misra, and Amar K Mohanty. Improving the impact strength and heat resistance of 3d printed models: structure, property, and processing correlationships during fused deposition modeling (fdm) of poly (lactic acid). *Acs Omega*, 3(4):4400–4411, 2018.

[25] Isyna Izzal Muna. Introduction to thermal degradation on additive manufactured polymeric composites. *Selected problems in mechanical engineering 2021*, In Magdalena Mieloszyk and Tomasz Ochrymiuk, editors.

[26] Kevin R Hart, Ryan M Dunn, Jennifer M Sietins, Clara M Hofmeister Mock, Michael E Mackay, and Eric D Wetzel. Increased fracture toughness of additively manufactured amorphous thermoplastics via thermal annealing. *Polymer*, 144:192–204, 2018.

[27] Radoslaw A Wach, Piotr Wolszczak, and Agnieszka Adamus-Wlodarczyk. Enhancement of mechanical properties of fdm-pla parts via thermal annealing. *Macromolecular Materials and Engineering*, 303(9):1800169, 2018.

[28] Tejas Pramod Naik, Inderdeep Singh, and Apurbba Kumar Sharma. Processing of polymer matrix composites using microwave energy: A review. *Composites Part A: Applied Science and Manufacturing*, page 106870, 2022.

[29] Caroline Lüders and Michael Sinapius. Fatigue of fibre-reinforced plastics due to cryogenic thermal cycling. *Journal of Composite Materials*, 53(20):2849–2861, 2019.

[30] Yottha Srithep, Paul Nealey, and Lih-Sheng Turng. Effects of annealing time and temperature on the crystallinity and heat resistance behavior of injection-molded poly (lactic acid). *Polymer Engineering & Science*, 53(3):580–588, 2013.

[31] Jing Zhou, Yingguang Li, Mengchen Zhang, Eyan Xu, and Tao Yang. Effect of lay-up configuration on the microwave absorption properties of carbon fiber reinforced polymer composite materials. *Materials Today Communications*, 26:101960, 2021.

[32] Jiayi Liu, Linling Xiang, and Tao Kan. The effect of temperature on the bending properties and failure mechanism of composite truss core sandwich structures. *Composites Part A: Applied Science and Manufacturing*, 79:146–154, 2015.

[33] Yao Ma, Qiang Huang, Yixin Yu, Yuan Dong, Hui Li, Guangyong Jin, and Chao Xin. Theoretical and experimental investigation on temperature rise of cfrp with long pulse laser irradiation. *Applied Sciences*, 11(22):10555, 2021.

[34] MC Lafarie-Frenot, S Rouquie, NQ Ho, and V Bellenger. Comparison of damage development in c/epoxy laminates during isothermal ageing or thermal cycling. *Composites Part A: Applied Science and Manufacturing*, 37(4):662–671, 2006.

[35] AR Ghasemi, A Tabatabaeian, and M Moradi. Residual stress and failure analyses of polymer matrix composites considering thermal cycling and temperature effects based on classical laminate plate theory. *Journal of Composite Materials*, 53(21):3021–3032, 2019.

[36] Frank KA Nyarko, Gabriel Takyi, Emeka H Amalu, and Muyiwa S Adaramola. Generating temperature cycle profile from in-situ climatic condition for accurate prediction of thermo-mechanical degradation of c-si photovoltaic module. *Engineering Science and Technology, an International Journal*, 22(2):502–514, 2019.

[37] Piyush K Dutta and David Hui. Low-temperature and freeze-thaw durability of thick composites. *Composites Part B: Engineering*, 27(3-4):371–379, 1996.

[38] David M Grogan, Sean B Leen, COA Semprimoschnig, and CM Ó Brádaigh. Damage characterisation of cryogenically cycled carbon fibre/peek laminates. *Composites Part A: Applied Science and Manufacturing*, 66:237–250, 2014.

[39] Ran Y Kim, Allan S Crasto, and Gregory A Schoeppner. Dimensional stability of composite in a space thermal environment. *Composites science and technology*, 60(12-13):2601–2608, 2000.

[40] Shambhu K Gupta and Mehdi Hojjati. Thermal cycle effects on laminated composite plates containing voids. *Journal of Composite Materials*, 53(4):489–501, 2019.

[41] Daniel S Adams, David E Bowles, and Carl T Herakovich. Thermally induced transverse cracking in graphite-epoxy cross-ply laminates. *Journal of reinforced Plastics and composites*, 5(3):152–169, 1986.

[42] MC Lafarie-Frenot and NQ Ho. Influence of free edge intralaminar stresses on damage process in cfrp laminates under thermal cycling conditions. *Composites Science and Technology*, 66(10):1354–1365, 2006.

[43] C Guigon, MC Lafarie-Frenot, Y Pannier, L Olivier, and C Rakotoarisoa. Impact of temperature and thermal cycling ageing on performance of 3d woven composites whit polymer matrix manufactured by rtm. In *European Conference on Composite Materials (ECCM16), Seville, Spain*, volume 22, page 26, 2014.

[44] Chao Zhang, Wieslaw K Binienda, Gregory N Morscher, Richard E Martin, and Lee W Kohlman. Experimental and fem study of thermal cycling induced microcracking in carbon/epoxy triaxial braided composites. *Composites Part A: Applied Science and Manufacturing*, 46:34–44, 2013.

[45] Isyna Izzal Muna and Magdalena Mieloszyk. Temperature influence on additive manufactured carbon fiber reinforced polymer composites. *Materials*, 14(21):6413, 2021.

[46] Marius Rimašauskas, Tomas Kuncius, and Rūta Rimašauskienė. Processing of carbon fiber for 3d printed continuous composite structures. *Materials and Manufacturing Processes*, 34(13):1528–1536, 2019.

[47] Peng Wang, Bin Zou, Hongchuan Xiao, Shouling Ding, and Chuanzhen Huang. Effects of printing parameters of fused deposition modeling on mechanical properties, surface quality, and microstructure of peek. *Journal of Materials Processing Technology*, 271:62–74, 2019.

[48] Isyna Izzal Muna and Magdalena Mieloszyk. The numerical model of additive manufactured carbon fiber reinforced polymer under mechanical loading. In *Nondestructive Characterization and Monitoring of Advanced Materials, Aerospace, Civil Infrastructure, and Transportation XVI*, volume 12047, pages 303–308. SPIE, 2022.

[49] Nabeel Maqsood and Marius Rimašauskas. Characterization of carbon fiber reinforced pla composites manufactured by fused deposition modeling. *Composites Part C: Open Access*, 4:100112, 2021.

Chapter 7

Selected issues of hydrogen energy in the maritime industry

Marek Pluciński

Institute of Fluid Flow Machinery, Polish Academy of Sciences,
Fiszera 14, 80-231 Gdansk, Poland

7.1 Introduction

The aim of the chapter is to check the possibility of using hydrogen as a ship fuel.

Renewable energy means energy derived from natural repetitive natural processes, obtained from renewable non-fossil energy sources. In EU-28, the share of energy from renewable sources in total final energy consumption increased in 2016-2019 from 9.2% to 10.2% [1].

Renewable energy sources (RES) are an alternative to traditional primary non-renewable energy carriers (fossil fuels). Their resources replenish in natural processes which practically allows to treat them as inexhaustible [1].

The International Convention for Prevention of Pollution from Ships (MARPOL) is the most important convention concerning counteraction of contamination of the marine ecosystem. Pollution are caused by ships operations or their random action.

The MARPOL Convention was signed on 02.11.1973 at International Maritime Organization. The 1978 MARPOL Protocol absorbed the Convention from 1973. Consolidated document took effect on 02.09.1983. In 1997, a new protocol was passed to change the Convention and Annex VI was pronounced which become operative on 19.05.2005. MARPOL had new versions over the years [2].

Annex VI was first document which delimit the sulphur oxide and nitrogen oxide emissions from the exhaust and forbid deliberate emissions of substances that reduce the ozone layer. Determined emission control zones set restricted limits for SOx, NOx and particulate matter. That document contains obligatory technical and operational energy efficiency measures [2].

In July 2019 the European Commission announced project called – The European Green Deal. The European Green Deal contains to reduce emissions by at least 55% by 2030, comparing to 1990 values [3].

This also applies to ships flying the flag of countries belonging to the European Union [3].

Intensive research and development work is needed to develop tech-

nologies that significantly reduce emissions substances harmful to the atmosphere.

7.2 Increasing the share of hydrogen in the energy sector

The Swiss alchemist Paracelsus was the first person to make the first observations of hydrogen. During his research, he threw metals into acids and collected their products into gas tanks reaction. In his notes, he wrote that air, formed as a result of a chemical reaction of hydrochloric acid with iron is explosive. Irish chemist Robert Boyle conducted research on hydrogen in 1671. It was the same as in the case of Paracelsus studies through the reaction of metals and acids. One hundred years later, Henry Cavendish isolated hydrogen as a separate substance and described the combustion reaction.

Initially, no practical use of this gas was found. Hydrogen was used in the 19^{th} and 20^{th} centuries as a gas used to fill airships and balloons in air transport.

Increasing energy demand and environmental degradation determined the need for new, safe, highly efficient and environmentally friendly energy technologies. Replacing fossil fuels with hydrogen allows for complete elimination of the emission of the substance harmful to the atmosphere.

On an industrial scale, hydrogen is currently produced mainly from natural gas in steam methane reforming (SMR) [4].

Hydrogen can be also selected by using partial oxidation (POX); catalitical partial oxidation (CPOX), combining both of the above autothermal reforming (ATR), cathalytic dehydrogenation, pyrolysis and electrolysis [4].

At a different point of the development there are other technologies such as photocatalytic processes, plasma reforming, membrane reactors, biological processes [4].

Tab. 7.1: Efficiency of hydrogen production technology [4].

Process	Fuel	Total efficienty,%
Steam methan reforming	natural gas or crude oil	$65 - 75^{1)}$
Partial oxidation	natural gas or crude oil	$50^{1)}$
Gasification	coal or biomass	$42.5 - 46.5^{2)}$
Pyrolysis	coal or biomass	$9 - 13^{2)}$

[1)] calorific value, [2)] heat of combustion

Further development of hydrogen energy is closely related to the use of renewable sources for energy production. Most dominated technology today is electrolysis. In Tab. 7.2 presents the characteristics of three types of electrolysers: alkaline, polymeric (PEM) and ceramic (SOEC). In operation, the alkaline and PEM electrolysers are currently the most popular [4].

Tab. 7.2: The state of development of electrolysers [4].

Indicator	Alkaline electrolysers	PEM	SOEC
Electrical efficienty [%, Wd]	63-70	56-60	74-81
Working pressure [bar]	1-30	30-80	1
Working temperature [K]	333-353	323-353	923-1273
Working time [h]	$60^4 - 90^4$	$30^4 - 90^4$	$10^4 - 30^4$

The economic development forecast assumes an increase in the production of hydrogen from electrolysis. Currently, the cost of electricity from renewable sources is quite high, which requires the improvement of production processes and consistency in climate policy. Well-understood support policy may be of key importance in the perspective of replacing fossil energy sources [4].

The development of fuel cells is of great importance for the implementation of the technology of energetic use of hydrogen [4].

Hydrogen and ammonia can be a flexible option for ship fuel. They can determine the decarbonisation of maritime transport.The cessation of hydrogen production from fossil fuels in order to replace it with production from renewable energy sources should be considered in the long term [4].

7.3 Isotopes and molecules of hydrogen

Hydrogen is the chemical element with the symbol H and atomic number 1. Hydrogen is the lightest element. At standard conditions hydrogen is a gas of diatomic molecules having the formula H_2. It is colorless, odorless, tasteless, non-toxic, and highly combustible. Hydrogen is the most abundant chemical substance in the universe, constituting roughly 75% of all normal matter [5]. Most of the hydrogen on Earth exists in molecular forms such as water and organic compounds. Isotopes protium ${}^1_1\mathrm{H}$ (one proton and one electron), deutrium ${}^2_1\mathrm{H}$ (one proton, one neutron and one electron) are stable. Tritium ${}^3_1\mathrm{H}$ is not stable and decays according to the reaction (e^--electron, v- electric neutrinos) [4].

$$ {}^3_1H \rightarrow {}^3_2 He + e^- + v $$

Protium makes up 99.985% of the hydrogen present on earth. Deutrium is about 0.015% and tritium is present in trace amounts. Hydrogen isotopes form diatomic molecules (H_2, D_2, T_2), differing in the orientation of proton spins. There are ortohydrogen (o-H_2) and parahydrogen (p-H_2). In the orthohydrogen form, the spins of the two nuclei are parallel. In the parahydrogen form the spins are antiparallel. Hydrogen is a mixture of both types in a molar ratio: 75% ortohydrogen and 25% parahydrogen in temperature 293.15K and higher. Selected physical properties of parahydrogen and natural hydrogen in natural conditions are presented in Tab. 7.3 [4].

Tab. 7.3: Selected physical properties of hydrogen (273.15K, 1.013bar) [4].

Value	p-H_2	n-H_2
Molar mass	2.016	2.016
Density, kg/m^3	0.0899	0.0899
Specific heat capacity c_p,J,(mol K)	30.35	28.59
Specific heat capacity c_v,J,(mol K)	21.87	20.3
Dynamic viscosity, μPa s	8.34	8.34
Speed of sound,m/s	1246	1246
Conductivity coefficient, W/(m K)	0.1826	0.1739
Dielectric constant	1.0005	1.00042
Compressibility factor	1.0005	1.00042
Prandtl number	0.6873	0.680

7.4 Transport and storage of hydrogen

Hydrogen pipelines are operated in the USA, Canada and Europe. The first European pipeline has been operating in Germany since 1939 (Essen). The characteristics of the pipeline are as follows: flow 2.52 kg/s, pressure 20 bar, temperature 300K, diameter of the pipeline 0.254 m, flow velocity 31 m/s and length 220 km. The longest European pipeline between France and Belgium is 400 km long, the pressure is 100 bar, the temperature is 300K and the diameter of the pipelines is 0.1 m. In the United States, there are currently more than 2 000 km of pipelines transporting hydrogen at a pressure of 30-80 bar. A new line of hydrogen transport pipelines is located in Belgium between at Antwerp and Feluy. Modern technologies are fully taken into account in the construction of these pipelines. The quality of the material complies with the standard API SPEC5L Grade X42. On the outer surface there is a polyethylene lining. The pipeline connections are welded [4].

There are basic types of potential geological stores: salt caverns, deep aquifers, empty oil fields and empty gas fields [4].

Tab. 7.4: Salt quarters for hydrogen storage – technical parameters [4, 6, 7].

Location/ parameters	**Teesside, UK**	**Chevron Phillips, Texas**	**AirLiquide, Texas**	**Moss Bluff Praxair,Texas**
Total volume, m^2	3*150000 [8] 3*70000 [7]	580000	906000	566000
Location depth, m	350-380	850-1150	1158-1524	1400
Pressure levels, MPa	4.5 - const	7-13.5	6.8-20	no data
Accumulated energy, GWh	25	83.3 ca. 2500 t H_2	no data	80

Based on the analyzed data, it can be indicated that the cheapest technologies for underground hydrogen storage are technologies using depleted oil and gas deposits and aquifers [4, 7].

Tab. 7.5: Total cost of different types of underground hydrogen storage facilities [9, 10].

Types of storage	**Salt caverns**	**Oil and gas fields**	**Aquifers**	**Rock caves**
Cost USD/kg H_2	1.61	1.23	1.29	2.27

From the point of view of accumulated power, the installations of pumped storage plants, compressed air energy storage (CAES), adiabatic compressed air energy storage (ACAES) and storage compressed hydrogen have the greatest potential [4].

Compression for small volume tanks is characterized by an efficiency of 5-30%, a product life cycle of about 20-50 years and an investment cost of about EUR 1050-3000/kW [4].

Tab. 7.6: Technical data of individual energy storage technologies [4].

Technology	Power, MW	Total efficiency, %	Investment cost, EUR/kW
Pumped storage power plant	100-5000	75-87	470-2700 (3300)
CAES	15-400	42-54	450-1150
ACAES	100-500	70	600-1200

7.5 Threats and safety protection

Physical hazards are related to the properties of metals and relate to the phenomenon of hydrogen embrittlement. It is a phenomenon dependent on temperature and pressure, hydrogen interaction time and the condition of the metal surface. The phenomenon of hydrogen embrittlement increases at a temperature of 200-300K. This phenomenon can be prevented by using oxygen coatings, appropriate metal alloys, eliminating the concentration of hydrogen and introducing special additives to the hydrogen [4].

Chemical hazards are related to the properties of hydrogen. This gas has a very low ignition energy of 0.02 mJ. Electrical and thermal equipment and open flame sources should be adequately isolated from places where a hydrogen system is used. Hydrogen also has a wide flammability range, which makes it quite dangerous [4].

The main risks of failure of hydrogen transport and storage modules are:

- hydrogen evolution at low temperatures, which may pose a threat to humans in the form of cryogenic burns,
- fire related to the uncontrolled release and ignition of hydrogen; this phenomenon causes a risk to humans in the form of direct exposure to flame or heat radiation; it can take the form of a jet fire, characterized by a long, steady flame. It can arise when hydrogen is released under high pressure, for example through an opening in a pipe wall,

- BLEVE phenomenon (boiling liquid expanding vapour explosion). This phenomenon may occur when the hydrogen containing tank is damaged. The danger to people is related to the pressure wave and metal parts detached from the tank. If the phenomenon is accompanied by a fire in the shape of a ball, the risk is also the direct effect of flame and heat radiation,
- an explosion occurs in the event of an uncontrolled release of hydrogen from a tank or pipeline; the threat to people and the environment is related to the impact of the pressure wave and elements of the damaged infrastructure [4].

7.6 Combustion reaction

Combustion is a chemical reaction in which combustible components combine with oxygen, accompanied by intense heat generation, causing the temperature to rise rapidly. In contrast to the usual reactions with large exothermic effects, there is a visible flame during combustion, which is a reaction zone in the gas phase. The theoretical result of the combustion reaction can be described by stoichiometric equations. These equations characterize the reactants and the resulting substances. In the case of exothermic reactions, we supplement their material result with the thermal effect, i.e. the energy generated during the reaction. The stoichiometric equation for the hydrogen combustion reaction is presented below [11].

$$2H_2 + O_2 \rightarrow 2H_2O + 483.2kJ$$

The rate of a chemical reaction can be defined as the change in the concentration of the reaction components per unit of time and per unit of volume. In the case of organized combustion of fuel it can be conditionally regarded as constant, because new streams of fuel and oxidant continuously flow into the combustion zone. The rapid course of the reaction is caused, among others, by chain mechanism of these reactions. Any chemical reaction involves the rearrangement of the chemical bonds between atoms during the transition of the reactants from the initial to the final state. However, the transformation of the input reactants from the initial state to the end

products of the reaction takes place not in one stage, but through a series of stages with the formation and decomposition of intermediate chemically active particles - atoms and radicals generated by the reaction itself [11].

These particles easily bind to the starting reagents and to each other, contributing to the formation of both end products and new active particles, capable of initiating a similar chain of reactions. The increasing self-generation of active particles causes a rapid increase in the rate of reaction in the form of an explosion [12].

The chain mechanism of the hydrogen combustion reaction with oxygen is one of the simplest and best known mechanisms. The initiation of this reaction is associated with the formation of atamic hydrogen as a result of the dissociation of molecular hydrogen [11]:

$$H_2 + M = H^{\cdot} + H^{\cdot} + M$$

Element M should be understood as the source of activation energy. An analogous initiation of the reaction by atomic oxygen is unlikely due to the very high dissociation energy of molecular oxygen. The diagram of the further course of the chain reaction is shown below [11].

$$\begin{array}{lllllll} & & & & & \rightarrow & H_2O \\ & & & \rightarrow & OH^{\cdot} + H_2 & \rightarrow & H^{\cdot} \\ & \rightarrow & O^{\cdot} + H_2 & \rightarrow & H^{\cdot} & & \\ H^{\cdot} + O_2 & & & & & & \\ & \rightarrow & OH^{\cdot} + H_2 & \rightarrow & H_2O & & \\ & & & \rightarrow & H^{\cdot} & & \end{array}$$

Fig. 7.1: Unit cycle of the hydrogen-oxygen chain reaction mechanism [11].

The result of such a single cycle is a reaction [11]:

$$H^{\cdot} + 3H_2 + O_2 \rightarrow 2H_2O + 3H^{\cdot}$$

Thus it can be concluded that each hydrogen radical gives rise to two water molecules (end product) and three new radicals. In fact, any hydrogen radical can now contribute to the chain development through a successful collision with an oxygen molecule. However,

a collision of two radicals is also likely, resulting in a stable hydrogen molecule [11]:

$$H^{\cdot} + H^{\cdot} \rightarrow H_2$$

A feature of gaseous fuel is that in a properly prepared mixture with air, the number of successful events favoring the development of the reaction chain exceeds the number of events closing the chain. Ignition of the mixture in one place causes the flame to spread very quickly over the entire volume of the mixture, which is called an explosion [11].

For the stoichiometric equation of the combustion reaction:

$$H_2 + 0.5O_2 \rightarrow H_2O$$

The heat of combustion is 10.785 [MJ/m^3] [11].

The stoichiometric formulas predict the theoretical binding of the fuel directly to oxygen. In fact, oxygen enters the reaction zone as a component of air. If we do not take into account the content of other gases in the composition of dry air, we can state that the air is a mixture of oxygen and nitrogen in the proportions [11]:

Tab. 7.7: Air volume and mass composition [11].

	Volume composition [% volume]	**Mass composition [%]**
Oxygen O_2	21.0	23.2
Nitrogen N_2	79.0	76.8

Hence it follows that in order to supply 1 m^3 of oxygen to the combustion zone, it is actually necessary to add 4.76 m^3 of air, because every 1 m^3 of air is accompanied by 3.76 m^3 of nitrogen that is not involved in combustion. One mole of any gas (taken as an ideal gas) occupies the same volume under the same conditions, therefore the hydrogen combustion reactions can be represented as follows [11]:

$$\underset{1m^3}{H_2} + \underset{2.38m^3}{(0.5O_2 + 0.5 * 3.76N_2)} = \underset{1m^3}{H_2O} + \underset{1.88m^3}{1.88N_2}$$

If oxygen or air is supplied in the theoretically needed (i.e. stoichometric) quantity, then it is impossible to completely burn the fuel [13, 14]. At the final stage of combustion, the fuel and oxygen molecules will be so separated by the mass of the already formed combustion products that they will not be able to meet and react with each other, despite the fact that their total number corresponds to the stoichiometric equations. There is a phenomenon of incomplete combustion. To ensure the complete burnout of combustible components, the actual amount of air supplied to the combustion zone V_p should be bigger than theoretical [11].

$$V_p = \alpha * V_p^t$$

The dimensionless quantity α is called the excess air coefficient (excess oxidant). The value depends on the type of fuel, the design of the burner and furnace, which determine the degree of perfect mixing of fuel and air in the combustion zone. As it results from the above considerations, the condition of effective fuel combustion in most technological processes is the principle $\alpha > 0$, otherwise incomplete combustion is inevitable. For common gas burners the excess air factor is 1.05-1.15. A certain paradox is that the absolutely necessary excess air in the flame does not wear out, but only provides the conditions for the complete combustion of the fuel. All excess air goes to the exhaust gas composition.

7.7 Fuel cells

Fuel cells directly convert the chemical energy of a fuel into electricity. This type of conversion is an important advantage of cells because the efficiency of converting one energy into another is not limited by heat engine theory. Thus, it is possible to obtain efficiencies that exceed the efficiency of converting heat into mechanical energy at the currently mastered temperatures of heat supply to the cycle in which the heat engine operates (steam turbine, gas turbine). In the energy sector, the use of fuel cells in small and medium-sized power units is considered, also as a dispersed source of heat and electricity. There are many criteria for classifying fuel cells. The basic division is the links of direct use of a given fuel

and indirect use of its conversion (reforming). A typical representative of the first group is a hydrogen and oxygen powered cell. The cell to which methane or biogas and the oxidant are fed belongs to the second group. An important criterion for the division is the operating temperature of the cell. There are low-temperature ($25-100°C$), medium-temperature ($100-500°C$), high-temperature ($500-1000°C$) and especially high-temperature (over $1000°C$) cells. The technological division criterion is the type of electrolyte [4].

Tab. 7.8: Types of fuel cells [4].

Types of fuel cell, FC	**Type of electrolyte**	**Working temperature**
Alcaline, AFC	solution KOH (35-50%)	$60 - 200°C$
Polymer electrolyte, PEFC	polymer membrane	$50 - 80°C$
Phosphoric acid, PAFC	phosphoric acid in high concentration	$160 - 220°C$
Direct methanol, MCFC	molten mixture of lithium and sodium carbonates (Li_2CO_3/Na_2CO_3) or lithium and potassium (Li_2CO_3/K_2CO_3)	$620 - 650°C$
Solid oxide, SOFC	zirconium dioxide stabilized with yttrium (ZrO_2/Y_2O_3)	$800 - 1000°C$

The most commonly used fuel today is hydrogen, and the oxidant is oxygen in the air. For the cells presented in Tab. 7.8, the fuel is respectively:

- pure hydrogen H_2,
- pure hydrogen (CO content of less than 100 ppm),
- pure hydrogen ($CO_2 > 1\%$),
- hydrogen, carbon monoxide, methane (and other hydrocarbons), CO_2 is required at the cathode [3].

The following direct fuels are also considered:

- methane (CH_3OH),

- formic acid, systematically named methanoic acid (HCOOH),
- ammonia (NH_3),
- hydrazine (N_2,H_3),
- methane (CH_4),
- elemental carbon (C) and others [3].

Apart from the classification resulting from the type of fuel, electrolyte or operating temperature, it is possible to divide the cells taking into account the design and technological criteria. Hence the division into criteria:

- monopolar and bipolar,
- flat and tubular,
- pressure and atmospheric,
- single-chamber and two-chamber [4].

A two-chamber cell is a classic cell in which the cathodes and anodes are separated. In a single-chamber cell, the fuel and the oxidant are brought into one space, and in this case the electrodes must be distinguished by electrocatalytic selectivity [4].

7.8 Combustion of hydrogen in spark ignition engines

Spark ignition combustion engines operate in such a way that the prepared and compressed fuel-air mixture is ignited by the spark of the spark plug. The fuel-air mixture is introduced into the cylinder and passes through working cycles in which the chemical energy of the fuel is converted into thermal energy and energy of compressed exhaust gas capable of performing mechanical work. The condition for ignition and proper combustion of fuel is an appropriate concentration of the fuel and an appropriately high energy of the electric spark. Contemporary spark ignition engines use mixtures with a composition that allows to minimize fuel consumption and emission of toxic compounds at maximum power output under given conditions. The rate of combustion of the mixture in the cylinder

affects the temperature and pressure. Too high temperature (above the auto-ignition temperature) will cause self-ignition. As a result of such combustion, shock waves are created, which reflect from the cylinder walls, creating an acoustic effect. This type of combustion is very disadvantageous because:

- reduces engine efficiency,
- causes unevenness of work,
- it loads the elements of the piston-crank system, which increases the failure rate,
- increases fuel consumption,
- causes overheating of engine components and deterioration of lubrication [15].

The use of hydrogen in spark engines causes a lot of controversy, mainly due to the power obtained compared to the use of gasoline. The engine power expenditure is often masked with some extremely important advantages, such as efficiency or the range of functionality. The basic differences of these two fuels in the combustion process result from extremely different properties such as: density, volume energy density, minimum ignition energy, damping gap, combustion rate, combustion temperature, etc. [16].

A leak in a hydrogen engine causes a rapid volatilization of hydrogen and diffusion over long distances. The flame will be directed upwards, burning takes a relatively short time [15].

The general characteristics of the hydrogen combustion process show that hydrogen-powered internal combustion engines have almost zero emissions of toxic components, and their efficiency may exceed that of a conventional gasoline engine. This is due to two unique properties of hydrogen:

- this fuel does not contain any toxic ingredients, therefore the only toxic gas resulting from combustion is NO_x, created by thermal dissociation of nitrogen from the atmospheric air,
- low flammability threshold of lean hydrogen mixtures enables stable combustion of the mixture under high dilution conditions [15].

Since the combustion of ultra-lean hydrogen mixtures takes place at low temperatures, the combustion conditions can be selected so that the emission from the engine is close to zero. This unique, efficient and clean hydrogen combustion process only takes place at low engine load. In the case of high engine load, very low ignition energy of the hydrogen-air mixture causes many unfavorable events resulting in a significant increase in the combustion temperature, which results in the formation of a significant amount of NO_x. High self-ignition temperature of hydrogen (585°C) means that it is the most suitable fuel for spark-ignition engines. Nevertheless, the ignition energy of a hydrogen-air mixture is less than the energy required to ignite a hydrocarbon-air mixture. The low energy of the hydrogen-air mixture means that it has a predisposition to self-ignition. Pre-ignition is also affected by overheated engine components: spark plugs, valves and debris from improper combustion. In recent years, significant progress has been made in the use of hydrogen as a fuel in spark engines [15].

7.9 Combustion of hydrogen in diesel engines

A typical diesel engine, like a spark ignition engine, is a 4-stroke engine with the following cycles: filling, compression, operation (mixture combustion and gas expansion) and exhaust. The cylinder of this engine has separate fuel and air supply. Air is first compressed in the cylinder to a pressure of 1.4-2.5 MPa. During compression, the temperature rises to 500-600°C and fuel is finally injected into the cylinder. After a few milliseconds the fuel will self-ignite. In the next milliseconds, the fuel is fed into the cylinder where it evaporates and burns. during combustion, the temperature of the gases in the cylinder can reach 1600-2000°C and the pressure 6-10 MPa. Combustion in a diesel engine takes place in a fuel/air ratio of 0.75/0.85 [15].

Hydrogen in mono fuel mode is not applicable to compression ignition engines. Hydrogen may be used in a dual fuel hydrocarbon pilot fuel mode.

The influence of hydrogen addition to diesel fuel on thermal efficiency, smoke emission, the amount of HC, CO and NO_x emitted, ignition delay, increase in cylinder pressure and combustion rate was investigated in [15, 17].

The tests were carried out on a single 3.7 kW water-cooled cylinder, at 1500 rpm, with a compression ratio of 1:15, where the injection pressure was 170 atm [15, 17].

The conducted research shows that:

- by adding 5-7% m/m H_2, the thermal efficiency of the tested system increases by approx. 2%,
- at low loads, the addition of hydrogen reduces the efficiency of the tested system,
- with a higher power output, the combustion rate is higher, and with a lower power output, it is lower due to the emission of a small amount of heat,
- the addition of hydrogen changes the opacity of exhaust gases by about 55% for diesel fuel,
- the addition of hydrogen increases NO_x by about 15% [15, 17].

The research shows that the addition of hydrogen in the amount of 5-7% m/m during diesel fuel combustion in self-ignition engines does not have a significant impact on the combustion effects of such mixtures [15, 17].

7.10 Combustion of hydrogen in gas turbines

Increased interest in hydrogen is shown, among others, by electricity producers using gas turbines for this purpose. Compared to classic technologies, using them seems simpler and more functional [15]. Hydrogen can be used as the primary fuel in gas turbines. Research on combustion processes of fuels with an increased share of hydrogen are carried out, among others in connection with the preparation of a new class of gas turbines in gas-steam systems integrated with coal gasification. Program of the United States Department of Energy (DOE)- the Advanced Hydrogen turbine Development pro-

gram, aims to increase the efficiency of the gas and steam system by 3-5%, reduce investment costs by 20-30% and reduce NO_x emissions to 2ppm (15% O_2). The main goal of the project is to develop new structures of the vane system with increased efficiency, cooling intensity and thermal resistance. Achieving the assumed emission level of nitrogen oxides requires the use of catalytic technology of their reduction in exhaust gases. In the coal gasification gas, the hydrogen content is usually in the range of 25-35% by volume. High power designs are available for this level of participation. In the case of a coal gasification installation with carbon dioxide capture, the share of hydrogen in the fuel exceeds 80%. Company MHPS announces the generation of pure hydrogen gas turbines for 2025 [4].

7.11 Examples of the use of hydrogen in ship propulsion

7.11.1 The Energy Observer

Energy Observer is the vessel autonomous in energy. The boat for ship propulsion uses the energy generated by solar panels and wind turbines. There is also a hydrogen obtained by electrolysis. Energy observer produces hydrogen directly on board. Hydrogen produced in the electrolysis process at a pressure of 30 bar is compressed to 350 bar. Then it is stored in tanks with a total capacity of 322 litres. It is equipped with hydrogen cells [18].

Principal parameters:

- length 30.5 meters,
- width 12.80 meters,
- weight 28 metric tons,
- speed 8-10 knots [18].

7.11.2 The ULSTEIN SX190 Zero Emission DP2

The ULSTEIN SX190 Zero Emission DP2 construction support vessel is Ulstein's first hydrogen powered offshore vessel featuring a Nedstack fuel cell power system. The design uses proven and

available technology, enabling clean shipping operations to reduce the environmental footprint of offshore projects [19].

Principal parameters:

- length: 99 m,
- beam: 23.4 m,
- deck area: 1000-1200 sqm,
- draught (min. In DP): 6.0 m,
- accommodation: 60-90 POB,
- propulsion thrusters: 2 x 1.280 kW,
- tunnel thrusters: 2 x 750 kW,
- deck strength: 10 t/m^2 [19].

The vessel in stage I is to work without refueling 4 days in a row. After refining the hydrogen drive, work autonomy on one hydrogen bunkering is expected to take up to 14 days. Hydrogen bunkering for the ship will be carried out by replacing containers with refueled fuel. The hydrogen will be delivered to the ship in specially prepared containers ISO according to the requirements of ADR (Agreement of 30 September 1957 concerning the International Carriage of Dangerous Goods by Road). The hydrogen tanks will be filled at the production site or hydrogen distribution [18].

7.12 Conclusions

In conclusion, all the above-mentioned methods of using hydrogen as ship fuel are possible. Pure hydrogen can be used in fuel cells, spark ignition engines and gas turbines. Hydrogen in mono fuel mode is not applicable to compression ignition engines. Hydrogen can be used in a dual fuel hydrocarbon pilot fuel mode. Hydrogen is essentially an inexhaustible, clean, storage and transportable energy medium. More research is needed to develop and adapt hydrogen technology in the maritime sector.

Bibliography

[1] Energia ze źródeł odnawialnych w 2020 roku. www.stat.gov.pl/obszary-tematyczne/srodowisko-energia/ energia/energia-ze-zrodel-odnawialnych-w-2020-roku,10,4.html, Głowny Urząd Statystyczny.

[2] International convention for the prevention of pollution from ships. www.imo.org/en/About/Conventions/Pages/International-Convention-for-the-Prevention-of-Pollution-from-Ships-(MARPOL).aspx, International Maritime Organization.

[3] The european green deal. www.ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal, European Commission.

[4] Tadeusz Chmielniak and Tomasz Chmielniak. *Energetyka wodorowa*. Wydawnictwo Naukowe PWN SA, 2020.

[5] Padi Boyd. What is the chemical composition of stars? Report, NASA.

[6] Ahmet Ozarslan. Large-scale hydrogen energy storage in salt caverns. *International journal of hydrogen energy*, 37(19):14265–14277, 2012.

[7] Anna S Lord, Peter H Kobos, and David J Borns. Geologic storage of hydrogen: Scaling up to meet city transportation demands. *International journal of hydrogen energy*, 39(28):15570–15582, 2014.

[8] Christoph Noack, Fabian Burggraf, Seyed Schwan Hosseiny, Philipp Lettenmeier, Svenja Kolb, Stefan Belz, Josef Kallo, K Andreas Friedrich, Thomas Pregger, Karl-Kiên Cao, et al. Studie über die planung einer demonstrationsanlage zur wasserstoff-kraftstoffgewinnung durch elektrolyse mit zwischenspeicherung in salzkavernen unter druck. 2015.

[9] Wade A Amos. Costs of storing and transporting hydrogen. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 1999.

[10] JL Gillette and RL Kolpa. Overview of interstate hydrogen pipeline systems. Technical report, Argonne National Lab.(ANL), Argonne, IL (United States), 2008.

[11] Aleksander Szkarowski. *Paliwa gazowe: podstawy efektywnego i ekologicznego wykorzystania*. PWN, 2020.

[12] Bernard Lewis and Guenther von Elbe. Experimental evidence for and determination of the predicted δ 1 level of the neutral oxygen molecule. *Physical Review*, 41(5):678, 1932.

[13] R Wilk. Podstawy niskoemisyjnego spalania. Technical report, Komisja Energetyki PAN, Oddzial Slaski, 2000.

[14] Stanisław Wójcicki. *Spalanie*. Wydawnictwa Naukowo-Techniczne, 1969.

[15] Jan Surygała. *Wodór jako paliwo*. Wydawnictwa Naukowo-Techniczne, 2008.

[16] LM Das. Hydrogen engines: a view of the past and a look into the future. *International Journal of Hydrogen Energy*, 15(6):425–443, 1990.

[17] M Senthil Kumar, A Ramesh, and B Nagalingam. Use of hydrogen to enhance the performance of a vegetable oil fuelled compression ignition engine. *International Journal of Hydrogen Energy*, 28(10):1143–1154, 2003.

[18] Course: Projektowanie obiektow plywajacych – projekt zespolowy. Technical report.

[19] Sx190. www.ulstein.com/vessel-design/sx190, Ulstein Design and Solutions B.V.

Chapter 8

Subgrid-scale modeling for particle motion based on multifractal scale-similarity in the enstrophy field

Michał Rajek

Institute of Fluid Flow Machinery, Polish Academy of Sciences,
Fiszera 14, 80-231 Gdansk, Poland

8.1 Introduction

Turbulent dispersed two-phase flows are indispensable part of many industrial and environmental processes. These can include among many others: fuel atomization and mixing, spray combustion, cyclone separation, pneumatic transport, bubble column reactors, but also volcanic processes, sediment transport, warm rain formation or even the motion of dust particles in protoplanetary disks.

A more profound understanding of the dispersed phase dynamics, particularly concerning preferential concentration, preferential sweeping, particle collisions and/or coalescence, is critical for further advancements in the aforementioned fields. Due to high complexity of such flows associated with a wide range of spatial and temporal scales, the dispersed phase is most often accounted for through various numerical methods comprehensively reviewed by Kuerten [1] and Elghobashi [2]. For the sake of clarity, it should be underlined that as long as dilute suspensions of particles smaller than the Kolmogorov length scales are concerned, the dispersed entities can be safely approximated as point-particles. Therefore, the Eulerian description of a carrier phase motion is typically combined with the Lagrangian tracking of such individual mass-points leading to the so-called point-particle direct numerical simulations or large eddy simulations. It is also worth mentioning that the carrier phase motion can also be solved by means of the Reynolds-Averaged Navier-Stokes simulations (RANS), however, this approach is beyond the scope of this work.

It must be stressed that despite significant growth in computing performance, direct numerical simulations (DNS), although providing the most accurate description, will not be feasible for the majority of practical applications for many years to come. For this reason, large eddy simulations (LES) have emerged as a promising tool for particulate two-phase flows in the past 20 years. As a result of spatial filtering only the large, energy-containing, eddies are resolved, allowing for reduced computational cost when compared to DNS. Small, unresolved eddies are typically accounted for through the so-called subgrid-scale (SGS) stress tensor found in the momentum conservation equation, often modeled in terms of an artificial viscos-

ity. Due to the lack of SGS motions, single- and two-point particle statistics obtained by means of point-particle DNS and LES diverge with decreasing particle inertia [3]. For this reason, point-particle LES most often require "two-stage" SGS modeling. In most practical applications the "first-stage" is meant to ensure the large-scale motions are to be properly resolved, whereas the "second-stage" should precisely enhance particle motions. Such an approach typically requires an explicit coupling between the stages to be enforced basing on additional assumptions. An ideal "single-stage" structural SGS model, in which small-scale motions are introduced through the available large-scale quantities, remains unreachable. In such an approach the synthetic small-scale velocity field would be used simultaneously to determine the SGS stress tensor in the momentum equation, but also to enhance the fluid velocity at the particle locations.

In this work we focus on the multifractal scale-similarity observed in the enstrophy field, and revisit the model of Burton & Dahm [4] developed in the framework of the "first-stage" SGS modeling. We also attempt to assess the possibility of extending the proposed model to particle-laden flows. Specifically, we compared the SGS vorticity magnitude field obtained from both the SGS multifractal model and the high-pass filtered homogeneous and isotropic turbulence.

8.2 Previous "first-stage" modeling approaches

If the filtering operator is applied to the Navier-Stokes equations governing the motion of an incompressible fluid, a set of equations describing the large-scale dynamics is obtained as follows:

$$\begin{aligned} (a) \quad & \frac{\partial \overline{u}_i}{\partial t} + \frac{\partial \overline{u}_i \overline{u}_j}{\partial x_j} = -\frac{1}{\rho}\frac{\partial \overline{p}}{\partial x_i} + \nu \frac{\partial^2 \overline{u}_i}{\partial x_i \partial x_j} - \frac{\partial \tau_{ij}}{\partial x_j} \\ (b) \quad & \frac{\partial \overline{u}_i}{\partial x_i} = 0 \end{aligned} \tag{8.1}$$

where u_i is the velocity field, p is the pressure field, ρ and ν are the density and kinematic viscosity of the fluid, respectively. The overbar $\overline{(\cdot)}$ denotes the filtering defined as the convolution with the filter function G:

$$\overline{u}_i = G * u_i, \quad \overline{u}_i(\mathbf{x}, t) = \int G(\mathbf{x} - \mathbf{x}', t) u_i(\mathbf{x}', t) d\mathbf{x}'. \tag{8.2}$$

The net impact of the subgrid-scale eddies on the resolved ones is determined through the SGS stress tensor defined as $\tau_{ij} = \overline{u_i u_j} - \overline{u}_i \overline{u}_j$. A key goal of the "first-stage" SGS modeling is to express τ_{ij} in terms of the resolved (large-scale) quantities.

The traceless part of the SGS stress tensor is often assumed proportional to the resolved rate-of-strain tensor $\overline{S}_{ij} = (\partial \overline{u}_i / \partial x_j + \partial \overline{u}_j / \partial x_i)$ according to [5]:

$$\tau_{ij} - \frac{1}{3} \tau_{kk} \delta_{ij} = -2 \nu_t \overline{S}_{ij}, \tag{8.3}$$

with the eddy viscosity $\nu_t \approx (C_s \Delta)^2 |\overline{S}|$, as first proposed by Smagorinsky [6]; C_s is a model constant, Δ is a filter width and $|\overline{S}| = (2 \overline{S}_{ij} \overline{S}_{ij})^{1/2}$ is the magnitude of the rate-of-strain tensor. The model is strictly dissipative and yields $\nu_t \geq 0$ even in laminar flows, hence it is not suitable for simulating transition, and requires additional adjustments near the walls. This model was further extended by Germano et al. [7] who proposed a dynamic evaluation of the model parameter C_s based on double-filtering of the large-scale velocity field. The primary advantage of this approach is that the model does not contain any user-defined parameters. The parameter C_s varies in time and space. It is also reduced near the walls and vanishes for well-resolved laminar flows. On the other hand, C_s exhibits large oscillations and thus has to be averaged spatially over homogeneous directions [8]. This model was further generalized for the LES of compressible flows and advection of passive scalars by Moin et al. [9]. Subsequently Lilly [10] proposed the use of least-squares technique to determine the model parameter C_s. Ghosal et al. [11] further extended the model for inhomogeneous flows.

Kraichnan [12] and Chollet & Lesieur [13] proposed a wave-number-dependent eddy-viscosity for homogeneous turbulence. In the lat-

ter, the eddy viscosity is defined as follows:

$$\nu_e(k|k_c) = C_K^{-3/2}[0.441 + 15.2\exp(-3.03k_c/k)]\sqrt{\frac{E(k_c)}{k_c}}, \quad (8.4)$$

where C_K is a model constant, k_c is the cutoff wavenumber, i.e. the maximum wavenumber in LES, $E(k_c)$ is the kinetic energy associated with the cutoff wavenumber. However, when compared to DNS, the model turns out to overpredict the energy spectrum deep in the inertial range, but, on the other hand, underpredicts $E(k)$ near the cutoff.

In general, functional models not only sufficiently accurately predict the transport of kinetic energy within the resolved scales but also the transfer from the resolved eddies to unresolved ones. The main drawbacks of such models, however, are that they generally cannot predict both the correlations between modeled and actual SGS quantities and the so-called backscatter, i.e. the transfer of energy from smaller to larger scales. According to recent findings, however, it seems not necessary for the SGS model to account for the backscatter in order to obtain accurate and stable LES [14].

Structural models, on the other hand, aim at reconstructing the actual small scales that directly influence the evolution of the resolved ones, and are usually used to address the aforementioned deficiencies of the functional approach. Bardina et al. [15, 16] proposed a scale-similarity model that presupposes equality between the smallest resolved and largest unresolved scales leading to:

$$\tau_{ij} \approx \overline{\overline{u_i}\,\overline{u_j}} - \overline{\overline{u_i}}\,\overline{\overline{u_j}} \quad (8.5)$$

The model accounts for the backscatter, but, on the other hand, turns out to underpredict the SGS dissipation. Hence, it is typically combined with the Smagorinsky model. Stolz & Adams proposed an approximate deconvolution method (ADM) procedure for LES in which a truncated series expansion of the inverse filter operator is used to obtain an approximation of the unfiltered velocity field. ADM, however, does not account for any explicit SGS motions and thus enhances scales near the cutoff only. Domaradzki proposed the so-called velocity estimation procedure developed in Fourier

space [17] and subsequently extended it to physical space [18]. In this approach the unfiltered velocity field appearing in the SGS stress tensor, τ_{ij}, is approximated at scales two times smaller than the filter width, i.e. $\Delta/2$, as the evolution of the resolved scales ($k \leq k_c$) is mostly governed by the interaction between scales for which the wavenumber is less than $2k_c$ [19]. Initial SGS velocity components are obtained through the deconvolution of the resolved velocity field. These are further adjusted to match the SGS velocity field obtained by means of nonlinear interactions between the resolved scales. Burton & Dahm [4, 20] proposed a new approach to modeling the unfiltered velocity components appearing in τ_{ij} based on the multifractal scale-similarity in the enstrophy field observed in high Reynolds number turbulence. The model accounts for the backscatter and shows good correlations between modeled and actual SGS quantities in both *a priori* and *a posteriori* tests. Scotti & Meneveau [21] calculated τ_{ij} explicitly from a synthetic SGS velocity field obtained through the so-called fractal interpolation. In this approach the resolved velocity field is interpolated down to scales much smaller than the LES grid scale by means of an iterative mapping. The model turns out not only to be sufficiently dissipative but also yields good correlations between the modeled and actual SGS quantities.

8.3 Previous "second-stage" modeling approaches

Ray & Collins [22] studied the so-called kinematic simulations (KS) as a means to enrich the fluid velocity at the particle locations in a filtered homogeneous and isotropic turbulence. In such an approach the fluid velocity seen by inertial particles, $\mathbf{u}^{\text{sgs}}$, is determined as follows:

$$\mathbf{u}^{\text{sgs}}(\mathbf{x},t) = \sum_{n=1}^{N}\sum_{m=1}^{M} (\mathbf{a}_{nm} \times \hat{\mathbf{k}}_{nm}) \cos(\mathbf{k}_{nm}\cdot\mathbf{x} + \omega_n t) + (\mathbf{b}_{nm} \times \hat{\mathbf{k}}_{nm}) \sin(\mathbf{k}_{nm}\cdot\mathbf{x} + \omega_n t). \quad (8.6)$$

where $\hat{\mathbf{k}}_{nm} = \mathbf{k}_{nm}/|\mathbf{k}_n|$. The energy is essentially distributed over N shells of radius $|\mathbf{k}_n|$, each containing M randomly oriented wavevectors. The velocity field is non-divergent, statistically stationary, homogeneous and isotropic. Coefficients $\mathbf{a}_{nm}$ and $\mathbf{b}_{nm}$ are randomly selected vectors with independent and normally distributed components, each with zero mean and variance $\sigma_n^2 = E(k_n)\Delta k_n$, where $E(k_n)$ is a prescribed energy spectrum typically obtained from a reference DNS. The Authors studied specific two-point particle statistics and reported some improvement for particles of larger inertia. Voßkuhle et al. [23] studied particle collisions in both DNS and KS-based synthetic velocity field and observed good agreement between the two only in the limit where the influence of particle inertia is negligible. Zhou et al. [24] combined KS with ADM and studied relative dispersion of flow tracers in LES of homogeneous isotropic turbulence. The Authors observed substantial improvement in the prediction of Lagrangian dispersion of passive particles for LES when compared to a reference DNS. The Authors extended the model to particle-laden flows [25], however, did not report any further improvement when compared to similar work reported by Ray & Collins [22]. Recently, Rajek & Pozorski [26] studied the impact of KS on particle settling velocity and collision-related statistics in *a priori* LES of homogeneous and isotropic turbulence. The Authors indicated that, due to the lack of the sweeping effect, and possibly SGS nonlinear interactions, KS is incapable of retrieving the instantaneous velocity field observed in DNS and, consequently, cannot precisely enhance particle motions in *a priori* LES. Bassenne et al. [27] extended the velocity estimation procedure developed in physical space [18] to particle-laden flows. In this approach SGS velocity field is iteratively reconstructed at scales smaller than the LES filter width, and thus, at least to some extent, allowing for approximate reconstruction of two-point particle statistics. The Authors reported a moderate improvement of particle statistics when compared to LES without any SGS model for particle motion. It should also be pointed out that the reconstructed SGS velocity field is enforced to be non-divergent through a classic Hemholtz decomposition which can be prohibitively expensive.

8.4 Multifractal SGS modeling for single-phase flows

Burton & Dahm [4, 20] proposed a novel approach to approximate the SGS velocity field components appearing in τ_{ij} basing on the multifractal representation of the SGS vorticity field. This approach has a strong theoretical foundation indicating that gradient fields, such as the enstrophy [28, 29] or the turbulence energy dissipation rate [30–32], exhibit multifractal scale-similarity over inertial-range scales in high Reynolds number turbulence. A self-similar or otherwise multifractal measure, is said to be statistically indistinguishable at all scales. The Authors proposed to retrieve the SGS velocity components from a multifractal SGS vorticity field by means of the Biot-Savart integral as follows:

$$\begin{aligned} \mathbf{u}^{\mathrm{sgs}}(\mathbf{x},t) &= \frac{1}{4\pi}\int_{\Omega_{\Delta}} \boldsymbol{\omega}^{\mathrm{sgs}}(\mathbf{x}',t)\times\mathbf{K}(\mathbf{x},\mathbf{x}')d\Omega_{\Delta}, \\ \mathbf{K}(\mathbf{x},\mathbf{x}') &= \frac{\mathbf{x}-\mathbf{x}'}{|\mathbf{x}-\mathbf{x}'|^3}, \end{aligned} \tag{8.7}$$

where $\boldsymbol{\omega}^{\mathrm{sgs}}$ is the SGS vorticity field within each resolved-scale grid cell of size Δ. It should be pointed out that the Biot-Savart integral, Eq. (8.7), is evaluated for each Δ-scale grid cell separately. The SGS vorticity field is approximated in two steps. During the first step, a scale-invariant multiplicative cascade distributes the SGS vorticity magnitudes within each Δ-scale grid cell according to:

$$|\boldsymbol{\omega}^{\mathrm{sgs}}|(\mathbf{x},t) = \left[Q_{\mathrm{sgs}}(2^{\mathcal{N}})^3\prod_{n=1}^{\mathcal{N}}\mathcal{M}_n(\mathbf{x},t)\right]^{1/2}, \tag{8.8}$$

where $\mathcal{N}\equiv log_2(\Delta/\lambda_\nu)$ is the number of cascade steps; $\lambda_\nu\approx 5.9\lambda_K$ is the viscous length scale [29] and λ_K is the Kolmogorov length scale. Multipliers $\mathcal{M}_n$ are randomly selected at each step of the multiplicative cascade from a scale-invariant distribution $P(\mathcal{M})$, precomputed by reversing the multiplicative process in the DNS velocity field [33, 34]. Q_{sgs}, i.e. the total SGS enstrophy within

each resolved-scale cell of size Δ, is approximated as:

$$Q_{\text{sgs}} = \int_{k_\Delta}^{k_\nu} \mathcal{Q}(k)dk = \kappa Q_\Delta \left[\left(\frac{k_\nu}{k_\Delta} \right)^{4/3} - 1 \right], \tag{8.9}$$

on condition that k_Δ lies within the inertial range, and thus the enstrophy spectrum $\mathcal{Q}(k)$ scales with the wavenumber as $\mathcal{Q}(k) \sim k^{1/3}$. In the above equation $\kappa \equiv (1 - 2^{-4/3})^{-1}$ and $Q_\Delta = \omega_i^\Delta \omega_i^\Delta$ where $\boldsymbol{\omega}^\Delta = \nabla \times \mathbf{u}^\Delta$ is the vorticity field associated with the scales range between Δ and 2Δ. The velocity field $\mathbf{u}^\Delta$ can be obtained through filtering according to $u_i^\Delta = \overline{u}_i - (\overline{u}_i)^{2\Delta}$ where $(\overline{u}_i)^{2\Delta}$ is the resolved velocity field, $\overline{u}_i$, filtered at scale 2Δ.

During the second step of the SGS vorticity field generation process, an additive cascade is introduced to ensure isotropic decorrelation of the SGS vorticity orientations between subsequent subgrid scales as the cascade progresses from the smallest resolved scales Δ down to the smallest, unresolved ones λ_ν. The Authors determined each component of the orientation unit vector at stage $(n+1)$ as:

$$\hat{e}_i^{n+1} = \hat{e}_i^n + \delta_i(\phi, \theta)^{n+1}, \tag{8.10}$$

where ϕ and θ are stochastic spherical decorrelation angles in a local coordinate system aligned with $\hat{\mathbf{e}}^\Delta$. However, it should be emphasized that the orientation vector, Eq. (8.10), is not a unit vector, i.e. $|\hat{\mathbf{e}}^{n+1}| = |\hat{\mathbf{e}}^n + \boldsymbol{\delta}(\phi, \theta)^{n+1}| \neq 1$, and requires additional normalization. Nevertheless, the Authors proposed to determine the SGS vorticity field as follows:

$$\boldsymbol{\omega}^{\text{sgs}}(\mathbf{x}, t) = |\boldsymbol{\omega}^{\text{sgs}}| \left(\mathcal{I}(\mathcal{N}) \hat{\mathbf{e}}^\Delta(\mathbf{x}, t) + (1 - \mathcal{I}) \sum_{n=1}^{\mathcal{N}} \boldsymbol{\delta}_n \right), \tag{8.11}$$

where $\boldsymbol{\delta}_n$ is the decorrelation vector in the orientation cascade, Eq. (8.10), and $\mathcal{I}$ is the intermittency factor defined as:

$$\mathcal{I} = \int_{\Omega_\Delta} \boldsymbol{\omega}^{\text{sgs}} \cdot \boldsymbol{\omega}^\Delta d\Omega_\Delta \Big/ \int_{\Omega_\Delta} |\boldsymbol{\omega}^{\text{sgs}}||\boldsymbol{\omega}^\Delta| d\Omega_\Delta. \tag{8.12}$$

The vorticity field in its current form, Eq. (8.11), was designed to retrieve the tendency of the strongest vortical structures to remain

aligned with the local strain rate tensor [4]. According to the Authors, this corresponds to high multiplier values $\mathcal{M}_n$. With respect to the aforementioned considerations, the SGS vorticity field should somehow take into account the correlation between $\mathcal{M}_n$ and $\boldsymbol{\delta_n}$. The Authors, however, assumed such correlations to be negligible and thus expressed the expected value of the small-scale vorticity field as:

$$\langle \boldsymbol{\omega}^{\mathrm{sgs}} \rangle = \mathcal{I}(\mathcal{N}) \langle |\boldsymbol{\omega}^{\mathrm{sgs}}| \rangle \hat{\mathbf{e}}^{\Delta}(\mathbf{x}, t) + (1 - \mathcal{I}) \langle |\boldsymbol{\omega}^{\mathrm{sgs}}| \rangle \sum_{n=1}^{\mathcal{N}} \langle \boldsymbol{\delta}_n \rangle. \quad (8.13)$$

From (8.8) and (8.13) they obtained

$$\langle \boldsymbol{\omega}^{\mathrm{sgs}} \rangle = \mathcal{I}(\mathcal{N}) (2^{\mathcal{N}})^{3/2} \langle (\mathcal{M}_1 \dots \mathcal{M}_{\mathcal{N}})^{1/2} \rangle Q_{\mathrm{sgs}}^{1/2} \hat{\mathbf{e}}^{\Delta}, \quad (8.14)$$

assuming the orientation cascade to be fully isotropic. i.e. $\langle \boldsymbol{\delta}_n \rangle \equiv 0$. Furthermore, the expected value of the SGS velocity field $\langle \mathbf{u}^{\mathrm{sgs}} \rangle$ was obtained through the Biot-Savart integral, Eq. (8.7), as follows:

$$\begin{aligned} \langle \mathbf{u}^{\mathrm{sgs}} \rangle = {} & \mathcal{I}(\mathcal{N}) (2^{\mathcal{N}})^{3/2} \langle (\mathcal{M}_1 \dots \mathcal{M}_{\mathcal{N}})^{1/2} \rangle \\ & \times \left(\frac{Q_{\mathrm{sgs}}}{Q_{\Delta}} \right)^{1/2} \frac{1}{4\pi} \int_{\Omega_{\Delta}} Q_{\Delta}^{1/2} \hat{\mathbf{e}}^{\Delta} \times \mathbf{K} \, d\Omega_{\Delta}, \end{aligned} \quad (8.15)$$

eventually leading to:

$$\langle \mathbf{u}^{\mathrm{sgs}} \rangle = \mathcal{I}(\mathcal{N}) (2^{\mathcal{N}})^{3/2} \langle \mathcal{M}^{1/2} \rangle^{\mathcal{N}} [2^{(4/3)\mathcal{N}} - 1]^{1/2} \sqrt{\kappa} \mathbf{u}^{\Delta}. \quad (8.16)$$

It is worth mentioning that if the Biot-Savart integral was evaluated properly, i.e over the entire vorticity field, such simplification could not be possible as Q_{sgs} is essentially a function of space and time. Finally the Authors approximate the SGS velocity components according to:

$$u_i^{\mathrm{sgs}} \approx \mathcal{I}(\mathcal{N}) (2^{\mathcal{N}})^{3/2} \langle \mathcal{M}^{1/2} \rangle^{\mathcal{N}} [2^{(4/3)\mathcal{N}} - 1]^{1/2} \sqrt{\kappa} \mathbf{u}^{\Delta}, \quad (8.17)$$

where the intermittency factor $\mathcal{I}$, Eq. (8.12), is further approximated as:

$$\mathcal{I}(\mathcal{N}) \approx C_{\mathcal{I}} 2^{-(2/3+3/2)\mathcal{N}} \langle \mathcal{M}^{1/2} \rangle^{-\mathcal{N}}, \quad (8.18)$$

where $C_{\mathcal{I}} \approx 0.37$ is a model constant obtained through *a priori* testing. Equations (8.17) and (8.18) constitute the multifractal SGS model and are used to approximate the SGS stress tensor according to:

$$\tau_{ij} = \overline{\overline{u}_i \overline{u}_j} - \overline{u}_i \overline{u}_j + \overline{\overline{u}_i u_j^{\mathrm{sgs}}} + \overline{u_i^{\mathrm{sgs}} \overline{u}_j} + \overline{u_i^{\mathrm{sgs}} u_j^{\mathrm{sgs}}}. \tag{8.19}$$

8.5 Multifractal SGS modeling for particulate two-phase flows

In this work we studied the multifractal SGS model introduced in Section 8.4 and took a necessary step to assess its possible extension to particulate two-phase flows. We specifically concerned the instantaneous structure of the SGS vorticity magnitude field, retrieved through the application of the multiplicative cascade given by Eq. (8.8), and subsequently compared it with the SGS vorticity magnitude field obtained from a high-pass filtered homogeneous and isotropic turbulence.

We performed the DNS of homogeneous isotropic turbulence by means of the standard pseudo-spectral method [35] in a periodic cube with an edge length of 2π. In Fourier space, equations governing the motion of an incompressible fluid become:

$$\left(\frac{\partial}{\partial t} + \nu k^2\right) \hat{\mathbf{u}}(\mathbf{k}, t) = \mathbf{P}(\mathbf{k}) \mathcal{F}(\mathbf{u} \times \boldsymbol{\omega}) + \hat{\mathbf{f}}(\mathbf{k}, t) \qquad \text{for} \quad k \leq k_{\max}, \tag{8.20}$$

where $\hat{\mathbf{u}}(\mathbf{k}, t)$ is the fluid velocity in Fourier space, $\mathcal{F}$ denotes the Fourier transformation, $P_{ij} = \delta_{ij} - k_i k_j / k^2$ $(i, j = 1, 2, 3)$ is the projection tensor, and $\hat{\mathbf{f}}$ is an artificial forcing used to sustain turbulence [36]. The computational domain was uniformly discretized with $N = 256^3$ points allowing to accurately resolve a turbulent flow with the Reynolds number based on the Taylor microscale $Re_\lambda \approx 107$. The spatial resolution was monitored by the value of $k_{\max}\eta$ which was kept larger than 1.5. An ideal SGS velocity field, i.e. deprived of any modeling errors, was furthermore obtained

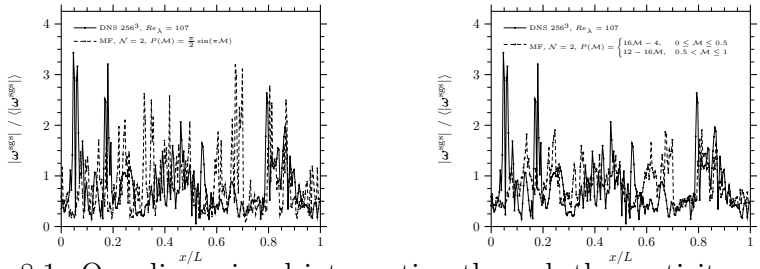


Fig. 8.1: One-dimensional intersection through the vorticity magnitude field.

through the high-pass filtering according to:

$$\hat{\hat{\mathbf{u}}}_{\text{sgs}} = \begin{cases} \hat{\mathbf{u}}(\mathbf{k}, t) & \text{if } |\mathbf{k}| > k_c \\ 0 & \text{otherwise,} \end{cases} \tag{8.21}$$

where $k_c = 30$ which corresponds to LES with a 64^3 grid resolution. Hence, the multiplicative process is expected to have at maximum $\mathcal{N} = 2$ steps. It should be emphasized that in such case, however, the multifractal model will introduce scales comparable to the smallest dissipative scales, λ_K, rather than viscous scales, λ_ν, thus reaching out beyond the model assumptions regarding the inertial range scaling, Eq. (8.9). The multifracal SGS vorticity magnitude field was obtained according to Eq. (8.8), where the scale-invariant distribution of multipliers was approximated according to [37] as:

$$P(\mathcal{M}) = \frac{\pi}{2}\sin\pi\mathcal{M}, \quad 0 < \mathcal{M} < 1 \tag{8.22}$$

In principle, the multiplicative cascade generates a total of $2^{\mathcal{N}}$ cells of size λ_ν in each spatial direction, however, only $2^{\mathcal{N}} - 1$ multipliers are randomly selected to ensure the conservation of quantity being distributed. Fig. 8.1 shows a simple 1D intersection through the vorticity magnitude field. It is readily observed that the multifractal model, Fig. 8.1 (left), is incapable of retrieving the instantaneous SGS vorticity field, at least in terms of the vorticity magnitudes and spatial correlations. Such deviations may result from the adopted scale-invariant distribution of multipliers, Eq. (8.22). To address

this issue, we defined an additional triangular distribution of multipliers, Fig. 8.2, as follows:

$$P(\mathcal{M}) = \begin{cases} 16\mathcal{M} - 4, & 0.25 \leq \mathcal{M} \leq 0.5 \\ 12 - 16\mathcal{M}, & 0.5 < \mathcal{M} \leq 0.75. \end{cases} \tag{8.23}$$

It is readily observed, Fig. 8.1 (right), that such a distribution ensures better agreement with the DNS SGS vorticitiy magnitude field. In general, the multiplier $\mathcal{M}_\epsilon$ between two successive scales at any point $(\mathbf{x}_i, t_i)$ is determined according to [32] as:

$$\mathcal{M}_\epsilon(\mathbf{x}_i, t_i) \equiv \frac{\int_{|\mathbf{x}-\mathbf{x}_i|<\epsilon/2} \int_{|t-t_i|<\epsilon/2} \mu(\mathbf{x}, t) \mathrm{d}\mathbf{x}\mathrm{d}t}{\int_{|\mathbf{x}-\mathbf{x}_i|<\epsilon} \int_{|t-t_i|<\epsilon} \mu(\mathbf{x}, t) \mathrm{d}\mathbf{x}\mathrm{d}t}, \tag{8.24}$$

where μ denotes the field being distributed. The multiplier distribution $P(\mathcal{M}_\epsilon)$ is determined at each scale ϵ and must subsequently be generalized for scales range at which the scale-similarity applies. Ideally, for an autonomous LES, the scale-invariant distribution of multipliers $P(\mathcal{M})$ should be parameterized by means of the resolved quantities only.

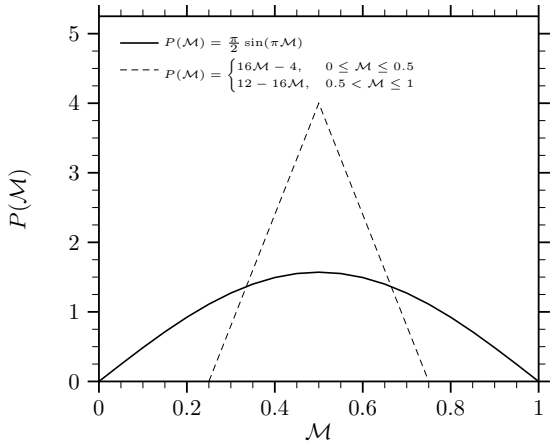


Fig. 8.2: Scale-invariant distribution of multipliers.

8.6 Conclusions and outlook

We revisited the multifractal SGS model introduced by Burton & Dahm [4] in the framework of the "first-stage" subgrid-scale modeling for LES of homogeneous and isotropic turbulence. This approach has a strong theoretical foundation indicating that gradient fields, particularly concerning the enstrophy field, exhibit multifractal scale-similarity over inertial-range scales in high Reynolds number flows. In this approach the scale-invariant multiplicative cascade, Eq. (8.8), distributes the SGS vorticity magnitude within each resolved-scale grid cell, whereas the additive cascade, Eq. (8.10), ensures isotropic decorrelation of the SGS vorticity orientations from the resolved scale Δ to the viscous scale λ_ν. Once the SGS vorticity field is determined, the SGS velocity field is obtained by solving the Biot-Savart integral, Eq. (8.7), for each resolved-scale grid cell separately.

If LES of particulate two-phase flows is concerned, the SGS velocity field should be explicitly included to enhance particle motions according to:

$$\begin{aligned} \frac{\mathrm{d}\mathbf{V}_p}{\mathrm{d}t} &= -\frac{\mathbf{V}_p - (\widetilde{\mathbf{u}} + \mathbf{u}^{\mathrm{sgs}})}{\tau_p} + \mathbf{g}, \\ \frac{\mathrm{d}\mathbf{x}_p}{\mathrm{d}t} &= \mathbf{V}_p, \end{aligned} \tag{8.25}$$

where $\mathbf{V}_p$ is the particle velocity, τ_p is the particle relaxation time, $\widetilde{\mathbf{u}}$ and $\mathbf{u}^{\mathrm{sgs}}$ are the fluid large-scale and subgrid-scale velocities at the particle locations, respectively, $\mathbf{g}$ is the gravitational acceleration, $\mathbf{x}_p$ is the particle position.

To assess the scope of applicability of the present approach we compared the SGS vorticity magnitude field obtained from the multifractal SGS model and the high-pass filtered homogeneous and isotropic turbulence. We observed notable discrepancies between the two, particularly concerning vorticity magnitudes and spatial correlations. The vorticity magnitude fields differ in detail, however, are expected to exhibit the same multifractal character [4, 32].

However, if we assume that $\mathbf{A}$ is an arbitrary vector field then:

$$\nabla \cdot (\nabla \times \mathbf{A}) \equiv 0, \tag{8.26}$$

meaning that the divergence of a curl is always equal to 0. Due to the stochastic nature of both the multiplicative cascade, Eq. (8.8), and the additive cascade, Eq. (8.10), the SGS vorticity field defined according to Eq. (8.11), is not non-divergent, ergo cannot be reversed through the application of the Biot-Savart integral, Eq. (8.7).

Moreover, in contrast to the present approach, the Biot-Savart integral is typically evaluated over the entire vorticity field allowing to account for the contribution from the surrounding cells, but also, enforcing the retrieved velocity field to be continuous and non-divergent. If the Biot-Savart integral is evaluated over artificially restricted region, as proposed by the Authors, the retrieved velocity field will obviously be discontinuous at cells boundaries. Such discontinuities are furthermore expected to influence particle motions leading to artificial clustering at the resolved-scale cells boundaries. To overcome this issue, the scale-invariant multipliers $P(\mathcal{M})$ should be somehow resampled to obtain continuous and divergence-free vorticity field and, consequently, to enable a proper evaluation of the Biot-Savart integral.

If particulate two-phase flows are concerned, and thus explicit SGS velocity components are required in the particle equation of motion, the multifractal SGS model, developed by Burton & Dahm [4, 20, 28] in the framework of single-phase SGS modeling, introduces non-physical "vorticity" field, and, subsequently, reverses it incorrectly using the Biot-Savart integral. Therefore, it seems that this approach is inappropriate in terms of retrieving the SGS velocity components. Moreover, due to inevitable computational cost associated with the evaluation of the Biot-Savart integral, the multifractal SGS model, does not bode well for future development of an SGS model enhancing particle motions.

Acknowledgments

The work was supported by the National Science Centre (NCN, Poland), research project 2018/30/Q/ST8/00341, and the Interdisciplinary Centre for Mathematical and Computational Modelling (ICM), grant G87-1145.

Bibliography

[1] J.G.M. Kuerten. Point-particle DNS and LES of particle-laden turbulent flow — a state-of-the-art review. *Flow, Turbulence and Combustion*, 97:689–713, 2016.

[2] S. Elghobashi. Direct Numerical Simulation of turbulent flows laden with droplets or bubbles. *Annual Review of Fluid Mechanics*, 51:217–244, 2019.

[3] C. Marchioli. Large-eddy simulation of turbulent dispersed flows: A review of modelling approaches. *Acta Mechanica*, 228:741–771, 2017.

[4] G.C. Burton and W.J.A. Dahm. Multifractal subgrid-scale modeling for large-eddy simulation. I. Model development and a priori testing. *Physics of Fluids*, 17:075111, 2005.

[5] J. Pozorski. Models of turbulent flows and particle dynamics. In J. Pozorski and J-P Minier, editors, *Particles in Wall-Bounded Turbulent Flows: Deposition, Re-suspension and Agglomeration*, pages 97–150. Springer, 2017.

[6] J. Smagorinsky. General circulation experiments with the primitive equations: I. The basic experiment. *Monthly Weather Review*, 91:99–164, 1963.

[7] M. Germano, U. Piomelli, P. Moin, and W.H. Cabot. A dynamic subgrid-scale eddy viscosity model. *Physics of Fluids*, 3:1760–1765, 1991.

[8] J. Fröhlich and W. Rodi. Introduction to large eddy simulation of turbulent flows. In B. E. Launder and N. D. Sandham, editors, *Closure Strategies for Turbulent and Transitional Flows*, pages 267–298. Cambridge University Press, 2002.

[9] P. Moin, K. Squires, W. Cabot, and S. Lee. A dynamic subgrid-scale model for compressible turbulence and scalar transport. *Physics of Fluids A: Fluid Dynamics*, 3:2746–2757, 1991.

[10] D.K. Lilly. A proposed modification of the Germano subgrid-scale closure method. *Physics of Fluids A: Fluid Dynamics*, 4:633–635, 1992.

[11] S. Ghosal, T.S. Lund, P. Moin, and K. Akselvoll. A dynamic localization model for large-eddy simulation of turbulent flows. *Journal of Fluid Mechanics*, 286:229–255, 1995.

[12] R.H. Kraichnan. Eddy viscosity in two and three dimensions. *Journal of Atmospheric Sciences*, 33:1521–1536, 1976.

[13] J.P. Chollet and M. Lesieur. Parameterization of small scales of three-dimensional isotropic turbulence utilizing spectral closures. *Journal of Atmospheric Sciences*, 38:2747–2757, 1981.

[14] A. Vela-MartĂn. Subgrid-scale models of isotropic turbulence need not produce energy backscatter. *Journal of Fluid Mechanics*, 937:A14, 2022.

[15] J. Bardina, J. Ferziger, and W.C. Reynolds. Improved subgrid-scale models for large-eddy simulation. In *13th Fluid and Plasmadynamics Conference*, page 1357, 1980.

[16] J. Bardina. *Improved turbulence models based on large eddy simulation of homogeneous, incompressible, turbulent flows*. PhD thesis, Stanford University, 1983.

[17] J. A. Domaradzki and E.M. Saiki. A subgrid-scale model based on the estimation of unresolved scales of turbulence. *Physics of Fluids*, 9:2148–2164, 1997.

[18] J.A. Domaradzki and K.-C. Loh. The subgrid-scale estimation model in the physical space representation. *Physics of Fluids*, 11:2330–2342, 1999.

[19] J.A. Domaradzki, W. Liu, C. Härtel, and L. Kleiser. Energy transfer in numerically simulated wall-bounded turbulent flows. *Physics of Fluids*, 6:1583–1599, 1994.

[20] G.C. Burton and W.J.A. Dahm. Multifractal subgrid-scale modeling for large-eddy simulation. II. Backscatter limiting and a posteriori evaluation. *Physics of Fluids*, 17:075112, 2005.

[21] A. Scotti and C. Meneveau. A fractal model for large eddy simulation of turbulent flow. *Physica D: Nonlinear Phenomena*, 127:198–232, 1999.

[22] B. Ray and L.R. Collins. A subgrid model for clustering of high-inertia particles in large-eddy simulations of turbulence. *Journal of Turbulence*, 15:366–385, 2014.

[23] M. Voßkuhle, A. Pumir, E. Lévêque, and M. Wilkinson. Collision rate for suspensions at large stokes numbers–comparing navier–stokes and synthetic turbulence. *Journal of Turbulence*, 16:15–25, 2015.

[24] Z. Zhou, S. Wang, and G. Jin. A structural subgrid-scale model for relative dispersion in large-eddy simulation of isotropic turbulent flows by coupling kinematic simulation with approximate deconvolution method. *Physics of Fluids*, 30:105110, 2018.

[25] Z. Zhou, S. Wang, X. Yang, and G. Jin. A structural subgrid-scale model for the collision-related statistics of inertial particles in large-eddy simulations of isotropic turbulent flows. *Physics of Fluids*, 32:095103, 2020.

[26] M. Rajek and J. Pozorski. Kinematic simulations as a subgrid-scale model for particle motion: a priori LES of homogeneous isotropic turbulence. *Journal of Physics: Conference Series*, accepted 2022.

[27] M. Bassenne, M. Esmaily, D. Livescu, P. Moin, and J. Urzay. A dynamic spectrally enriched subgrid-scale model for preferential concentration in particle-laden turbulence. *International Journal of Multiphase Flow*, 116:270–280, 2019.

[28] G.C. Burton. *A multifractal subgrid-scale model for the large-eddy simulation of turbulent flows.* PhD thesis, University of Michigan, 2003.

[29] J.A. Mullin and W.J.A. Dahm. Dual-plane stereo particle image velocimetry measurements of velocity gradient tensor fields in turbulent shear flow. II. Experimental results. *Physics of Fluids*, 18:035102, 2006.

[30] C. Meneveau and K.R. Sreenivasan. Simple multifractal cascade model for fully developed turbulence. *Physical Review Letters*, 59:1424, 1987.

[31] C. Meneveau and K.R. Sreenivasan. The multifractal nature of turbulent energy dissipation. *Journal of Fluid Mechanics*, 224:429–484, 1991.

[32] R.D. Frederiksen, W.J.A. Dahm, and D.R. Dowling. Experimental assessment of fractal scale similarity in turbulent flows. Part 3. Multifractal scaling. *Journal of Fluid Mechanics*, 338:127–155, 1997.

[33] K.R. Sreenivasan. Fractals and multifractals in fluid turbulence. *Annual Review of Fluid Mechanics*, 23:539–604, 1991.

[34] A.B. Chhabra and K.R. Sreenivasan. Scale-invariant multiplier distributions in turbulence. *Physical Review Letters*, 68:2762, 1992.

[35] S.A. Orszag and Patterson G.S. Numerical simulation of three-dimensional homogeneous isotropic turbulence. *Physical Review Letters*, 28:76–79, 1972.

[36] V. Eswaran and S.B. Pope. An examination of forcing in direct numerical simulations of turbulence. *Computers & Fluids*, 16:257–278, 1988.

[37] H. Fayed and S.A. Ragab. Large-eddy simulation of particles and bubbles collisions in homogeneous isotropic turbulence. In *43rd AIAA Fluid Dynamics Conference*, page 3183, 2013.

Chapter 9

Reference free damage detection by guided waves with using FBG sensors

Sara Sarbaz

Institute of Fluid Flow Machinery, Polish Academy of Sciences,
Fiszera 14, 80-231 Gdansk, Poland

9.1 Introduction

In this chapter fundamental of guided wave-based structural health monitoring is reviewed. After that principles and advantages and disadvantages of transducers used for the SHM system are described. Additionally, there is a short study of reference-free damage detection techniques. In the last section, an experiment has been conducted to investigate one of the reference-free damage detection method using "the theory of instantaneous baseline" approach.

9.2 Guided wave-based structural health monitoring

The process of damage detection in aerospace, civil, and engineering science infrastructure is part of structural health monitoring (SHM). Here, any changes to the dimension and geometry or material properties including changes to the boundary conditions and system connectivity, which adversely affect the system's performance are considered as damage. Within the last decades, a rapid increase in the amount of research associated with SHM has been seen. The increased interest in SHM and its associated potential for significant life-safety and economic benefits has motivated the requirement for this theme [1].

SHM systems provide a useful tool for damage detection in engineered structures which ends up in the reduction of the upkeep costs and improvement of reliability of the structures. Several SHM techniques like vibration-based [2], strain-based [3], and electromechanical impedance-based techniques [4] are developed. Another candidate is a guided wave-based SHM technique, which has attracted substantial attention within the past 20 years. Guided wave-based schemes offer some advantages such as application of cheap and light-weight transducers and it can easily incorporate within the structure, beside that due to covering a high range of frequency excitation can detect minute damages. Unlike low-frequency vibration technique, few numbers of transducers can perform measurements

a wide range of area. The trade-offs are essentially the complexities related to guided wave propagation and their interaction with damage. Firstly, the guided waves are often multi-modal and hence one wave input converts into different wave modes make the response analysis more complicated to. Secondly, most of the guided waves are dispersive, i.e., the form of the wave changes when it propagates. In addition to this, the waves are reflected by the discontinuities and boundaries, resulting in complex response, which is strongly influenced by environmental factors, instrumentation, and ambient noises, and altered in operational conditions. The above phenomena often hide the damage signature, which might be sorted out through proper selection of transducers, a good understanding of the wave characteristics, and signal processing schemes [5].

9.3 Fundamental of guided-wave based SHM

Primarily, ultrasonic elastic wave propagating through bounded structural media of guided wave is used in the context of SHM. Lamb wave is one of the most common sorts of guided waves that have been studied for the application to SHM. Lamb waves propagate in thin plates and bounded shells. For thicker plates Rayleigh propagates near the free surface of the media and at a higher frequency. Another, similar wave in plate-like structure is the shear horizontal (SH) wave, commonly studied for laminated composite plates [6].

Analysis of propagation is based on the governing differential elastic-dynamic equations and associated boundary conditions. To obtain the dispersion features these equations are transformed to the frequency-wavenumber domain. Propagating modes and wave velocity as a function of frequency are relations between group/phase velocity and frequency which define wave propagation behaviour. In addition to dispersion characteristics, the governing equations are also solved in the time domain to evaluate the wave amplitudes, reflections from edges and boundaries, and interaction with damage [7].

9.3.1 Modeling and simulation of guided waves

Numerical, analytical, and semi-analytical techniques for modelling methodologies of guided waves have been introduced to date. Although some reviews of computational mechanics have been conducted [8–11], FEM (finite element method) remained the most common modelling technique to validate and study guided wave experiments to evaluate the efficacy of different post-processing techniques. FEM has been also used as a great solution for the simulation of wave propagation in laminated structures to overcome the problems caused by anisotropy and layered structure [5]. Recently 3-D modelling of composite is more in the center of attention for researchers in this field. A comparative study on the suitability of several higher-order FE techniques including the spectral element method (SEM) for simulation of Lamb waves has been discussed [6].

Simulation and visualization of Lamb waves based on the finite differences method (FDM) were investigated in [11]. This method was later adopted by [12, 13] for damage detection in the anisotropic plate. Numerical approaches include distinctive methods such as the FDM method investigated by [14] and wave finite elements (WFE) reviewed in [15]. In a parallel manner work on analytical methods for the study of guided wave motion progressed and many papers have also introduced analytical modelling of scattering and interaction of the guided wave with damages [16]. To anticipate the A0 Lamb wave scattering at delamination, in reference [17] an approach based on Mindlin plate theory and Born approximation is presented. Damage localization in composite laminates based on an analytical technique that comes from the relation between propagation directions and anisotropic was presented by Li *et al.* [18].

9.4 Damage detection and localization by guided-wave based SHM

Several approaches for SHM have been investigated by researchers. Farrar [1] and C.Boller [19] described the integration of SHM into the aircraft design process and the way to validate the established

loads monitoring process and compared it to an emerging damage monitoring solution. Creating a network of sensors integrated into the structure with the spacing between sensing points similar to or smaller than the scale of the anticipated damage. The sensing requirements for an individual sensor can be accomplished with conventional NDT techniques. This approach is a biologically inspired nervous system and relies on a complete sensing system integration into the baseline structure. Making such a technique more practicable, crucial developments in sensor technology are demanded. Some progress has been indicated in [20] showing how several techniques can be applied together and integrated into a structure.

The basic idea of damage localization in the Lamb waves-based technique is that damage can dissipate the traveling wave in a damaged plate [21]. Processing the scattered signal contains some information about the damage such as the (location, size, orientation, type). Time-of-flight (ToF), is an important element for damage localization which is extracted from the scattered signal. ToF is calculated by the travel time of a wave packet from an actuator to a sensor through a given path [22].

Many of these techniques implement the baseline signal subtraction approach for damage diagnosis. In this approach, the signal obtained from a defect-free structure called the baseline signal is compared with the actual signal recorded during routine inspections. Any considerable differences between these signals can be recognized as critical damage. However, ambient temperature variations, unavoidable unpredictability in the PZT installation, and breaking down of mechanical properties can affect the accuracy of this common approach [23].

9.5 Actuators and sensors used in SHM

As mentioned previously, actuators and sensors are important parts of an SHM system. Choosing an appropriate transducer is affected by many factors such as compatibility with the geometry of the structures, sensitivity to mechanical responses, the efficiency of stress or energy transfer, lower level of interference of transducer

dynamics with the wave propagation, frequency of operation, environmental conditions, etc. [24].

The following subsections describe, in brief, these transducers and their implementation for guided wave generation and sensing.

9.5.1 Piezoelectric transducer (PZT)

Piezoelectric transducers have been most commonly used among all the other kinds of transducers for actuation and sensing of guided waves. In the beginning, PZT wedges and combs were being used for ultrasonic NDT as both actuators and sensors of guided waves [25, 26]. Due to a lack of integrability to the structure, PZT has been replaced by piezoelectric wafer transducers (PWT) for guided wave actuation and sensing [27–29]. Piezoelectric wafer active sensors (PWAS) provide SHM systems with some benefits such as being easily embedded in the structures, lower cost, lighter weight, and capability to manufacture with low thickness. Apart from what has been mentioned as advantages of these transducers, PWT encounters some challenges with incorporation into structures with curved and unusual geometry. Some attempts have been performed by researchers to overcome this limitation such as using piezo-composite actuators to achieve more flexibility [30].

9.5.2 Doppler vibrometer

The laser Doppler velocimeter/vibrometer (LDV) is a noncontact optical sensing tool for accurately measuring vibration velocity in the point of structure. The noncontact nature of the instrument makes it particularly attractive for use on lightweight structures where measurement interaction must be minimized. Real-time scanning LDV's have recently been introduced to measure fluid flow velocity profiles rapidly.

9.5.3 Fiber Bragg grating sensors (FBG)

Recently FBGs have gained researchers' attention because of their attractive properties and capabilities. The FBG is a narrow band wavelength filter written in an optical fiber, which reflects a certain

narrow bandwidth of light out of the total input light to the fiber, transmitting the other wavelength. The reflection is due to the unique structure of the FBG in which the refractive index of the core is periodically perturbed along the grating length [31].

There are several advantages of FBGs which can be taken into account. First of all, they enhance ultrasonic detection with only a single lead-in lead-out connection which provides multiple sensing locations. FBG sensor is made up of glass which is environmentally more stable and has well above the lifetime of the FRP structures. Additionally, because of the diminutive size can be embedded in FRP composites during the manufacturing of the composite part with no effect on the strength of the part. This sensor is appropriate for organizing since it encompasses a narrowband with a wavelength working extend and thus can be exceedingly multiplexed. This non-conductive sensor can operate in electromagnetically noisy environments due to immunity to electromagnetic fields. It is also a suitable choice for remote sensing because of its low transmission loss, which provides monitoring of the signals from long distances [32, 33]. Furthermore, the FBG sensor provides the measurement of multiple parameters such as load/strain, vibration, and temperature [34].

The problem with the application of GW in composite structures is the high attenuation of the waves and complex propagation pattern due to the material anisotropy. The attenuation of the elastic waves can be countered by the use of more sensors which is not ideal due to the additional wiring requirements and the increase in weight. The Fiber-optic (FO) sensors can be multiplexed (many sensors on a single fiber) and hence are a good solution to reduce the weight issue and overcome the wiring issues. The FO sensors-based fiber Bragg grating (FBG) sensors are extremely popular but their use for GW detection has been limited due to their low sensitivity [35]. Although the recent advances in the use of the edge filtering approach [36] and the remote bonding of the FBG sensors [37], it is possible to obtain good sensitivity, thus paving the way for their use in structures.

Despite all the advantages of FBG sensors, they are passive sensors and have directional sensitivity. In order to apply them to the

existing techniques of baseline-free damage detection there might be a few challenges. One of the promising techniques to improve the sensitivity of FBG is edge-filtering, in which on the linear region of the reflectivity spectrum of the FBG a narrow band laser source is set with its output wavelength. As it is shown in Fig. 9.1 [38], any minor shift in the λ ends up in a large change in the output intensity, which is detected by a photodetector. The schematic setup for this technique is illustrated in Fig. 9.1 [38].

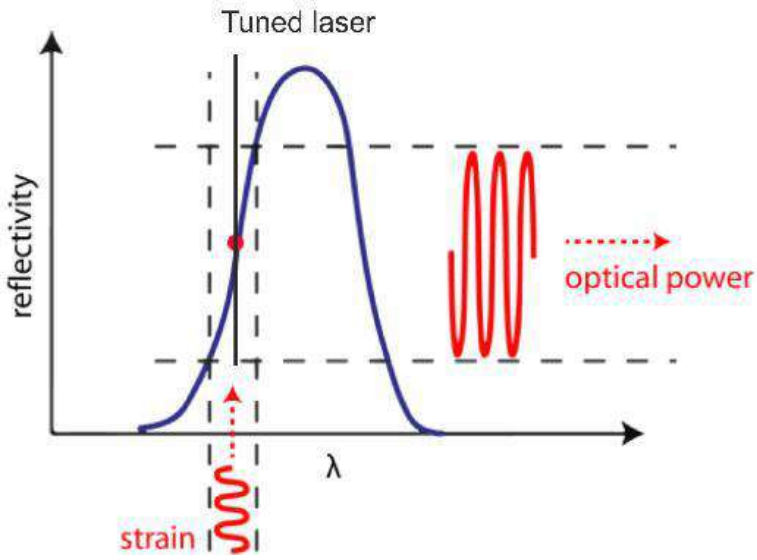


Fig. 9.1: Ilustration of edge filtering approach [38]. Used by author's permission.

The primary benefit of FBG sensors being used in edge-filtering domain is being self-referencing and independent of fluctuating light levels as the perturbation information is encoded in wavelength. In contrast to other types of optical fiber sensors, this system has immunity to source power and connector losses. There are further advantages of FBGs can be taken into account such as linear responses over distinctive orders of magnitude. Therefore FBGs are considered very important components for monitoring composite material and the development of smart structure technology [39].

One of the major disadvantages of using an FBG arises from the fact that the extraction of relevant features from the measured signals requires a high signal-to-noise ratio (SNR) in the captured signal. This is because the measured signal has such a minimal amplitude and we need to discern any small change in that signal from the background noise. Several factors reduce the SNR of the FBG. First of all, the cylindrical optical fiber geometry reduces the signal transfer to the optical fiber, because the contact area between the optical fiber and structure is significantly smaller as compared to disc/patch type sensors such as PZT disks/wafers.

The physics of the fiber-wave coupling and the use of this phenomenon for damage detection has been the focus of several studies [40–42].

9.6 Reference free damage detection

Looking through the current reference-free damage detection approaches, almost all of these techniques compare the measured response of the structure with measurement taken when the structure is known to be healthy (baseline). Although these methods seemed to be practical in laboratory tests, their application to real structures in in-service conditions has been very limited due to some reasons. Above all, baseline condition measured is not always accessible, besides that baseline measurements are often corrupted by noise or affected by changes in ambient conditions. Changes in environmental and operational conditions such as fluctuations in temperature, variation in surface moisture, and varying loading conditions can all cause the response of a structure to change significantly from the baseline measurement. The presence of the various ambient condition in the use of a standard baseline comparison method may lead to false damage indication. Additionally, standard baseline comparison methods are unable to detect damage that exists in a structure before the installation of the health monitoring system, therefore, they are only useful for detecting future damage [43]. As a result, a method that can be applied to a structure without the baseline measurements is extremely demanded.

9.6.1 Non-linear ultrasonic based technique

Some of the reference-free GW-based damage detection techniques make use of non-linear ultrasonic. Non-linear analysis of guided waves enables the detection of small changes in the microstructure of the material that do not affect the linear elastic constants or result in detectable scattering [44]. The basis and methodology for utilizing nonlinear ultrasonic guided waves for early damage detection are discussed by [44]. The basic concept of this technique is a distortion of the guided waves coming about from the microstructural changes that cause higher harmonics, which are then representative of the early stages of damage. Lim *et al.* [45] studied a reference-free fatigue crack detection technique using nonlinear ultrasonic modulation. A nonlinear modulation technique can be employed when low frequency (LF) and high frequency (HF) inputs are generated by two surface-mounted PZT transducers that are applied to a structure, the presence of a fatigue crack can provide a nonlinear mechanism, and create spectral sidebands around the frequency of the HF signal. The crack-induced spectral sidebands are isolated using a combination of linear response subtraction (LRS), synchronous demodulation (SD), and continuous wavelet transform (CWT) filtering [45].

The same as a bulk wave, changes in an elastic medium can be recognized by any nonlinearity in an ultrasonic Lamb wave however, its nonlinearity is affected by angle, partial wave vectors, and higher-order elastic constants [46]. Parameter β' is widely-accepted as relative nonlinear which can be defined by:

$$\beta' = \frac{A_2}{{A_1}^2} \tag{9.1}$$

where A_1 and A_2 are the amplitudes of the fundamental wave and the second harmonic wave, respectively.

In nonlinear ultrasonic technique, to extract the amplitude information at the exact fundamental and harmonic frequencies while reducing the influence of the reflected and scattered waves, a data process based on the job done in [47] and is shown in Fig. 9.2.

First, a complex wavelet transform plots the signal in the time-frequency domain. The input signal distributes its main energy and exhibits no harmonic components. The signal detected by the PSFBG balanced sensor exhibits its energy concentration, which agrees well with the middle frequency of the input signal. The signal also shows a clear harmonic component with a low amplitude. Then, the amplitude envelopes at both fundamental and harmonic frequencies will be extracted.

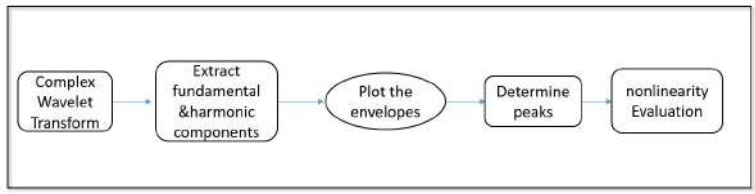


Fig. 9.2: Flow chart of the data processing method [47].

These methods are not sensitive to linear defects, also the damage-sensitive signal is often overlapped by background noise.

9.6.2 Reference-free damage detection using the time-reversal method

As it was mentioned before, one of the main problems of working with Lamb waves is effect of dispersion. Time reversal (TR) has been introduced by some researchers as a technique to overcome this issue [48, 49].

In the time-reversal method, an input signal can be reconstructed at an excitation point (point A) if an output signal recorded at another point (point B) is remitted to the source point (point A) after being reversed in a time domain as illustrated in Fig. 9.3. This time reversibility (TR) of waves is based on the spatial reciprocity and time-reversal in variance of linear wave equations [49, 50].

The idea of applying the time-reversal concept to Lamb waves has been introduced before by [50, 52]. A specific narrowband input waveform and a wavelet-based signal filtering technique are employed to enhance the TR of Lamb waves to rebuild the input signal.

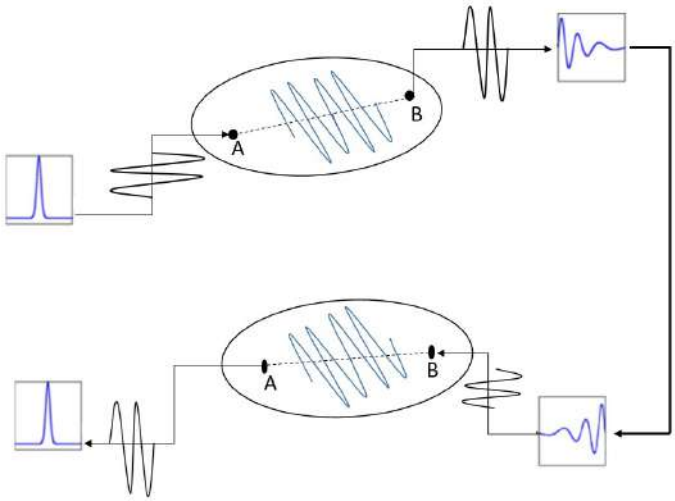


Fig. 9.3: Concept of time-reversal principle [51].

The time-reversal process depends on frequency therefore, different frequency components of the broad-band excitation are scaled distinctively over the time-reversal process which leads to making the complete reconstruction of the input signal for Lamb wave propagations. As a result, using a narrowband excitation could be practical to improve the TR process. Hyun Woo Park [51] enforced TR for Lamb waves, to detect certain types of defects by examining the deviation of the reconstructed signal from the original input signal without any past baseline data.

To localize and estimate the size of damages in plate-like structures pulse-echo mode has been used by [49, 50]. However, this method requires more complicated approaches for multiple damage detection as the pulse-echo process is likely to detect only the most distinct defect. Therefore investigating pulse-echo mode in plate-like structures by using a dense array of sensors to cover the entire boundary of the specimen seems impractical [51].

As fundamental of TR is based on linear responses in case of existence of any source of nonlinearity its linearity will be disturbed.

Therefore, comparison between the original input signal and rebuilt signal leads to damage detection such as recognition of crack or fiber breakage. The majority of regular damage detection techniques use the measurements of the healthy condition of the structure as the baseline data which is highly under the effect of environmental conditions. As a result, normalization is required to be conducted to distinguish signal changes made by defects from those disorders caused by natural variations of the system [51].

An enhanced time-reversal method using wavelet signal processing

To minimize the dependency of the time-reversal operator on the frequency and maximize the signal-to-noise ratio, a designed narrowband input waveform propagates on a structure. Application of narrowband input signal prevents the time-reversal operator from having large variations around the driving frequency also, signal processing will be more convenient and repeatable [53] attempted to eliminate noise in the ultrasonic signal by the same approach.

Modified time-reversal method (MTRM)

The capability of the time-reversal concept can be improved by a modified time reversal process. Regularly in structural health monitoring systems transducers are located in a circular or rectangular pattern. Interestingly in MTRM only one transducer is usually placed in a central location and works as an actuator and the others could function as sensors. One of the main advantages of this method is reducing the hardware such as signal generator wiring, etc. Moreover, it enables sensing using non-contact methods like vibrometer for ground-based SHM. The principle of the MTRM method contains several stages beginning with sending A Lamb wave tone burst signal from transducer A and recording at transducer B, which continues by reversing the received signal at B in time. After that time-reversed signal will be resent to transducer B. At last, by comparing the captured signals to the actuated ones where both signals are normalized damage detection can occur [54].

9.6.3 Instantaneous baseline approaches (IBM)

Instantaneous baseline structural health monitoring is based on detecting damage without the use of baseline data by obtaining an "instantaneous baseline" measurement of structure investigation. For plate-like structure and shells with free surfaces which is propagated by Lamb waves, dispersive waves travel in two distinct modes, known as symmetric and ant symmetric modes. A comprehensive analysis of Lamb waves is given by Viktorov [55] and Rose [56] . One possible approach could be placing transducers on a structure and then pitch-catch Lamb wave propagation is used to obtain common features of undamaged sensor-actuator paths, which is considered as a baseline. Transducers must be located at equal lengths so that structural features and material properties are uniforms among transducers. In an undamaged anisotropic structure, the initial Lamb wave arrival signals for different paths before any boundary reflections will remain unchanged. In presence of damage along with one of the paths, the Lamb wave signal obtained from that path is different from the rest signals. The reference-free separation of the undamaged path from other signals occurs by creating a statistical baseline using undamaged features. An instantaneous baseline is made by signals from undamaged paths and damage will be indicated by comparison of signal differences. In structures with more complicated geometry features, sensor placement becomes a significant issue to maintain the structural features symmetric between the paths. Principles of nondestructive evaluation techniques such as ultrasonic and thermography testing, which assume that data can be collected over a large area of a structure and that most of those data will be recorded over undamaged sections, can be applied for the instantaneous baseline method. In the instantaneous baseline method, the majority of sensor-actuator paths are considered to be recorded as healthy paths, in this case, the damaged path can be identified by comparing that with the baseline [57].

The instantaneous baseline method relies on the fact that Lamb wave signals recorded for two equally spaced paths will be identical if no damage is present in the vicinity of the paths. Assuming

the case when for healthy strucuture, both material properties and geometric parameters are the same. Besides that, the electromechanical properties of the transducers caught by Lamb waves must remain without any change. Any minor deviation in bonding condition and electrical or mechanical properties of sensors result in differences in Lamb wave signals from undamaged paths. The piezoelectric active-sensor technique is used to enhance the structural health monitoring system by providing the constant monitoring condition of the piezoelectric sensors and actuators installed on the structure. This technique is capable of evaluating both the bonding condition between a piezoelectric transducer and its host structure, as well as the mechanical and electrical properties of the device. The success of the instantaneous baseline method depends on the conditions of transducers as the method is based on the fact that the shape, amplitude, and frequency of Lamb waves recorded for undamaged paths are the same and that an instantaneous baseline can be created from undamaged measurements. Therefore, the ability to diagnose the condition of PZT transducers is critical to developing this method. The sensor diagnostics technique, explained by Park *et al.* [58] , can be used to detect transducers error at various frequencies and compare plots of the imaginary part of the admittance versus the frequency of surface-bonded piezoelectric [58].

The instantaneous baseline method (IBM) is not able to process pitch-catch pairs of different lengths so Jianxi Qiu proposed a technique based on distance compensation for pitch-catch pairs of different lengths [59]. Moreover, distance compensation can be employed for the inspection of non-symmetrical structures, which is a significant improvement over the traditional IBM method. Using the propagation characteristics of a certain Lamb wave mode, signals received by a transducer pair of shorter length can be predicted to a longer wave traveling distance. Thus, a baseline in a new environment can be extracted if the data are fused to reveal the damaged paths [59].

For the methods based on the instantaneous baseline, a distance compensation algorithm for pitch-catch pairs of different lengths is

proposed by Qiu *et al.* [60]. Salmanpour *et al.* [61] have proposed a method that maps a baseline area onto the interrogation area and overcomes the need for a baseline. Despite these methods proposed there are several open questions regarding the sensitivity of the method for different types of damage, and the effect of ambient conditions on these techniques which have limited their use in real applications, and hence the search for a definitive baseline-free method is still ongoing.

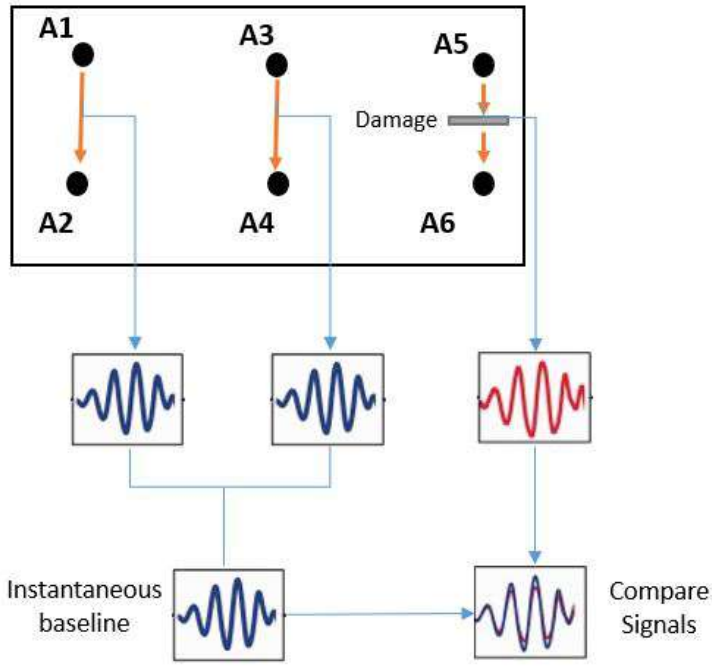


Fig. 9.4: Methodology for instantaneous baseline [59].

However, this method also has some shortages in practical applications. Firstly, before the distance compensation, the material properties in the calculation of relatively precise dispersive curves should be obtained. Thus, there is one more step in the inspection for the measurement of the material. Secondly, the compensation precision is affected by the wave propagation distance. Therefore, the

inspection size is limited. The transducers should be arranged at the locations where damage easily occurs.

To use the instantaneous baseline method for damage detection, the existing techniques such as distance compensation and cross-correlation need to be customized to take into account the directionality. The directionality will be compensated analytically. Fig. 9.4 shows the methodology concisely. In addition to the directionality, the passive nature will result in fewer signals for comparison.

For the advanced signal processing approach, the instantaneous baseline may be used, where the measurement from one actuator-sensor (AS) pair measured on the fly is used as a baseline, and the correlation of other AS pairs is calculated. The path affected with damage will have a lower agreement with the assumed baseline. This method overcomes the need for healthy baseline measurement ahead of time [59].

9.7 A sample experiment for reference-free damage detection using FBG sensor

To investigate the theory of instantaneous baseline, an experiment has been carried out to investigate one of the reference-free detection using "the theory of instantaneous baseline" approach. This study focuses on reference-free damage detection using GW-based FBG sensors.

9.7.1 Motivation

The baseline free approaches detect some features in the signal which are present due to the presence of the damage. This approach will allow us to overcome the need for a pre-measured baseline. This allows wider applicability of the method to existing structures that are suspected to have deteriorated. Also, the instantaneous baseline techniques are inherently more robust to ambient condition changes which allow the application to a structure in different ambient conditions.

9.7.2 Experimental Setup

A homogeneous ASL plate with dimension of (50×50×0.1) cm is used for this research. A tunable laser (Apex AP1000) coupled with FBG (Femto Fibertech) was used. GW signals were generated using PZT and sensed based on FBG and edge-filtering technique. To help sensing the GW signals a pre-amplified photo receiver was located along with an oscilloscope (National Instruments, PXI 5105) via optical coupling and with a higher sampling frequency. The wave excitation signals were amplified via an amplifier (Krohn Hite7500) and then supplied to PZTs (NCE51-CTS Corporation). The schematic represented in Fig. 9.5, illustrates ALS with the arrangement of PZTs. Six symmetric pitch-catch Pairs of PZTs (labeled as "A_1-A_7","A_2-A_8" and respectively) of identical length are adjusted on the surface of the plate. Two damage scenarios were simulated by adding magnets of ϕ 1.8 cm at various Locations D_1 (X=20 cm, Y=33 cm), D_2 (X=30 cm, Y=30 cm).

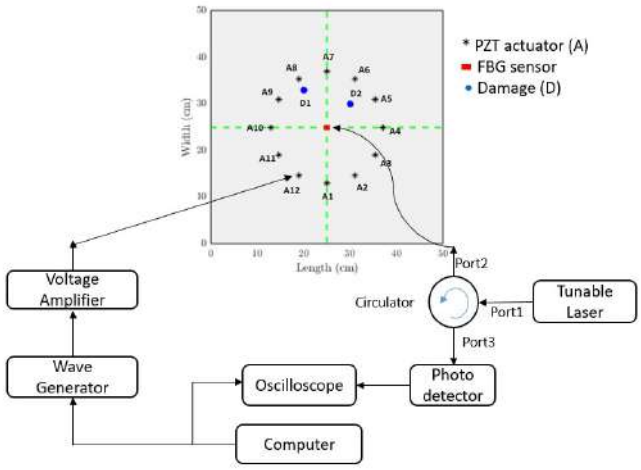


Fig. 9.5: Schematic of the experimental setup.

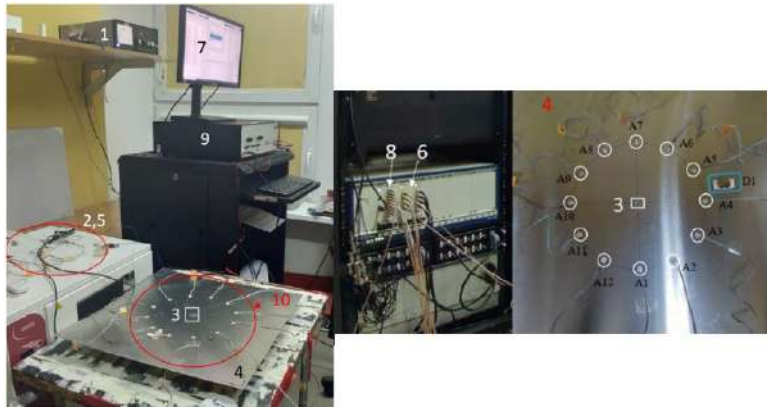


Fig. 9.6: Experimental setup:(1) Laser, (2),(5) Circulator, Photo detector units,(3) FBG,(4) Plate, (6) Oscilloscope, (7) GW monitoring, (8) Signal generator, (9) Amplifier, (10) PZT network [62].

9.7.3 Methodology

Various damage indices were previously proposed by the researchers for damage identification. Methods such as root mean square (RMS), signal differences (SD), and root mean square deviation (RMSD) helps in better signal paths separation. But, due to the directionality of the FBG sensor, it is difficult to apply such techniques. In this study, damage identification will be investigated based on the cosine distance (CD) method or which is independent of amplitude or scaling [62].

9.7.4 Results

The signals obtained from the actuator-sensor network are analysed by plotting the signals obtained for all pairs of actuators. As is shown in Fig. 9.7 for the first damage scenario the differences between the signals obtained from A_2 and A_8 are considerably higher in comparison to the rest pairs for example A_1-A_7. By calculation of these differences, either using cosine distance or cross-correlation damage can be detected independently from data of healthy signal. Preliminary measurements have validated proper

operation of measurement setup and proposed methodology.

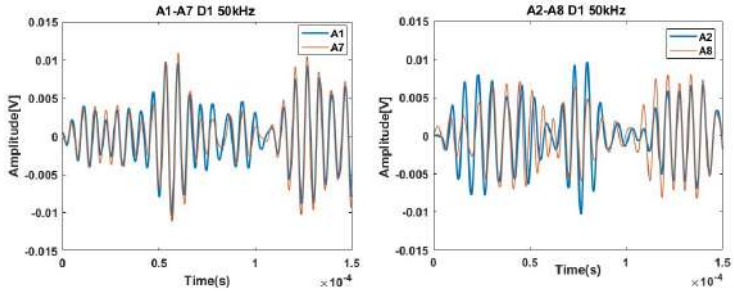


Fig. 9.7: Signal difference visualization for A_2-A_8 and A_1-A_7.

9.8 Conclusion

In this chapter the basic of guided-wave based structural health monitoring and its fundamental were discussed.As an important element of a SHM system various actuators and sensors used for sensing were explained. Studying these approaches shows that most of damage detection techniques are based on data from the healthy structure. In other words they are not reference- free.To improve a technique which is independent from healthy structure data, a short review of reference-free techniques was conducted. In the last section an experiment has been carried to find a new reference-free approach was described. Comparison of signals obtained from the experiment, revealed that using instantaneous base-line method with setting pairs of actuators in a circular pattern with an FBG sensor located in the center could be a possible reference free technique.

Acknowledgments

The author acknowledge the support provided by Sonata 16 project titled: Guided waves based reference-free structural health monitoring using fiber Bragg grating sensors granted by National Science Center, Poland, grant no. (2020/39/D/ST8/00188) .

Bibliography

[1] Charles Farrar and Keith Worden. An introduction to structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851):303–315, 2007.

[2] Swagato Das, P Saha, and SK Patro. Vibration-based damage detection techniques used for health monitoring of structures: a review. *Journal of Civil Structural Health Monitoring*, 6(3):477–507, 2016.

[3] YY Li. Hypersensitivity of strain-based indicators for structural damage identification: A review. *Mechanical Systems and Signal Processing*, 24(3):653–664, 2010.

[4] Shishir Kumar Singh, Rohan Soman, Tomasz Wandowski, and Pawel Malinowski. A variable data fusion approach for electromechanical impedance-based damage detection. *Sensors*, 20(15):4204, 2020.

[5] Chunhui Yang, Lin Ye, Zhongqing Su, and Michael Bannister. Some aspects of numerical simulation for lamb wave propagation in composite laminates. *Composite structures*, 75(1-4):267–275, 2006.

[6] Christian Willberg, Sascha Duczek, JM Vivar Perez, David Schmicker, and Ulrich Gabbert. Comparison of different higher order finite element schemes for the simulation of lamb waves. *Computer methods in applied mechanics and engineering*, 241:246–261, 2012.

[7] Zhongqing Su, Lin Ye, and Ye Lu. Guided lamb waves for identification of damage in composite structures: A review. *Journal of sound and vibration*, 295(3-5):753–780, 2006.

[8] Ajay Raghavan. *Guided-wave structural health monitoring.* PhD thesis, 2007.

[9] Zhongqing Su, Lin Ye, and Ye Lu. Guided lamb waves for identification of damage in composite structures: A review. *Journal of sound and vibration*, 295(3-5):753–780, 2006.

[10] Christian Willberg, Sascha Duczek, Juan Miguel Vivar-Perez, and ZA Bin Ahmad. Simulation methods for guided wave-based structural health monitoring: a review. *Applied Mechanics Reviews*, 67(1), 2015.

[11] BC Lee and WJ Staszewski. Modelling of lamb waves for damage detection in metallic structures: Part i. wave propagation. *Smart materials and structures*, 12(5):804, 2003.

[12] Shankar Sundararaman and Douglas E Adams. Modeling guided waves for damage identification in isotropic and orthotropic plates using a local interaction simulation approach. *Journal of vibration and acoustics*, 130(4), 2008.

[13] Matthew Obenchain and Carlos Cesnik. Guided wave interaction with hole damage using the local interaction simulation approach. *Smart materials and structures*, 23(12):125010, 2014.

[14] R Balasubramanyam, D Quinney, RE Challis, and CPD Todd. A finite-difference simulation of ultrasonic lamb waves in metal sheets with experimental verification. *Journal of Physics D: Applied Physics*, 29(1):147, 1996.

[15] MN Ichchou, J-M Mencik, and W Zhou. Wave finite elements for low and mid-frequency description of coupled structures with damage. *Computer methods in applied mechanics and engineering*, 198(15-16):1311–1326, 2009.

[16] MN Ichchou, J-M Mencik, and W Zhou. Wave finite elements for low and mid-frequency description of coupled structures with damage. *Computer methods in applied mechanics and engineering*, 198(15-16):1311–1326, 2009.

[17] Ching-Tai Ng, Martin Veidt, LRF Rose, and CH Wang. Analytical and finite element prediction of lamb wave scattering at delaminations in quasi-isotropic composite laminates. *Journal of Sound and Vibration*, 331(22):4870–4883, 2012.

[18] Bing Li, Yongquan Liu, Kezhuang Gong, and Zheng Li. Damage localization in composite laminates based on a quantitative expression of anisotropic wavefront. *Smart Materials and Structures*, 22(6):065005, 2013.

[19] Christian Boller. Next generation structural health monitoring and its integration into aircraft design. *International Journal of Systems Science*, 31(11):1333–1349, 2000.

[20] Xinlin Qing, Amrita Kumar, Chang Zhang, Ignacio F Gonzalez, Guangping Guo, and Fu-Kuo Chang. A hybrid piezoelectric/fiber optic diagnostic system for structural health monitoring. *Smart Materials and Structures*, 14(3):S98, 2005.

[21] A Ródenas, GA Torchia, G Lifante, E Cantelar, J Lamela, F Jaque, L Roso, and D Jaque. Refractive index change mechanisms in femtosecond laser written ceramic nd: Yag waveguides: micro-spectroscopy experiments and beam propagation calculations. *Applied Physics B*, 95(1):85–96, 2009.

[22] Claude Fendzi, Nazih Mechbal, Marc Rebillat, Mikhail Guskov, and G Coffignal. A general bayesian framework for ellipse-based and hyperbola-based damage localization in anisotropic composite plates. *Journal of Intelligent Material Systems and Structures*, 27(3):350–374, 2016.

[23] Mira Mitra and S Gopalakrishnan. Guided wave based structural health monitoring: A review. *Smart Materials and Structures*, 25(5):053001, 2016.

[24] Daniel Balageas. Non-destructive evaluation of cooperative structures (ndecs). *Encyclopedia of Structural Health Monitoring*, 2009.

[25] W Zhu, JL Rose, JN Barshinger, and VS Agarwala. Ultrasonic guided wave ndt for hidden corrosion detection. *Journal of Research in Nondestructive Evaluation*, 10(4):205–225, 1998.

[26] MJ Santos, JB Santos, AM Amaro, and MA Neto. Low velocity impact damage evaluation in fiber glass composite plates using pzt sensors. *Composites Part B: Engineering*, 55:269–276, 2013.

[27] Victor Giurgiutiu. Tuned lamb wave excitation and detection with piezoelectric wafer active sensors for structural health monitoring. *Journal of intelligent material systems and structures*, 16(4):291–305, 2005.

[28] Victor Giurgiutiu. *Structural health monitoring: with piezoelectric wafer active sensors*. Elsevier, 2007.

[29] Lingyu Yu and Victor Giurgiutiu. Multi-mode damage detection methods with piezoelectric wafer active sensors. *Journal of intelligent material systems and structures*, 20(11):1329–1341, 2009.

[30] RA Badcock and EA Birt. The use of 0-3 piezocomposite embedded lamb wave sensors for detection of damage in advanced fibre composites. *Smart Materials and Structures*, 9(3):291, 2000.

[31] Minh Phong LUONG. Encyclopedia of structural health monitoring. 2009.

[32] Kenneth Hill and Gerald Meltz. Fiber bragg grating technology fundamentals and overview. *Journal of lightwave technology*, 15(8):1263–1276, 1997.

[33] Alan Kersey, Michael Davis, Heather Patrick, Michel LeBlanc, Koo, Askins, Putnam, and Joseph Friebele. Fiber grat-

ing sensors. *Journal of lightwave technology*, 15(8):1442–1463, 1997.

[34] Isabelle Riant, Salim Gurib, Josselyne Gourhant, Pierre Sansonetti, Christian Bungarzeanu, and Raman Kashyap. Chirped fiber bragg gratings for wdm chromatic dispersion compensation in multispan 10-gb/s transmission. *IEEE Journal of selected topics in Quantum Electronics*, 5(5):1312–1324, 1999.

[35] Chris Baldwin, John Niemczuk, Jason Kiddy, Peter Chen, Martin Christiansen, and Shiping Chen. Structural testing of navy vessels using bragg gratings and a prototype digital spatial wavelength domain multiplexing (dswdm) system. *Naval engineers journal*, 114(1):63–70, 2002.

[36] Daniel Betz, Graham Thursby, Brian Culshaw, and Wieslaw J Staszewski. Acousto-ultrasonic sensing using fiber bragg gratings. *Smart Materials and Structures*, 12(1):122, 2003.

[37] Daniel Betz, Graham Thursby, Brian Culshaw, and Wieslaw J Staszewski. Acousto-ultrasonic sensing using fiber bragg gratings. *Smart Materials and Structures*, 12(1):122, 2003.

[38] Rohan Soman, Junghyun Wee, and Kara Peters. Optical fiber sensors for ultrasonic structural health monitoring: a review. *Sensors*, 21(21):7345, 2021.

[39] Damien Kinet, Patrice Mégret, Keith Goossen, Liang Qiu, Dirk Heider, and Christophe Caucheteur. Fiber bragg grating sensors toward structural health monitoring in composite materials: Challenges and solutions. *Sensors*, 14(4):7394–7419, 2014.

[40] Zensheu Chang and Ajit Mal. Scattering of lamb waves from a rivet hole with edge cracks. *Mechanics of materials*, 31(3):197–204, 1999.

[41] Yoonyoung Chung, Hyun Ho Kim, Sangryun Lee, Eunho Lee, Seong Won Kim, Seunghwa Ryu, and Kilwon Cho. Ubiquitous graphene electronics on scotch tape. *Scientific reports*, 5(1):1–8, 2015.

[42] Jonathan Coleman, Umar Khan, Werner Blau, and Yurii Gunko. Small but strong: a review of the mechanical properties of carbon nanotube–polymer composites. *Carbon*, 44(9):1624–1652, 2006.

[43] Steven Anton, Daniel Inman, and Gyuhae Park. Reference-free damage detection using instantaneous baseline measurements. *AIAA journal*, 47(8):1952–1964, 2009.

[44] Cliff J Lissenden, Yang Liu, and Joseph L Rose. Use of non-linear ultrasonic guided waves for early damage detection. *Insight-Non-Destructive Testing and Condition Monitoring*, 57(4):206–211, 2015.

[45] Hyung Jin Lim, Hoon Sohn, Martin P DeSimio, and Kevin Brown. Reference-free fatigue crack detection using nonlinear ultrasonic modulation under various temperature and loading conditions. *Mechanical Systems and Signal Processing*, 45(2):468–478, 2014.

[46] Yanxun Xiang, Mingxi Deng, Fu-Zhen Xuan, and Chang-Jun Liu. Experimental study of thermal degradation in ferritic cr–ni alloy steel plates using nonlinear lamb waves. *Ndt & E International*, 44(8):768–774, 2011.

[47] Sara Sarbaz, MA Motamedi, and Saied Irani. Crack detection in cantilever beam using nonlinear vibration of multiple frequency excitations. In *8th International Conference on Acoustics and Vibration, Shahid Beheshty University, Tehran, Iran*, 2018.

[48] Claire Prada and Mathias Fink. Separation of interfering acoustic scattered signals using the invariants of the time-reversal operator. application to lamb waves characterization. *The Journal of the Acoustical Society of America*, 104(2):801–807, 1998.

[49] Ros Ing and Mathias Fink. Time recompression of dispersive lamb waves using a time reversal mirror-application to flaw

detection in thin plates. In *1996 IEEE Ultrasonics Symposium. Proceedings*, volume 1, pages 659–663. IEEE, 1996.

[50] Ros Ing and Mathias Fink. Time-reversed lamb waves. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 45(4):1032–1043, 1998.

[51] Hyun Woo Park, Hoon Sohn, Kincho H Law, and Charles R Farrar. Time reversal active sensing for health monitoring of a composite plate. *Journal of Sound and Vibration*, 302(1-2):50–66, 2007.

[52] DN Alleyne, TP Pialucha, and P Cawley. A signal regeneration technique for long-range propagation of dispersive lamb waves. *Ultrasonics*, 31(3):201–204, 1993.

[53] Agostino Abbate, Jeff Koay, Julius Frankel, Stephan Schroeder, and Pankaj Das. Signal detection and noise suppression using a wavelet transform signal processor: application to ultrasonic flaw detection. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 44(1):14–26, 1997.

[54] Ryan Watkins and Ratneshwar Jha. A modified time reversal method for lamb wave based diagnostics of composite structures. *Mechanical Systems and Signal Processing*, 31:345–354, 2012.

[55] Igor Viktorov. *Rayleigh and Lamb waves: physical theory and applications. Transl. from Russian. With a foreword by Warren P. Mason.* Plenum press, 1967.

[56] Joseph Rose. *Ultrasonic guided waves in solid media.* Cambridge university press, 2014.

[57] Steven Anton, Daniel Inman, and Gyuhae Park. Reference-free damage detection using instantaneous baseline measurements. *AIAA journal*, 47(8):1952–1964, 2009.

[58] Gyuhae Park, Charles R Farrar, Francesco Lanza di Scalea, and Stefano Coccia. Performance assessment and validation

of piezoelectric active-sensors in structural health monitoring. *Smart Materials and Structures*, 15(6):1673, 2006.

[59] Jianxi Qiu, Fucai Li, Saqlain Abbas, and Yanping Zhu. A baseline-free damage detection approach based on distance compensation of guided waves. *Journal of Low Frequency Noise, Vibration and Active Control*, 38(3-4):1132–1148, 2019.

[60] Jianxi Qiu, Fucai Li, Saqlain Abbas, and Yanping Zhu. A baseline-free damage detection approach based on distance compensation of guided waves. *Journal of Low Frequency Noise, Vibration and Active Control*, 38(3-4):1132–1148, 2019.

[61] Mohammad Saleh Salmanpour, Zahra Sharif Khodaei, and Mohammad Hossein Aliabadi. Instantaneous baseline damage localization using sensor mapping. *IEEE Sensors Journal*, 17(2):295–301, 2016.

[62] Kaleeswaran Balasubramaniam, Rohan Soman, and Pawel Malinowski. Guided waves based damage detection based on quadrant area using optical fiber sensors. 2021.

Chapter 10

Experimental and numerical study of the additively manufactured carbon fibre reinforced polymers with fibre Bragg grating sensors

Torkan Shafighfard

Institute of Fluid Flow Machinery, Polish Academy of Sciences,
Fiszera 14, 80-231 Gdansk, Poland

10.1 Introduction

Rapid prototyping (RP) or additive manufacturing (AM), generally referred to as 3D printing, have all seen significant growth and attention in recent years. Additive manufacturing has definite advantages over conventional technology. It can theoretically construct any shape structure without molds, as well as speed up manufacturing times and reduce processing steps [1]. It has been extensively used in a variety of industrial applications, including the aerospace, automotive, dental, architectural, and medical sectors for the fabrication of various prototypes, particularly with complex geometries, as a result of recent technological advancements and the current trend toward a sustainable environment. As a result, it is anticipated that AM would usher in the third industrial revolution, complementing production line assembly, which began to rule industry in the previous century [2]. The advancement of technology was not restricted to manufacturing processes. The urge to make materials more efficient, such as smart materials, has been one of research and industry's major goals in relation to fabrication techniques. As a result of its high stiffness-to-weight ratio, corrosion resistance, decorativeness, and thermal stability, composite materials have seen a significant increase in demand over the last decade. They are, nonetheless, excellent for Structural Health Monitoring (SHM) in aerospace and mechanical engineering applications due to their inherent sensing capabilities [3, 4].

In recent years, AM of composite materials have gained interest in both an academical and industrial sectors. It was well understood that with minor processing problems, continuous fibres can be introduced directly to the polymer filament to produce a noticeable improvement in mechanical properties [5]. Among those techniques, FDM has become more popular since it uses a wide range of thermoplastic materials and is more affordable than other AM methods while having short lead time [6]. In this approach the initial components are mounted on a moving system in the form of a continuous filament. Based on a computer-aided design (CAD) file, the filament is heated and extruded into the nozzle layer by layer to create the entire object [7]. A key factor for a success-

ful printing process in this regard is the quality of the filaments, specifically their dimensional precision, surface finish, and uniformity [8]. The mechanical qualities of CFRP can be significantly increased through FDM printing and used as structural parts in the aerospace industry applications. On the other hand, FDM of CFRP exhibits excellent potential in the vehicle fields as inexpensive and lightweight honeycomb constructions [9].

In different literature, the mechanical characterisation of CFRP and pure polymer samples has been studied during the FDM processes [10–12]. However, a few investigations were carried out for the temperature elevation and reduction effect on the thermal behaviour of composite structures specifically to the heat deflection temperature of CFRPs [13, 14]. Also, the application of embedded FBG sensors on the AM composite structures has not been commonly studied in literature.

In order to examine the impact of temperature elevation and demotion on the mechanical behaviour of those samples, CFRP specimens were additively built using the FDM technique and tested in an environmental chamber. In order to mimic the thermo-mechanical behaviour of the created specimens under thermal stress, the fabricated specimens were also modelled in the Finite Element Analysis software program, ABAQUS. Using experimental work done in thermal chamber with FBG sensors incorporated in the CFRP material structure, the simulation's output was verified. The materials and procedures used to examine how temperature (above and below zero) affects CFRP additively synthesized with Fiber Bragg Grating (FBG) sensors. The outcomes of both numerical and experimental experiments are then given. Also, an effect of FBG sensor location (embedded) was investigated to understand whether FBG sensor within the sample could capture the thermal strain with reasonable error. It was studied for both elevated and subzero temperatures. The findings are then discussed and conclusions were provided.

10.2 Material and Methods

The analysed structures of AM sample with embedded FBG is presented in Fig. 10.1. The polymer sample was reinforced with carbon fibres.

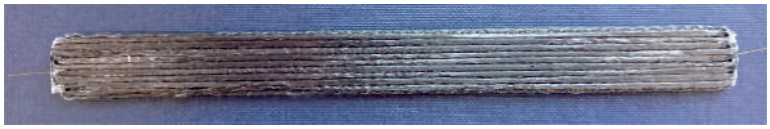


Fig. 10.1: Analysed sample: CFRP and FBG sensors'

The modified fused deposition modeling (mFDM) approach was employed to make the CFRP sample (150 mm x 15 mm x 2 mm) utilizing Geeetech's MeCreator 2 3D printer. Four layers of Polylactic Acid (PLA) matrix and continuous carbon fibre reinforcement make up the carbon fibre reinforced polymer (CFRP) sample. Referred to [15] for the fabrication details. The printing parameters are provided in Tab. 10.1.

Tab. 10.1: Parameters used for fabricating the sample'

Layer height (mm)	**First layer speed** (mm/s)	**Printing speed** (mm/s)	**Extruder temperature** ($^{\circ}C$)	**Bed temperature** ($^{\circ}C$)	**Extrusion multiplier** (mm)	**Extrusion width** (mm)	**Fan speed** (%)
0.5	1.20	4	200	70	0.6	1.6	50/80/100

Each sample had an FBG sensor (identified as "w") implanted during the AM procedure. In the CFRP sample, the sensor was positioned between the second and third layers. The primary axis of the samples was parallel to the orientation of the FBG sensor. The optical fibres with the sensors were parallel to the carbon fibre reinforcing direction for the CFRP sample.

Next, the Finite Element Method (FEM) was used to describe the material behaviour corresponding to the effect of temperature on the CFRP sample (FEM). Abaqus CAE software was employed to perform the numerical calculations for both subzero and elevated temperatures. Modeling the composite structure was carried out using the ply-by-ply technique inside this software. In this ap-

proach, 3D deformable solid is partitioned to the number of layers in the composite structure. Thereafter, the orientations were given based on the stacking sequences of the structure for each layer separately. Coupled temperature displacement solution is chosen to solve the problem. 9000 quadratic hexahedral completely integrated elements with a total number of 45069 nodes made up the FEM model (C3D20RT). The boundary conditions take into account the settings of the experimental study as well as the sample position in the environmental chamber on the shelf. It should be noted that the bottom portion of the sample, which was placed on a shelf in the measurement chamber, was restricted from moving in the thickness direction, while one side of the sample was restricted from moving in the y-direction and another side was restricted in the x-direction to allow the sample to expand. Also, the Abaqus software received the temperature elevation as a time-amplitude loading was applied for elevated and subzero temperatures which was exactly matched with the values given by an environmental chamber. Furthermore, in order to solve the convergence problem the size of the seeds are set at an identical distance of 1 mm along each edge after trial and error. Finally, because of the initial temperature of the chamber, the sample was defined at the predefined temperature of 20°C. The mechanical properties defined by the manufacturer except the thermal expansion coefficient which was found experimentally are shown in Tab. 10.2, which was given to the software through properties module.

Tab. 10.2: Mechanical properties of Fibre and Matrix.

	E [GPa]	ρ **[g/cm^3]**	ν	$\mathbf{X}_t$ **[MPa]**	**v [%]**	κ **[W/mK]**	**C [J/Kg·C°]**
Fibre	230	176	0.33	3530	24.5	10.46	794
Matrix	2.315	1.24	0.29	51	75.5	0.13	1800

Fig. 10.2 displays the sample that was modelled in Abaqus as well as the boundary conditions. For both the elevated and subzero temperatures the same model was utilized, but with different temperature definition for each sample. The temperature was read through the environmental chamber in each defined time-slot and input to the numerical simulation.

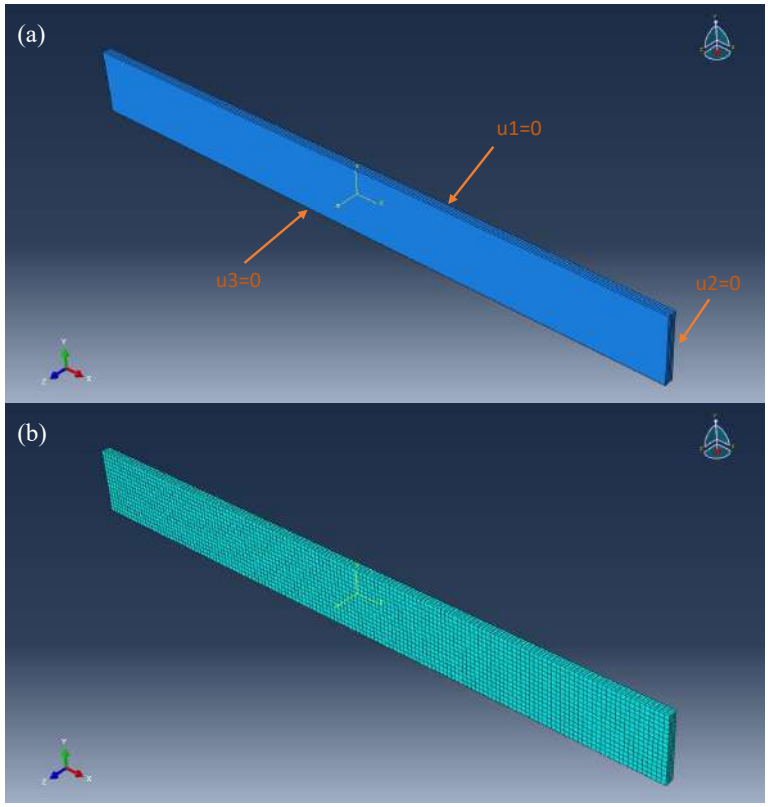


Fig. 10.2: FEM model: (a) boundary conditions, (b) meshed model.

10.3 Results and Discussion

The temperature influence test was separated into two sections: high temperatures rising from 10 °C to 50 °C, and subzero conditions falling from 10 °C to -50 degrees °C. The assumption in both situations was a 5°C step between two consecutive temperatures. The samples were kept on a lattice shelf during the research so they could expand in all directions. The measurements were made using a MyDiscovery DM600C environmental chamber (Angelantoni Test Technologies Srl, Italy). Interrogator si425-500 (Micron Optics, USA) with a measurement frequency of 1 Hz was used to measure the FBG sensors. The reference temperature for

the experimental and numerical studies was set at 20 °C. In order to understand the temperature change with respect to the time, Fig. 10.3 represents the environmental chamber program parameters for Elevated and Subzero temperatures. Thus, thermal strain in each time slots could be captured through combination of the previous and next figure.

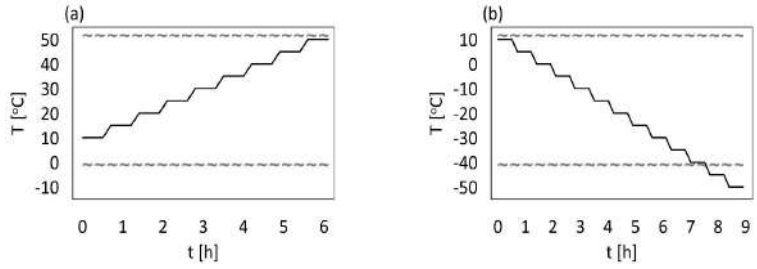


Fig. 10.3: Environmental chamber program parameters: (a) elevated (b) subzero temperature.

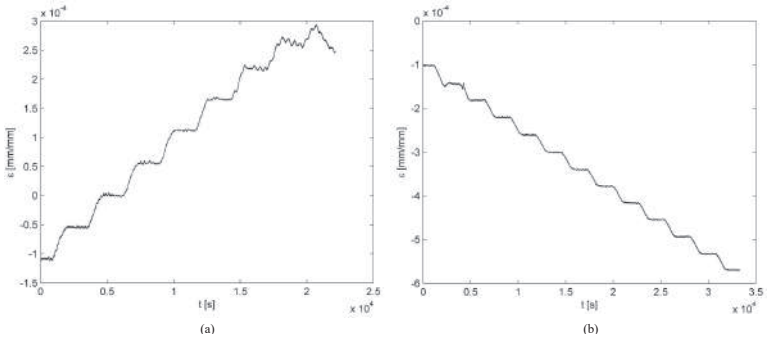


Fig. 10.4: Total strain for demoted temperatures given by embedded sensor.

The strain values of CRFP for the elevated temperatures is shown in Fig. 10.4 (a). It is worth mentioning that the behaviour of the composite structure in higher temperatures (around above 45°C) as shown in Fig. 10.4 (a) is different than lower temperatures and it was due to the fact that the PLA material inside the CFRP has not a constant coefficient of thermal expansion value within the

temperature elevation range.

The strain values of CRFP for the demoted temperatures is shown in Fig. 10.4(b). The strange behaviour of CFRP which existed in elevated temperatures is not a case in subzero temperatures and the thermal strain of the structure behaves normally.

A mesh convergence study was conducted and Fig. 10.5 represents that 1 mm edge seed size would be sufficient to get the correct numerical values and it is not required to decrease the mesh size in order to have the efficient computation.

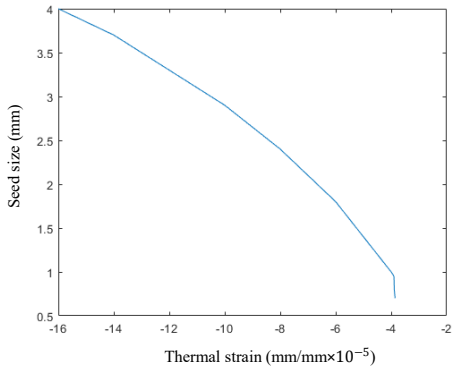


Fig. 10.5: Mesh convergence study.

An average thermal strain given by numerical simulation for second time slot is shown in Fig. 10.6.

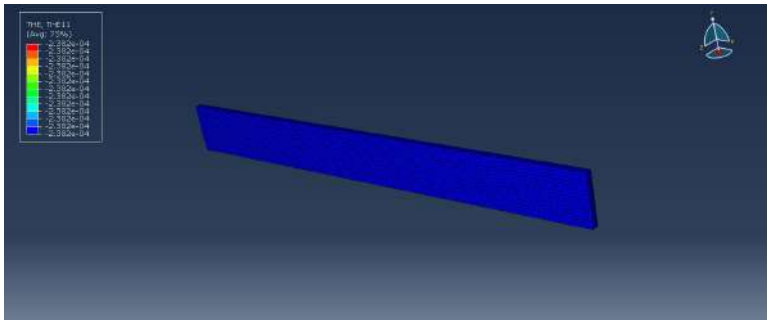


Fig. 10.6: Average thermal strain for second time-slot.

A comparison of strain values determined experimentally and numerically for CFRP sample in elevated temperature for embedded and attached FBG sensors is presented in Tab. 10.3. It should be noted that the thermal strain values given by Abaqus in each time-step is the average strain calculated by the software within the specimen. The average difference between the numerical values and the experimental values given by FBG embedded inside the sample is around 5%. The difference is found by subtracting the value of experimental result from the numerical one in which the result is divided by the experimental result for each temperature. Thereafter, the average of all values is considered as a final result. However, the thermal strain values provided by the FBG attached to the surface are not consistent with those given by ABAQUS. It was recognized that when sensors were applied to sample surfaces, variations were recorded that may be linked to the FBG sensors' placement on the carbon fibre and matrix or on the bundle of carbon fibre. A difference between the thermal strain values of Abaqus and the strains given by the embedded FBG sensor could be justified by the fact that in the numerical simulation the sample is considered to be an ideal specimen without any void inside it. Also, environmental effects were neglected in the Abaqus modelling.

Tab. 10.3: Numerical (ABAQUS) vs. experimental (FBG) strain values.

Temperature (°C)	**Num.** ($\times 10^{-5}$)	**Exp.** ($\times 10^{-5}$)	
	ABAQUS	**FBG surface**	**FBG center**
10	-3.95	-8	-4
15	-2.38	-5	-2.3
20	-0.1	1	0.26
25	1.95	1.5	2
30	4	6	4
35	6	15	6
40	6.12	18	6.1
45	8	20	8

A comparison of strain values determined experimentally and numerically for CFRP sample in demoted temperature is presented in

Tab. 10.4. The average difference between the numerical values and the experimental values given by FBG embedded inside the sample is around 4%. Similar to the elevated temperatures the FBG on the surface cannot capture the correct average thermal strain and the difference between these values and numerical simulations is obvious. Thus, the location of FBG sensor on the surface affects the value of thermal strain it captures.

Tab. 10.4: Numerical (ABAQUS) vs. experimental (FBG) strain values.

Temperature (°C)	**Num.** ($\times 10^{-4}$)	**Exp.** ($\times 10^{-4}$)	
	ABAQUS	**FBG surface**	**FBG center**
-50	-1.8	-3	-1.78
-40	-1.6	-2	-1.63
-30	-1.4	1	-1.4
-20	-1.18	2	-1.2
-10	-0.95	6	-1
0	-0.7	10	-0.65

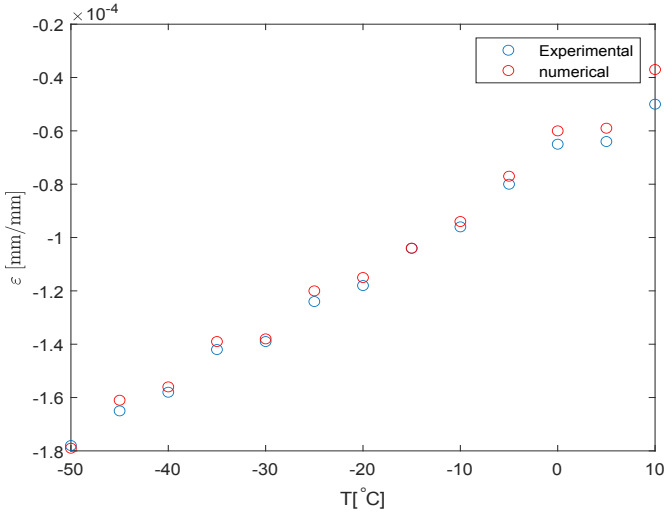


Fig. 10.7: Comparison of CFRP strain for subzero temperatures.

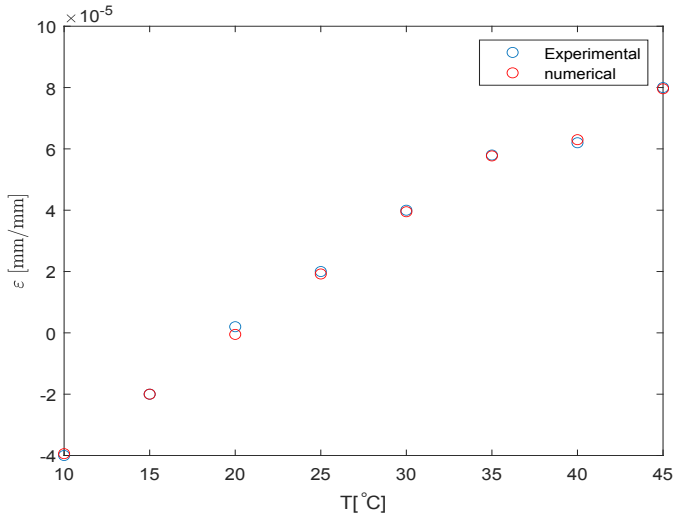


Fig. 10.8: Comparison of CFRP strain for elevated temperatures.

Fig. 10.7 and Fig. 10.8 shows the values noted experimentally and given by simulation numerically. The experimental results are presented for the embedded sensor as the mean mechanical strain values for stable temperatures during the steps. Fig. 10.7 represents the experimental and numerical values for the subzero temperatures. There is a good agreement between the strain values given by FEM and FBG sensors. Fig. 10.8 shows the experimental and numerical values for the subzero temperatures.

10.4 Conclusions

This chapter investigated the thermal behaviour of additively manufactured composite (CFRP) with embedded and attached FBG sensors for both elevated and demoted temperatures. FBG sensors were implanted into and affixed to the surfaces of the samples in both cases throughout the production process. It was shown how the Green method of manufacturing, e.g. AM, could be applied to fabricate CFRP structures containing FBG sensors in order to capture the thermal behaviour of structure.

The embedded sensor in the CFRP sample displays the genuine strain values in the element. However, the attached sensor only detected the behaviour of the carbon fibre, hence the measurement of thermal strain for FBGz cannot be verified. Thus, it was concluded that the FBG sensor inserted in the sample's midsection monitors the CRFP specimen's thermal behaviour with greater accuracy than the FBG sensor attached to the outer surface. This was as a result of the FBG sensor being manually glued to the exterior surface. Consequently, the FBG sensor that was originally incorporated during 3D printing might be used to study the behaviour of complicated materials. The percentage discrepancy between the experimental and numerical findings around 5% for both the elevated and subzero temperatures. Though, the difference is less in demoted temperatures which might be due to the low impact of lower temperatures on the structural behaviour of the polymer which is not a case in high temperatures.

The FEM approach may be used to calculate thermal strain values in an AM element for the temperature range under consideration for both subzero and elevated temperatures as the results are in good agreement with the experimental studies. Finally, As a natural component of the SHM systems during the composite structures 3D printing and post-processing, the embedded FBG sensors might be used to show the thermal behaviour of those materials.

Acknowledgement

This research is supported by the following project: 'Additive manufactured composite smart structures with embedded fibre Bragg grating sensors (AMCSS)' funded by the National Science Centre, Poland under M-ERA.NET 2 Call 2019, under grant agreement no. 2019/01/Y/ST8/00075.

The author would like to thank the researchers from the Kaunas University of Technology, Lithuania for the sample manufacturing. ABAQUS calculations were carried out at the Academic Computer Centre in Gdańsk (Poland).

Bibliography

[1] Wenqing Wang, Lu Zhang, Xingjie Dong, Jianqin Wu, Qing Zhou, Suwen Li, Chujing Shen, Wei Liu, Gang Wang, and Rujie He. Additive manufacturing of fiber reinforced ceramic matrix composites: Advances, challenges, and prospects. *Ceramics International*, 2022.

[2] Torkan Shafighfard and Magdalena Mieloszyk. Model of the temperature influence on additively manufactured carbon fibre reinforced polymer samples with embedded fibre bragg grating sensors. *Materials*, 15(1):222, 2021.

[3] Torkan Shafighfard, Eralp Demir, and Mehmet Yildiz. Design of fiber-reinforced variable-stiffness composites for different open-hole geometries with fiber continuity and curvature constraints. *Composite Structures*, 226:111280, 2019.

[4] Doug Munghen, Vincent Iacobellis, and Kamran Behdinan. Incorporation of fiber bragg grating sensors in additive manufactured acrylonitrile butadiene styrene for strain monitoring during fatigue loading. *International Journal of Fatigue*, 154:106485, 2022.

[5] Fuda Ning, Weilong Cong, Jingjing Qiu, Junhua Wei, and Shiren Wang. Additive manufacturing of carbon fiber reinforced thermoplastic composites using fused deposition modeling. *Composites Part B: Engineering*, 80:369–378, 2015.

[6] Meysam Mohammadi Zerankeshi, Sayed Shahab Sayedain, Mobina Tavangarifard, and Reza Alizadeh. Developing a novel

technique for the fabrication of pla-graphite composite filaments using fdm 3d printing process. *Ceramics International*, 2022.

[7] Samuel Clinton Daminabo, Saurav Goel, Sotirios A Grammatikos, Hamed Yazdani Nezhad, and Vijay Kumar Thakur. Fused deposition modeling-based additive manufacturing (3d printing): techniques for polymer material systems. *Materials today chemistry*, 16:100248, 2020.

[8] Ana Pilar Valerga, Moisés Batista, Jorge Salguero, and Frank Girot. Influence of pla filament conditions on characteristics of fdm parts. *Materials*, 11(8):1322, 2018.

[9] Xingshuang Peng, Miaomiao Zhang, Zhengchuan Guo, Lin Sang, and Wenbin Hou. Investigation of processing parameters on tensile performance for fdm-printed carbon fiber reinforced polyamide 6 composites. *Composites Communications*, 22:100478, 2020.

[10] Kouider Bendine, Dennis Gibhardt, Bodo Fiedler, and Alexander Backs. Experimental characterization and mechanical behavior of 3d printed cfrp. *European Journal of Mechanics-A/Solids*, 94:104587, 2022.

[11] Ning Su, Robert S Pierce, Chris Rudd, and Xiaoling Liu. Comprehensive investigation of reclaimed carbon fibre reinforced polyamide (rcf/pa) filaments and fdm printed composites. *Composites Part B: Engineering*, 233:109646, 2022.

[12] Nabeel Maqsood and Marius Rimašauskas. Characterization of carbon fiber reinforced pla composites manufactured by fused deposition modeling. *Composites Part C: Open Access*, 4:100112, 2021.

[13] Vinayak C Gavali, Pravin R Kubade, and Hrushikesh B Kulkarni. Mechanical and thermo-mechanical properties of carbon fiber reinforced thermoplastic composite fabricated using fused deposition modeling method. *Materials Today: Proceedings*, 22:1786–1795, 2020.

[14] Isyna Izzal Muna and Magdalena Mieloszyk. Temperature influence on additive manufactured carbon fiber reinforced polymer composites. *Materials*, 14(21):6413, 2021.

[15] Marius Rimašauskas, Tomas Kuncius, and Rūta Rimašauskienė. Processing of carbon fiber for 3d printed continuous composite structures. *Materials and Manufacturing Processes*, 34(13):1528–1536, 2019.

Chapter 11

Different approaches towards fabrication of photoelectrochemical cells dedicated for solar to fuel energy conversion

Manjunath Veeranna Shinnur

Institute of Fluid Flow Machinery, Polish Academy of Sciences,
Fiszera 14, 80-231 Gdansk, Poland

11.1 Introduction

The exponential growth in energy demand has resulted from the growth of the world population and advances in civilization. Despite the fact that fossil fuels are not sustainable and have serious environmental and health consequences, they remain the primary source of energy because they are cheap, easily available, effective, and have capacity to generate huge amounts of electric energy in a single location. During the combustion of fossil fuels, greenhouse gases such as methane, carbon dioxide, and nitrous oxide are emitted in huge quantities. Due to the increasing rate of civil and industrial growth, it is also projected that emissions of these gases will increase over time [1, 2] . The current state of greenhouse gases and predicted future state (assuming no change in energy sources) will result in weather changes, sea-level rise, severe health problems, ecosystem changes and other effects. Humans are threatened by climate change and health concerns caused by fossil fuels [1, 3, 4]. According to International Energy Agency (IEA) 2021 prediction [5], global oil demand would decline from 90 barrels a day to about 24 barrels a day by 2050 and it also called for no fossil fuel cars beyond 2035 in order to limit global heating and to achieve net zero emissions by 2050. Peak oil camp claims that the world will run out of oil in 2052, gas in 2060, and coal in 2088, but that subsequent finds such shales will keep the supply expanding [3]. To address these issues, different nations have begun to reassess their energy strategies and policies. Developing new devices that are highly efficient, improving the efficiency of existing technologies and transition into renewable energy resources are most considered routes to reach the aim of sustainable future [1, 3, 6].

Renewable energy can meet two-thirds of global energy demand and contribute to the majority of greenhouse gas emission reductions required between now and 2050 to limit global average surface temperature rise below 2°C [4, 7]. Solar energy is the most interesting renewable energy source since it is the largest renewable source (with 120,000 TW reaches the Earth's surface which is 104 times more than current global energy demand) [8, 9].The use of carbon dioxide (CO_2) as a chemical feedstock intends to save fossil fuels

and reduce greenhouse gas (GHG) emissions by providing an alternative carbon source can eventually close the carbon cycle. These characteristics may provide the impression that CO_2 usage is sustainable [10–12]. Hydrogen is storable, clean and environmental beneficial fuel that produces water on combustion, without release of greenhouse gases, particulates or any pollutants into atmosphere. Cognately, there are some disadvantages of hydrogen energy such as hydrogen is a highly flammable and volatile substance due to which working with hydrogen fuel becomes dangerous, storage and transportation of hydrogen gas is difficult as it is a lightest gas. Currently, around 95% of hydrogen comes from fossil fuels, steam reforming of natural gas and petroleum, with remaining 5% comes from water electrolysis. Since, this process uses non-renewable resources and consumes huge amount of energy for production of hydrogen thus this process is considered to be non-sustainable and not economically viable [3, 12, 13]. The cost of solar cells has reduced dramatically through several decades of development, with lifespan exceeding 25 years, and there is hope for increasing application in the future. However, because solar energy is intermittent, a really large-scale implementation will necessitate an equally big energy storage capability [6, 12]. One approach in- volves artificial photosynthesis, a chemical process converting solar energy into fuel by photo-assisted water splitting to generate hydrogen or reduce CO_2 emissions. To attain this technological goal, high-efficiency and low-cost photocatalyst must be developed. A photocatalyst is generally a semiconductor material which alters the rate of chemical reaction on exposure to light [6, 8, 13]. Artificial photosynthesis can be achieved in one of three ways: photovoltaic cell combined with electrolysis; photocatalytic cell; photoelectrochemical (PEC) cell. Each approach has its own arguments for and against.

This chapter includes brief description of the state of art within the research area of photoelectrochemical tandem cells for water splitting and CO_2 reduction. There will be discussed the challenges involved in water splitting and CO_2 reduction process and fabrication of efficient photoelectrodes for PEC tandem cells facilitating the maximum solar energy conversion efficiency.

11.2 Overview – Photoelectrochemical Cells

From past few decades, there has been a lot of research towards photoelectrochemistry based on photoactive semiconducting electrodes, especially in the application of these systems to solar energy [14]. During the early development of semiconductors applied in electronic devices, the modern era of photoelectrochemistry began at Bell Laboratories. Researchers at Bell Laboratories immersed several semiconductors, such as Ge and TiO_2, and studied their electrochemical response in light and dark [15]. In 1966 Gerishcher was the first to study the electrochemical behaviour of semiconductors under illumination and the studies explained the energetics and kinetics of dark and photo induced charge transfer between semiconductor and electrolyte interface [15–17]. Research on photoelectrochemical cells went overwrought after the 1973 oil crisis, which stimulated a worldwide search for alternative energy sources [18]. In 1972 Fujishima and Honda investigated the electrochemical photolysis of water at the semiconductor electrode, by conducting series of experiments using n-type semiconductor TiO_2 (rutile) which provided a base for further research to find the connection between semiconductor, photoelectrochemistry and solar energy [16, 18–20]. In 1991 O'Regan and Gratzel developed a low cost thin film photoelectrochemical PV device (Dye sensitized solar cells).

Photoelectrochemical cells (PECs) uses light as an energy source to generate electricity. Each cell includes one or two semiconducting photoelectrodes, with auxiliary metal and reference electrodes, immersed in an electrolyte [21, 22]. A photoelectrochemical cell dedicated for artificial photosynthesis is made up of a pair of electrodes, one or both of which are composed of a photoactive substance, and an electrolyte that ensures the flow of current in the closed electric circuit. The photoactive electrode material plays role as a photocatalyst, absorbing and converting solar energy at the oxidation-reduction potentials required for water splitting or CO_2 reduction reactions [23].

Fig. 11.1 presents the classification of photoelectrochemical cells. Under illumination on the surface of n-type semiconductors, an ox-

idation process occurs, while a reduction reaction takes part at p-type element. Solar energy is converted into electrical energy without changing the free energy of the redox electrolyte (ΔG=0) in the electrochemical photovoltaic cell, where a narrow bandgap semiconductor and a redox couple is present. In photoelectrosynthetic cell there are two effective redox couples in electrolyte, and free energy of electrolyte changes due to material illumination. Photosynthetic cells are further classified into photoelectrolysis cell with positive free change and photocatalytic cell with negative free energy change where light stimulates cell response in a contra-thermodynamic path [15, 23, 24].

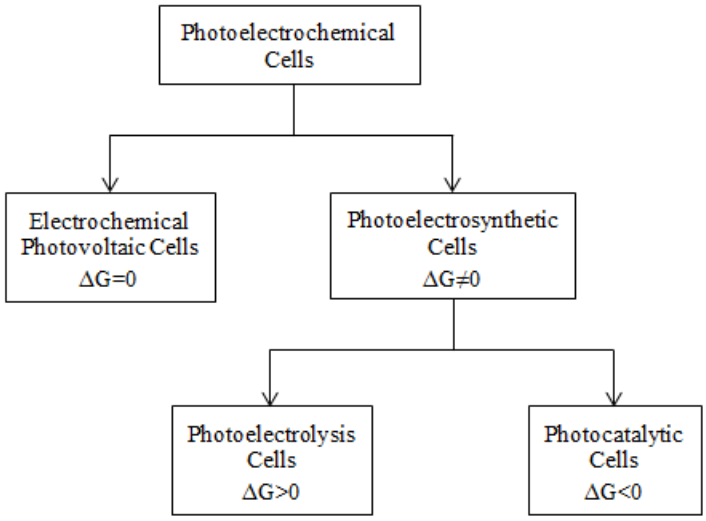


Fig. 11.1: Classification of photoelectrochemical cells. Source based on reference [15].

Semiconductor Electrolyte Interface

When a particular semiconductor is immersed in an electrolyte, all phenomena related with photoelectrochemical systems are based on the formation of a semiconductor-electrolyte junction.

Upon contact a space charge layer arises in the semiconductor. The

chemical potential of electrons in semiconductors is determined by the Fermi level of the semiconductor. On the other hand the chemical potential of liquid electrolytes are governed by the potential of the redox couples present in the electrolyte. These potentials are also associated with the electrolyte's Fermi level [15].

When the initial Fermi level in an n-type semiconductor is higher than the initial Fermi level in the electrolyte, electrons are transferred from the semiconductor to the electrolyte, resulting in equilibration of the two Fermi levels (or chemical potentials). In the semiconductor, this results in a positive space charge layer (also known as a depletion layer because the region is depleted of majority charge carriers). The conduction and valence band edges are bent upwards as a result, creating a potential barrier to further electron flow into the electrolyte. When p-type semiconductors have an initial Fermi level lower than the electrolyte, an inverted but equivalent scenario develops [20].

In the electrolyte, a charged layer exists adjacent to the solid electrode interface, known as the Helmholtz layer. This layer has a significant impact on the band bending that occurs when the semiconductor equilibrates with the electrolyte. The band bending would be expected to equal the difference in initial Fermi levels between the two phases without the Helmholtz layer [23, 25].

11.3 Photoelectrochemical Cells for Solar Water Splitting

11.3.1 History of Water Splitting

Water electrolysis is a well-known process that has been refined over the course of two centuries. PEC devices for solar water splitting integrate essential design characteristics taken from commercial electrolysis. Polymer Electrolyte Membrane (PEM) electrolysers first developed in the 1960s, with the benefits of a faster response to power input and the ability to deliver hydrogen at high pressure, but at the cost of more complex reactors and expensive materials [26]. Commercial electrolysers of both types could cur-

rently achieve efficiencies of up to 80%. With the development of increasingly efficient and less expensive photovoltaic (PV) technologies in recent decades, the concept of combined photovoltaic-electrolysis developed. The essential concept of PV-electrolysis is basically coupling PV-cell to an electrolyser. The best integrated PV-electrolysis system has a solar to hydrogen (STH) efficiency of 30% after 48 hours. PV cells made of crystalline silicon might have efficiency of up to 27.6% [27]. But combining these high efficient PV cells and electrolysers results in expensive devices, need of power input and some losses.

Photoelectrochemical (PEC) water splitting method is considered to be an ideal technology for producing hydrogen. One of the most essential tasks in the PEC water splitting process is photogenerated carrier migration [27]. In 1972, Fujishima and Honda for the first time demonstrated the water photolysis process by titanium dioxide electrode as photoanode. However due to its large band gap of 3-3.2eV, TiO_2 absorbed only 4% of the AM 1.5 solar spectrum. The hydrogen evolution reaction (HER) and the oxygen evolution reaction (OER) take place on two different photoelectrodes, allowing separation of the generated product. It satisfies all of the criteria for environmentally friendly and sustainable development, making it the most promising technique of hydrogen production so far [28, 29]. In PEC water splitting, semiconductor materials are used as photocatalysts, and the photogenerated electrons and holes react with water to produce H_2 and O_2, respectively. There are many p-type semiconductors studied as photocathode material in PEC for water reduction, including InP, $GaInP_2$, WS_2, $Cu(In,Ga)Se_2$, Cu_2O and GaP. On the other hand, semiconductors such as TiO_2, WO_3, $BiVO_4$, $ZnFe_2O_4$, Fe_2O_3, TaON and Ta_3N_5 are studied as photoanode materials in PEC for water oxidation [30]. The maximum photocurrent value is based on the semiconductor's band gap. Some semiconductor materials have high theoretical photocurrent but in experimental process the generated photocurrent density can be much lower than its theoretically estimated value. This is due to the recombination of photogenerated electron-hole pair with energy loss during transportation. Therefore, it is significant to facilitate the transport process of photogenerated electron-hole pair to

achieve high photoelectric conversion efficiency [31]. Above mentioned data are collected from recent research in the field of PEC cells for energy conversion.

11.3.2 PEC Water Splitting Mechanism

Photoelectrochemical cell is composed of 3 main components: an anode, a cathode, and an electrolyte. As already mentioned, oxygen evolution reaction (OER) oxidizes water to produce oxygen at the anode, whereas the hydrogen evolution reaction reduces hydrogen ions to hydrogen gas at the cathode (HER). Depending on the PEC cell's arrangement, the cathode, anode, or both can be photoactive semiconductors that absorb light. This method not only avoids the time-consuming manufacturing procedure, but it also lowers the cost of the system [29]. Despite significant research employing a variety of semiconductor configurations, there is still much work to be done to achieve the desired efficiency and stability. The optimal solar to hydrogen (STH) efficiency for a practical based photo-catalytical system is 10%. This efficiency is important to down the H_2 production costs. Decomposing water into H_2 and O_2 is a thermodynamically endothermic reaction (ΔG0 = +237.2 kJ/mol H_2). This implies that the decomposition reaction requires additional energy [32]. The amount of Gibb's free energy required to split the water molecule into H_2 and O_2 is +237.2 kJ/mol which corresponds to ΔE0 $\approx$ 1.23 eV. The photoelectrode must provide a minimum of 1.23 eV per electron to run this process. The water decomposition can be written in the following two reactions as shown in the Eq.11.1 and Eq.11.2 [33].

Water splitting half reaction at photoanode (OER):

$$H_2O \xrightarrow{Light} 2H^+ + 1/2O_2 + 2e^- \quad \Delta E = 1.23\, vs\, NHE \tag{11.1}$$

NHE: Normal Hydrogen Electrode

Water splitting half reaction at photocathode (HER):

$$2H^+ + 2e^- \xrightarrow{Light} H_2 \quad \Delta E = 0.00\, vs\, NHE \tag{11.2}$$

Overall reaction:

$$H_2O \xrightarrow{Light} H_2 + 1/2O_2 \quad \Delta E = 0.00\ vs\ NHE \tag{11.3}$$

Water splitting half reaction performed in PEC cells using either photoanode or photocathode system. By absorbing a photon, a semiconducting photoelectrode (photoanode/photocathode) in a PEC cell generates electrons in the conduction band and holes in the valence band. At the photoanode/electrolyte contact, these photogenerated holes can oxidize the water molecule. Photo-generated electrons, on the other hand, form hydrogen at the photocathode/ electrolyte [34] interface.

When a semiconducting photoelectrode is immersed in an electrolyte, charge transfer occurs between the photoelectrode and the electrolyte until equilibrium is reached. In the case of a photoanode, upward band bending happens at the photoanode's surface to match the semiconductor's Fermi level to the water redox level. At the photocathode surface, down-ward band bending is formed to align the Fermi level with the water redox level. The concentration of electrons/holes at the photoanode/ photocathode-electrolyte interface is depleted as a result of this energy level band bending, and this region is known as the depletion layer. An electric field is developed as a result of the creation of a depletion layer, which is beneficial for charge carrier separation [34, 35].

Photogenerated electrons and holes created on photoelectrodes under illumination can theoretically drive the HER and OER, as long as the bandgap of the photoelectrode is equal to or bigger than the minimum energy (1.23 eV) for splitting water molecules. Although there is a large difference between reported and theoretically expected efficiencies of PEC cells, many efforts have been made to develop an efficient PEC water splitting cell. Low efficiency results from the low light absorption by photoelectrode, high recombination, and kinetic losses. Furthermore, a better STH efficiency is required to attain a commercially feasible level, which is a fundamental criterion [36]. For unassisted overall water splitting, the configuration of dual-absorber tandem systems exhibits more advantages compared to traditional one side configuration. Photoanode/

photocathode systems or photoelectrode/ photovoltaic (PEC/PV) coupled devices can be used to produce dual-absorber PEC tandem devices [37, 38]. As a result, dual-absorber tandem devices can provide enough driving force for self-driven solar water splitting while also increasing the fraction of solar energy collected.

11.3.3 PEC/PEC Tandem Cell

The use of a photocathode and a photoanode coupled in series to construct a PEC/PEC tandem cell, in which sunlight can flow through the n-type semiconductor and then the p-type semiconductor. More specifically, photons with shorter wavelengths in the solar spectrum are absorbed by the top electrode in this configuration (the photoanode), while the photons with longer wavelength are transmitted and these photons are captured by bottom electrode in this configuration (the photocathode) [30]. Then, photons can be used step by step at the sunlight absorption path. Because each photoelectrode only needs to contribute a portion of the potential for overall water splitting, the two semiconductors with narrower band gaps can be chosen over the single photoelectrode from a thermodynamic viewpoint. As indicated in Fig. 11.2, each photoelectrodes band gap should cover the O_2/H_2O or H^+/H_2 standard electrode potential, and the photoanode's conduction band is more negative than the photocathode's valence band. More photons can be absorbed and a higher STH efficiency can be achieved with a smaller band gap in the semiconductor material [36]. In conclusion, while selecting semiconductor materials, ensure that the energy required for water splitting is derived entirely from the light source. A PEC cell with a TiO_2 photoanode and a p-GaP photocathode was presented in 1975 by Yoneyama *et al.* [39]. Without any external bias, the n-TiO_2/p-GaP tandem cell was found to generate both hydrogen and oxygen. Because of the high internal resistance of the n-TiO2/p-GaP cell, it exhibited only 0.25% STH efficiency. Xue *et al.* demonstrated [40] an unbiased PEC cell with nanoporous $BiVO_4$ photoanode co-catalysed by cobalt borate sandwiched with TiO_2 layers. $BiVO_4$ is co-catalysed with ultra small cobalt borate through photo-assisted electrodepostion process. The charge carri

rs properties of Cobalt borate co-catalysts was aided by conformal TiO_2 layer placed on $BiVO_4$ surface. The $BiVO_4$ photoanode produced a remarkable stable photocurrent of 2.5 mA/cm^2 at 1.23 V vs. RHE (reversible hydrogen electrode). This solar-driven water splitting tandem cell exhibited 3% STH conversion efficiency.

Yin *et al.* [41] designed an integrated unbiased solar water splitting tandem cell using FeOOH modified $TiO_2/BiVO_4$ photoanode and Cu2O as photocathode which is shown in the Fig. 11.3. This tandem cells achieved STH of 0.46% which is calculated from J-V data. Next, for 2.5 h, the current density of PEC tandem cell dropped to 0.07 mA/cm^2 but still remains about 58% of initial current density (1.2 mA/cm^2). The drop in the current density value was due to the poor stability of Cu_2O photocathode. The amount of H_2 and O_2 evolved by this tandem cell without bias were 2.36 μmol/cm^2 and 1.09 μmol/cm^2 respectively after 2.5 h of testing.

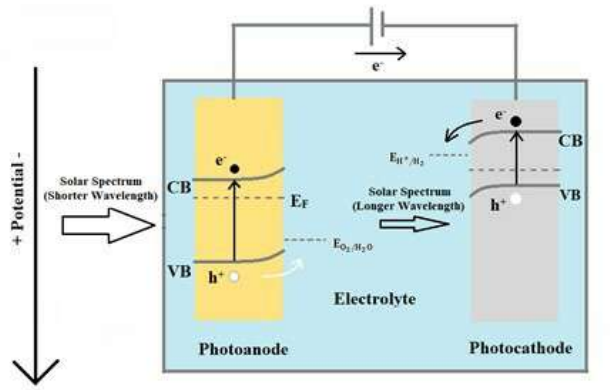


Fig. 11.2: Illustration of PEC/PEC tandem cell configuration (VB and CB-valence and conduction band respectively, EF-Fermi level).

To obtain n-type and p-type semiconductor properties, Leygraf *et al.* employed Mg-doped and Si-doped Fe2O3 electrodes; the resulting PEC tandem cell of polycrystalline Fe2O3 exhibited 0.05% STH efficiency in 0.1 M Na_2SO_4 [42] solution. Grimes *et al.* demonstrated a PEC tandem cell composed of p-type Cu-Ti-O nanotube array

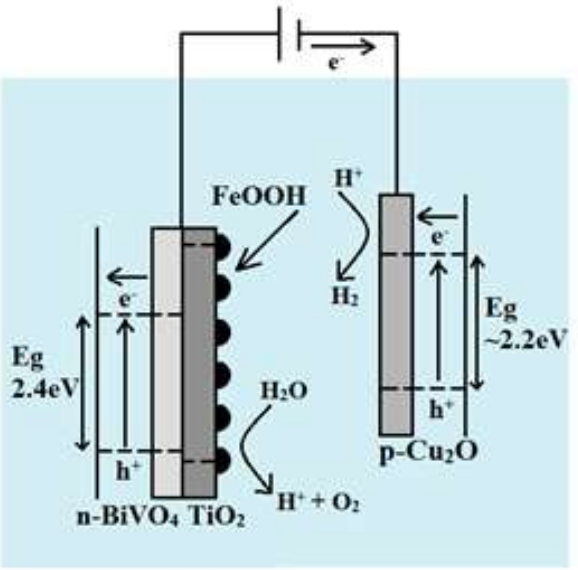


Fig. 11.3: Schematic diagram of FeOOH modified $BiVO_4$-Cu_2O PEC tandem cell (Eg-Energy band gap).

photocathode and n-type TiO_2 photoanode and this cell achieved a STH efficiency of 0.3% [43]. Finally, the PEC/PEC tandem device has a theoretical maximum STH efficiency of around 29.7% [36]. However, there are a number of problems that can lead to the system's inefficiency in real conditions. Higher efficiency may be achieved if the onset potential regions of the two photoelectrodes overlapped more. But the performance of the tandem device could be limited if one of the photoelectrodes has a lower photocurrent. The high internal resistance loss in PEC/PEC configuration contributes to the low efficiency. Furthermore, considering the corrosion and stability issues of both photoelectrodes, selecting an electrolyte becomes problematic. Despite the fact that the PEC/PEC system has a high theoretical STH efficiency, large-scale implementation could take a long time. Above mentioned data are obtained from recent research in the field of PEC cells for energy conversion.

11.3.4 Determination of Efficiency for PEC Water Splitting

There are various efficiency standards determination of PEC cell operation effectiveness. The characterization of photoelectrode materials performance can be carried out according to following Eqs. 11.4-11.7 [34, 44, 45]:

- Solar to Hydrogen Efficiency (ηSTH): It can be expressed as ratio between total energy generated to total energy input from sunlight irradiation.

$$STH = \frac{Total\ Energy\ Generated}{Total\ Energy\ Input} = \frac{\Delta G \times r_{H_2}}{P_{light} \times S} \tag{11.4}$$

  where, ΔG is Gibbs free energy (237 kJ.mol^{-1}), r$_{H_2}$ is rate of hydrogen production (mols^{-1}), P$_{light}$ is incident light intensity (100 mWcm^{-2}) and S is surface area of photoelectrode (cm^2)

- Applied Bias Photon-to-Current Efficiency (ABPE): An abillty of photoelectrode materials to perform the water splitting reaction is defined by the applied bias resulting in recorded current under simultaneous electrode illumination.

$$ABPE = \frac{Total\ Energy\ Generated}{Total\ Energy\ Input} = \frac{J_{ph} \times (1.23 - V)}{P_{light}} \tag{11.5}$$

  where, J$_{ph}$ is photocurrent density at a specified applied volt age, V is the applied voltage and P$_{light}$ is incident light intensity (100 mWcm^{-2}).

- Incident Photon to Current Efficiency (IPCE): The measurement is carried out at fixed incident wavelengths to evaluate the performance of photoelectrode. Typically the wavelength are selected by particular step from the radiation generated by solar simulator. IPCE is defined the number of photogenerated charge carriers contributing to the photocurrent per incident photon.

$$IPCE = \frac{Total\ energy\ of\ converted\ electrons}{Total\ energy\ incident\ photon} = \frac{1240 \times J_{ph}}{P_{light} \times \lambda} \tag{11.6}$$

where, J_{ph} is photocurrent density at specific wavelength, P_{light} is incident light intensity at specific wavelength and λ is wavelength of incident light.

- Absorbed Photon to Current Conversion Efficiency (APCE): Optical losses such as reflection and transmission are ignored when calculating IPCE, despite the fact that they can have a significant impact on the final IPCE value. The absorbed photon-to-current conversion efficiency (APCE) is frequently used to determine the quantum efficiency in attempt to correct for these optical losses. It is defined as number of photogenerated charge carriers contributing to photocurrent per absorbed photon which is termed as internal quantum efficiency.

$$APCE(\lambda) = \frac{IPCE(\lambda)}{1 - R - T} \times 100 \tag{11.7}$$

where, R and T are optical reflection and transmission respectively.

11.3.5 Challenges in PEC Tandem Cell for Water Splitting

Various researchers have aimed to solve the issues discussed earlier in the chapter by modifying photoelectrodes in a number of different ways, such as:

(i) Tandem configuration to generate enough photovoltage for water splitting,

(ii) Surface passivation layers to overcome recombination losses,

(iii) Incorporating a cocatalyst to improve kinetics,

(iv) Integrating protection layers to address the stability issue,

(v) Elemental doping to improve the conductivity.

Though these modifications were carried out, the experimental efficiency values of most photoelectrode materials were found to be far from expected theoretical values [34]. In order to drive the HER and OER, there will always be concentration-dependent and kinetic overpotential losses associated with electron transfer processes

at semiconductor/liquid junctions. Taking into account losses, the energy required for water photolysis will be around 1.6-2.4 eV. The possible losses that influence the efficiency of PEC are as follows:

(i) reflection losses,

(ii) quantum yield losses,

(ii) absorption losses,

(iv) collection losses [44, 46].

The bandgap and absorption coefficient of the photoelectrode material are closely related to photo-induced charge carrier generation. The output of a photoelectrode under light irradiation (photocurrent) is the measurable quantity that determines the overall efficiency of photoelectrodes. Many investigations have addressed light absorption properties, band edge position, crystallographic quality and physical properties in order to develop materials that are suitable for visible light-driven overall water splitting. In further sections of the chapter, strategies to overcome these challenges or limitations are discussed in detail.

11.4 Photoelectrochemical Cells for CO_2 Reduction

11.4.1 Background of CO_2 Photoreduction

CO_2 conversion is not a straight forward task as it is an extremely stable molecule, generally produced by burning of fossil fuels. Conversion of CO_2 into its useful state by activation or reduction is a scientifically challenging problem. For CO_2 conversion photoelectrochemical (PEC), phototcatalysis, electrocatalysis, thermocatalysis, radiolysis and biochemical techniques have all been used previously. Among them PEC has been found to be one of the most promising techniques for converting CO_2 into fuels and value-added products (e.g., methane, formate, methanol, ethanol and many more) under solar light irradiation. This technique is also called as artificial photosynthesis [47].

PEC reduction of carbon dioxide combines the benefits of pho-

tocatalysis with electrolysis, and represents a viable path to solar-driven chemical fuel synthesis from CO_2 [48]. The photocatalytic conversion of CO_2 into fuels using H_2O, utilizing solar energy as the driving force, has received a lot of attention since Halmann's study in 1978 [49]. In general, photoconversion systems for CO_2 reduction can be carried out in two configurations. The first configuration of the photocatalytic system (PC) uses a suspension of photocatalyst particles in a solvent for the reduction dissolved CO_2. This type of system is made up of widely accessible photocatalysts and simple devices, where solar energy acting as the single driving factor for CO_2 reduction. However, because both reduction and oxidation reactions take place on separate surface sites of the same photocatalyst particle, resulting in mixture of products. The main drawback of this configuration, in the absence of hole scavenger the reduced product from CO_2 might undergo re-oxidation by photogenerated holes or by produced O_2 [50]. The photoelectrochemical (PEC) cell system is the other configuration, which consists of a semiconducting photoelectrode (working electrode), a counter electrode, and frequently a reference electrode. The photoelectrode harvests light to generate charge carriers and performs a half-cell reaction to reduce CO_2 on the photocathode, while the counter electrode performs the other half-cell reaction. As a result, the reduction and oxidation products can be easily separated. Furthermore, the maximum theoretical solar conversion efficiency for PC system is predicted to be 46.7% whereas for PEC system it is reported as 23.8%. Application of external bias would further increase the maximum theoretical efficiency of PEC system [51–53].

11.4.2 Principle of photoelectrochemical CO_2 reduction

During the PEC CO_2 conversion, the electrode material is immersed in the electrolyte solution that ensures closing the whole circuit. When the Fermi level (EF) of the electrode material or semiconductor is not equal to that of the electrolyte, electron transfer process occur. The reactions occur at the electrode/ electrolyte interface to reduce the EF difference between the electrode and the

electrolyte, implying band bending or changes in the band diagram. The concentration of an electroactive species determines its EF. Until the EF levels are equal, electron transfer happens at the interface [54]. P-type semiconductors act as photocathodes in PEC processes, while n-type semiconductors take photoanodes role.

At the conduction band (CB) and valence band (VB), respectively, electrons and holes are produced in the semiconductor when light is irradiated. The predominant carriers for n-type (EF is near the CB) and p-type (EF is just above the VB) are electrons and holes, respectively. As a result, for n-type and p-type, upward and downward band bending could occur at the interface, respectively. To improve CO_2 conversion efficiency, the band bending phenomenon is required to separate photogenerated electrons and holes at the electrode-electrolyte interface [55].

CO_2 is a very stable molecule, and its C=O bond has a greater dissociation energy of 750 $kJmol^{-1}$ than many other chemical bonds, such as C-H (430 $kJmol^{-1}$) and C-C (336 $kJmol^{-1}$), this indicates that the transformation of CO_2 requires a lot of energy. Furthermore, carbon is in its maximum oxidation state in CO_2, thus CO_2 reduction can result in a wide range of products which includes gaseous products such as CO and CH_4 and higher hydrocarbons such as CH_3OH, HCOOH [56]. In fact, utilizing p-type semiconductors to make photocathodes is a typical technique for PEC CO_2 reduction (Fig. 11.4a). p-Si, p-Cu_2O, p-GaP, p-InP and p-CdTe silicon, metal oxides, sulfides, tellurides, phosphides, and other p-type semiconductors have all been exploited and investigated. Above mentioned data are obtained from recent research in the field of PEC cells for energy conversion.

Furthermore, metal materials (e.g., Pd, Au, Ag, and Cu) are commonly used as cathodes in electrochemical CO_2 reduction, owing to their high activity for CO_2 adsorption and activation. As a result, using an effective electrochemical CO_2 reduction catalyst as cathode and a photoanode for harvesting solar energy and delivering photogenerated electrons is another viable strategy to PEC CO_2 reduction (Fig. 11.4b). These two methods both utilizes photoelectrodes for harvesting light. A PEC cell combining photoanode and

photocathode has also been developed in the frames of the device construction to maximize light absorption which highly fit for CO_2 reduction and H_2O oxidation (Fig. 11.4c) [51, 57].

Tab. 11.1 shows the potentials in relation to a normal hydrogen electrode (NHE) reference in an aqueous solution at pH 7. Depending on the different reaction mechanisms and pathways, CO_2 reduction can produce CH_4, CO, CH_3OH, HCOOH, C_2H_6, C_2H_5OH, and even $H_2C_2O_4$. The CO_2 reduction mechanism involves first step of the formation of a $CO_2^{\cdot-}$ intermediate, the reaction has negative redox potential of -1.9V vs NHE.

Tab. 11.1: Thermodynamics potentials vs the normal hydrogen electrode (NHE) at pH 7 for various CO_2 reduction products. Reprinted with the permission from reference [50].

Product	**Reaction**	**E_0 (V vs NHE)**
$CO_2^{\cdot-}$ intermediate	$CO_2 + e^- \rightarrow CO_2^{\cdot-}$	-1.90
Hydrogen	$2H_2O + 2e^- \rightarrow 2OH^- + H_2$	-0.41
Methane	$CO_2 + 8H^+ + 8e^- \rightarrow CH_4 + H_2O$	-0.24
Carbon monoxide	$CO_2 + 2H^+ + 2e^- \rightarrow CO + H_2O$	-0.51
Methanol	$CO_2 + 6H^+ + 6e^- \rightarrow CH_3OH + H_2O$	-0.39
Formic acid	$CO_2 + 2H^+ + 2e^- \rightarrow HCOOH$	-0.58
Ethane	$CO_2 + 14H^+ + 14e^- \rightarrow C_2H_6 + 4H_2O$	-0.27
Ethanol	$CO_2 + 12H^+ + 12e^- \rightarrow C_2H_5OH + H_2O$	-0.33
Oxalate	$CO_2 + 2H^+ + 2e^- \rightarrow H_2C_2O_4$	-0.87

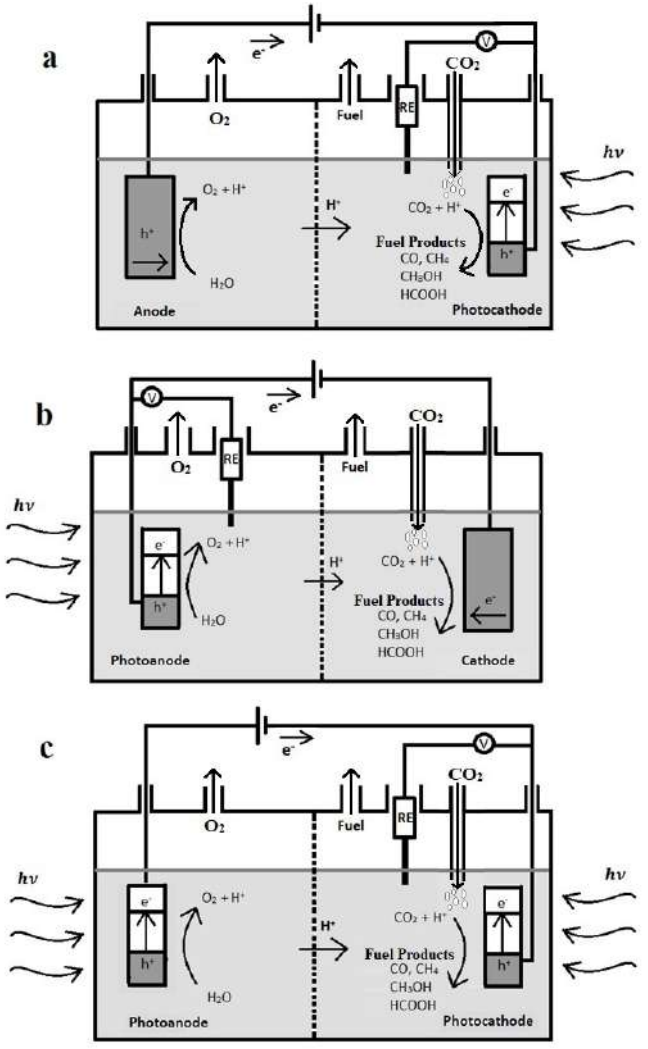


Fig. 11.4: Schematic diagram for PEC CO_2 reduction in water using semiconductor as (a) photocathode, (b) photoanode and (c) both photocathode and photoanode (RE-Reference electrode and $h\nu$ – *Photon energy*).

11.4.3 Photoanode and Photocathode System for CO_2 Reduction

Photoanode and photocathode system is an ideal model for CO_2 reduction on photocathode side and water oxidation on photoanode side without any external electrical bias. It is very difficult for materials selected as photoanode and photocathode to meet perfectly all the listed requirements [57]:

(i) Suitable band gap to absorb sufficient visible light and prevent recombination of photogenerated electron-hole pairs;

(ii) Suitable band positions to establish Z-scheme junction to ensure electrons to photocathode from photoanode, leaving holes in photoanode for subsequent redox reactions;

(iii) Both photoanode and photocathode remain stable in aqueous electrolyte for long time;

(iv) Efficient CO_2 electrocatalyst applied on photocathode to promote CO_2 reduction process.

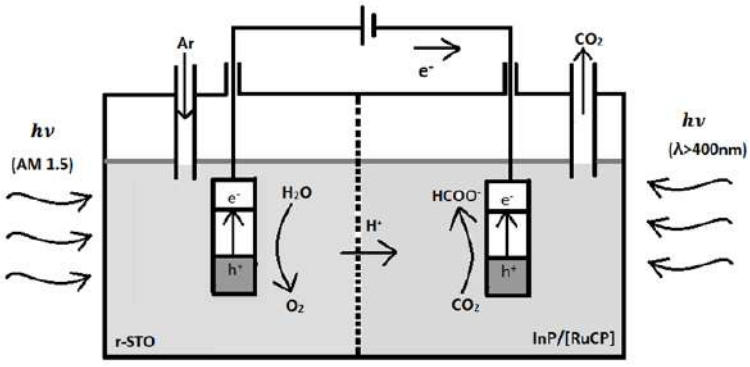


Fig. 11.5: Schematic representation of the photoelectrochemical reduction of CO_2 with two-electrode configuration in Z-scheme system without no external bias.

Sato *et al.* [58] developed a Z-scheme configuration comprising of p-type lnP/Ru complex polymer hybrid photocatalyst as photocathode and TiO_2 photoanode for PEC reduction of CO_2 to formate

without external bias. The selectivity for HCOO- was found to be >70% and conversion efficiency of solar to chemical energy achieved was 0.03-0.04%. To further improve the efficiency and reaction selectivity it's important to optimize energy-band configurations, and catalyst structure. Bias free design is important for sustainable practical application of artificial photosynthesis using solar irradiation to produce organic species such as alcohols, hydrocarbons and syngas. Arai *et al.* [59] developed two electrode configuration in Z-scheme system composed of indium phosphide and ruthenium complex polymer (RuCP) electrocatalyst, lnP/[RuCP] hybrid photocathode and reduced strontium titanate ($SrTiO_3$) photoanode as shown in Fig. 11.5 to reduce CO_2 to formate. They achieved 0.14% conversion efficiency from solar to chemical energy.

11.4.4 Adsorption and activation of CO_2

Adsorption and activation of CO_2 on the surface of heterogeneous catalysts are both critical for subsequent reduction reactions and suppression of competing HER. CO_2 adsorption results in the formation of partially charged species $CO_2^{\cdot-}$ through interaction with surface atoms. Several strategies such as increasing surface defects, catalyst surface area and noble-metal co-catalysts can promote the chemisorption of CO_2 [50]. He *et al.* [60] reported that anatase (101) facet played a significant role in adsorbing CO_2 and promoting electron transfer from surface of TiO_2 to CO_2 in the process of CO_2 reduction, apart from its role in photogeneration and separation of electron-hole pair. Liu *et al.* synthesized single crystalline ultrathin $ZnGa_2O_4$ nanosheets with over 99.6% (110) facets exposed for photoreduction of CO_2, image of ultrathin scaffold $ZnGa_2O_4$ nanosheet is shown in the Fig. 11.6a. The reduction in the thickness of nanosheets down to several nanometers offered a high specific area and resulted in high CO_2 adsorption ability and exhibited around 35% enhancement in photocatalytic conversion performance of CO_2 to CH_4 in comparison to meso-$ZnGa_2O_4$ nanosheets which is shown in Fig. 11.6b [61]. Anpo *et al.* [62] synthesized zeolite catalysts containing highly dispersed TiO_2. These Ti containing zeolites catalysts exhibited almost 10 times greater

photocatalytic efficiency for CH_4 generation and to achieve high selectivity for CH_3OH generation when compared to TiO2 catalysts in the form of powder.

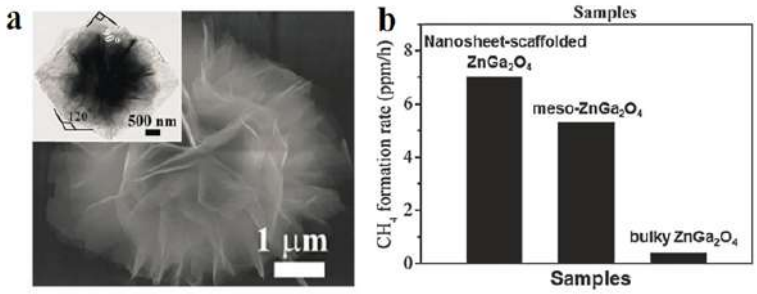


Fig. 11.6: (a) Field emission scanning electron microscopy (FE-SEM) and Transmission electron microscopy (TEM) images of ultra-thin ZnGa2O4 nanosheets and (b) photocatalytic CH4 production by various ZnGa2O4 samples. Reprinted with the permission from reference [61].

Surface defects, such as oxygen vacancies, act as the most reactive sites on metal oxide surfaces, which are capable of altering the structure and change the electrical and chemical properties semiconductor surface [50, 62]. Oxygen vacancies play a vital role in governing adsorption and activation of CO_2. Lee *et al.* [63] investigated electron induced dissociation of CO_2 adsorbed at the oxygen vacancy defect on TiO2 (110) surface using scanning tunnelling microscopy (STM). The surface area of TiO2 (110) before and after thermal diffusion of CO_2 molecules away from oxygen vacancy sites, reveals the oxygen vacant sites under CO_2. These interactions between oxygen vacancies and CO_2 molecules would lead to lower activation energy barrier and enhances the trapping and activation of CO_2.

Oxygen atoms in CO_2 molecule has lone pair of electrons which could be donated to surface Lewis acid centres. On the other hand C atom in CO_2 molecule can gain electrons from Lewis base centres such as oxides ions forming carbonate like species. Thus, CO_2 molecule can act as an electron donor and acceptor simultaneously.

Therefore, another strategy to improve CO_2 adsorption and activation is by introducing functional basic sites on to photocatalyst surface. Some basic oxides or hydroxides such as NaOH, CaO, Ga_2O_3 and MgO act as electron donor for interaction of C atoms in CO_2 molecule. Wang *et al.* found that the addition of MgO on the surface of TiO_2 resulted in high amount of CO generation on photoreduction of CO_2 and more efficient formation of CH_4 during CO_2 photoreduction was observed by adding MgO onto Pt-TiO_2 [50]. In summary, an effective strategy for surface modification of heterogeneous photocatalysts can significantly improve CO_2 adsorption and activation, as well as minimize the barriers to subsequent CO_2 reduction.

11.4.5 Evaluation parameters for PEC CO_2 reduction

- Product evaluation rate and catalytic density: The product evolution rate is commonly normalized over the photoelectrode's effective area (i.e., $molh^{-1}cm^{-2}$) or the mass of the catalyst on the photoelectrode (i.e., $molh^{-1}{g_{cat}}^{-1}$). The catalytic current density is usually normalized over the photoelectrode area (i.e., $mAcm^{-2}$). The reaction efficiency might be directly reflected by the product evolution rate and catalytic current density values.
- Quantum yield: It is a crucial parameter to be considered when assessing photocatalyst performance. Internal and apparent quantum yields are defined as following Eqs. 11.8-11.9.

$$Internal\ quantum\ yield = \frac{\alpha \times n_{product}}{n_{photon}} \times 100\% \qquad (11.8)$$

$$Apparent\ quantum\ yield = \frac{\alpha \times n_{product}}{n'_{photon}} \times 100\% \qquad (11.9)$$

  where, α is the number of electrons required for product evolution, $n_{product}$ and n'_{photon}, is the number of absorbed and incident photons respectively.
- Farafaic Efficiency (FE): Various products will be generated during the PEC CO_2 reduction process due to the complex reaction route. As a result, selectivity is a crucial parameter to

evaluating CO_2 reduction. Faradaic efficiency (FE) has been widely used in PEC CO_2 reduction to evaluate the selectivity of targeted products and is calculated using the following Eq. 11.10:

$$FE = \frac{\alpha \times n_p \times F}{Q} \tag{11.10}$$

where, α is number of electrons required for product evolution, n_p is the molar number of targeted products, Q is the total passed charges and F is faraday's constant.

- Turnover number and turnover frequency: In the PEC reduction of CO_2, the activity of catalytic active sites is an accurate parameter that can be measured using turnover number (TON) and turnover frequency (TOF). The following Eqs. 11.11-11.12 are used to calculate TON and TOF.

$$TON = \frac{n_p}{n_c} \tag{11.11}$$

$$TOF = \frac{n_p}{n_c \times t} \tag{11.12}$$

where, n_p is the total number of targeted products, n_c is the molar number of catalyst and t is the reaction time.

11.4.6 Challenges in PEC Cell for CO_2 reduction

Some limitations in the photoelectrochemical process for CO_2 reduction are in Sec. 11.3.5. Now, more challenges will be described. The synthesis of carbonaceous products often requires the transfer of more electrons (eight and six, for CH_4 and CH_3OH generation) than in the case of HER (two-electron transfer). The low solubility of CO_2 in aqueous solutions (e.g., 0.033 $molL^2$ of $KHCO_3$ at 25°C under 1 atm) presents a kinetic challenge in CO_2 reduction, resulting in mass transfer limitations for carbonic product generation. The photoelectrodes in PEC systems, in particular, are fixed, and there is a high concentration gradient of reactants between the electrode's surface and the bulk of the electrolyte. For high conversion efficiency, mass-transfer limitations in electrode kinetics should be carefully examined and overcome [64]. Fabrication

of the photoelectrodes and loading of co-catalysts are complicated and challenging tasks in CO_2 reduction process. Also, to loading of co-catalysts, the structure optimization between photoelectrodes and co-catalysts should be considered carefully to achieve high solar to chemical efficiency. Many p-types semiconductors (Cu_2O, GeP, ZnTe) are used as photocathodes due to their relatively negative CB edge position, which can favours large driving force in CO_2 reduction reactions. Most of these semiconductors suffer from photocorrosion due to their poor stability in aqueous electrolytes. The stability of these photocathodes can be improved by coating of surface with protective layers using various techniques such as atomic layer deposition (ALD), sputtering deposition and electron beam evaporation. The difficult working mechanisms and reaction pathways for the photocathode-photoanode system and the PV-PEC system should be gain wider attention in order to clearly demonstrate the mechanism, making them more acceptable and feasible in the practical use of CO_2 reduction processes.

11.5 Photoelectrode Materials

To perform efficient PEC reactions one of the critical factor is to choose photoelectrode materials with excellent stability, appropriate energy band gaps with respect to redox potentials of water, good carrier conductivity, and excellent absorption of visible light. But it's practically impossible for a material to possess all these properties. Many semiconducting photoelectrode materials have been discovered in the past, but their conduction and valence band potentials are incompatible with water redox levels. Many metal oxides based photoelectrodes were intensively investigated for water splitting due to their favorable semiconducting properties and chemical stability. In addition, for practical uses, photoelectrode materials should be low-cost and made of earth-abundant elements. This is important in establishing the need for solar-to-fuel devices to be scaled up at a reasonable cost. However many of those metal oxides exhibited large energy band gap, which resulted in poor visible light absorption and poor electrical characteristics [34].

TiO_2 has been investigated as a photoanode for PEC water splitting, but its performance has been limited due to its large energy band. Various attempts have been made to overcome poor absorption problem of TiO_2 by using dyes with good visible light absorption ability, as well as elemental doping to modify the energy band gap. Other materials such as $BiVO_4$ was considered for photoanode and has been explored extensively for PEC water oxidation.

Researchers obtained 90% of the theoretically predicted STH efficiency, but this material has low chemical stability [38, 40]. WO3 is another candidate that may perform oxygen evolution reactions under visible light, but due to its large energy band gap and poor charge dynamics, it shows poor photocurrent performance [34].

Other examples is Cu_2O that has been explored extensively as photocathode because to its optimal energy band gap (2.0-2.2 eV) and excellent semiconducting properties such as mobility. However it suffers from aqueous stability. To achieve PEC performance from Cu_2O, a protective layer is always necessary. The PEC performance of silicon-based photoelectrodes integrated with various co-catalysts for water splitting is high, but a trade-off between the co-catalyst loading on the surface of Si and light absorption is a major obstacle.

11.6 Photoanode

11.6.1 TiO_2 Photoanode

Titanium dioxide (TiO_2), an n-type semiconductor with a bandgap of 3.2 eV, it has been frequently used in PEC water splitting research, due to its strong photostability, non-toxicity, abundance, and low cost. The broad bandgap, on the other hand, restricts its ability to utilize visible light, and the low charge carrier mobility promotes electron-hole recombination [65]. To improve PEC performance, a common approach concerns increase the surface-to-volume ratio and formation of nanostructured TiO_2 photoelectrode. Especially one-dimensional (1D) nanostructures have attracted a lot of attention, because of its potential to provide a facile electron transport pathway along the longitudinal direction.

Grimes *et al.* reported anodized TiO_2 nanotubes for solar water splitting as shown in the Fig. 11.7a and Fig. 11.7b [66].

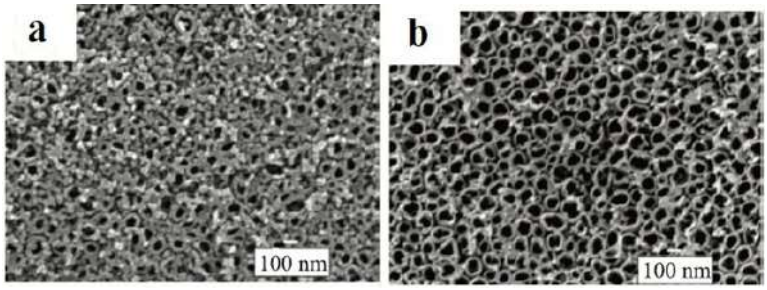


Fig. 11.7: Different TiO_2 morphology through nanostructuring: Anodized TiO_2 NTs at (a) 5°C and (b) 25°C. Reprinted with the permission from reference [66].

The modified TiO_2 nanotube achieved a photoconversion efficiency of 6.8%. Kim *et al.* [67] reported facile and effective surface functionalization of TiO_2 NTs through $TiCl_4$ mediated solution treatment and its effects on charge transfer and transfer properties for PEC water-oxidation. Flame method (i.e., rapid annealing using a high temperature flame under reducing atmosphere) surface treatment generated oxygen vacancies on the nanotubes surfaces, which can improve charge transfer efficiency by increasing the donor concentration. Apart from 1D nanostructure, 2D nanostructures have also showed promising results. Butburee *et al.* [68] demonstrated 2D anatase TiO_2 porous single crystalline nanostructure (PSN) on fluorine-doped tin oxide substrate (FTO) through ion exchangeinduced pore-forming technique. This technique successfully generated significant porosity in single-crystalline nanostructures with maintaining outstanding charge mobility and causing no structural damage. For photo-electrochemical water splitting, the PSN TiO_2 film delivered a photocurrent of 1.02 mAcm-2 at a relatively low potential of 0.4 V vs RHE, which is close to the theoretical value (1.12 mAcm-2) of TiO2. Subramanyam *et al.* [69] reported that metal oxide photoelectrodes sensitized with the narrow-band-gap semiconductor Bi2Se3 improved the PEC performance by extending the light response beyond the visible region. The incorporation

of Bi_2Se_3 into TiO_2 electrode, the extent of hydrogen production was found to be increased. TiO_2/Bi_2Se_3 photoelectrode exhibited a photocurrent density of 1.76 mAcm-2 at 1.23V, which is three times higher than that of pure Bi_2Se_3 nanoflowers. TiO_2/Bi_2Se_3 photoelectrode resulted the STH efficiency of 1.01% at 0.6V and IPCE efficiency of 10.5%.

11.6.2 Other Photoanode Materials

Hematite (Fe_2O_3) is the 4th most ample element on earth (6.3% by weight). Due to its high natural abundance, Fe_2O_3 is a promising photoanode material with high chemical stability, low toxicity, and low cost. It also has a band-gap value of 1.9 to 2.32 eV, which allows for visible light absorption and thus hematite can attain a theoretical maximum STH efficiency of 15% [34]. Hematite, on the other hand, has a conduction band position that is substantially more positive than the proton reduction potential, hence it can only be used for PEC water oxidation in the presence of external bias. Hematite has some disadvantages which includes [46]:

(i) Short charge carrier life time, due to fast charge carrier recombination,

(ii) Relatively low absorption co-efficient and,

(iii) Poor water oxidation kinetics.

To overcome these limitations many strategies were applied to improve the activity of hematite based photoelectrodes. Gratzel *et al.* demonstrated Si-doped hematite modified with IrO_2 co-catalyst to achieve a photocurrent density of 3.75 mAcm-2 [70]. Youn *et al.* [71] synthesized single crystalline hematite photoanode by two step annealing treatment. In addition to it, photoanode was Pt-doped to improve charge transfer characteristics and, an oxygen evolution co-catalyst (Cobalt-phosphate) to modify the surface properties. This hematite photoanode showed 4.32 $mAcm^{-2}$ photoelectrochemical water oxidation current at 1.23V vs RHE under solar irradiation (100 $mWcm^{-2}$). This photocurrent value corresponds to 34% of maximum theoretical limit exhibited for hematite with band gap of 2.1 eV. Wang *et al.* [57] fabricated $Ni(OH)_2/IrO_2$ co-catalyst on

Ti doped hematite photoanode for solar water splitting. It was reported that The Ni(OH)2 cocatalyst was found to be able to store charges (holes) from the photoanode.

$BiVO_4$ is a promising n-type semiconductor with a direct bandgap of 2.4 eV that has excellent water oxidation capability and it always acts as a visible light-active metal oxide photoanode in PEC water splitting systems. The theoretical STH efficiency of this material under regular AM 1.5 solar irradiation is 9.2% percent, which is close to the commercialized requirement for solar water splitting [63]. In 2013 Abdi *et al.* [72] developed first BiVO4 tandem cell with CoPi modified gradient W-doped $BiVO_4$ photoanode achieved a STH efficiency of 4.9% which is slightly better than Fe2O3 based tandem cell. In 2018 Xue *et al.* [40] demonstrated a highly active undoped $BiVO_4$ co-catalyzed with cobalt borate and surface modified with TiO_2 passive layer to increase the carrier life time and to enhance the charge separation and transportation of carriers. The resulting BiVO4 photoanode achieved a remarkable stable photocurrent of 2.5 $mAcm^{-2}$ at 1.23 V vs RHE. The tandem cell composed of TiO_2/Cobalt Borate-modified $BiVO_4$ photoanode with amorphous Si-photocathode achieved an STH efficiency of 3%. Above mentioned data are obtained from recent research in the field of PEC cells for energy conversion.

11.6.3 Synthesis of TiO_2 nanotubes via Electrochemical Anodization

Ti anodization was found in electrochemical literature since 1950's which divulged detailed information about this technique and also about valve metal behaviour. Among others, niobium, zirconium, and titanium is naturally covered by protective oxide film, however the thickness of the oxide film can be increased by oxidation method known as electrochemical anodization. Anodic oxidation is an electrochemical method that controls the thickness and the morphology of oxide layer formed over a substrate material. Even with complex surface patterns [73].

In last few decades researchers have achieved control over the anodizing process in order to develop engineered TiO_2 surfaces. This

technique includes polarizing titanium by imposing electric current between titanium specimen and a counter electrode, this causes metal atoms to oxidize to Ti^{4+} ions, which combines with oxygen anions present in the electrolyte to form a compact oxide layer on the substrate surface. This growth mechanism involves both inward and outward growth due to field assisted ions migration with prevalence of oxygen ions charge carriers across the oxide layer to reach the metal surface where Ti^{4+} ions are produces [74].

Electrochemical parameters such as applied voltage, current, electrolyte composition, pH and operating temperature all these parameters have effect on the oxide layer properties. Electrochemical cell setup, counter electrode material and different electrochemical cell geometries have a vital effect on the material morphology. Some of the most important cell geometries adapted to synthesis good oxide layers are shown in the Reference [75].

11.6.4 Quantum Dots on Photoanodes

Quantum Dots (QDs) light absorbers are commonly employed to decorate a host semiconductor with a larger bandgap than the QDs. QDs-based photoelectrodes can be made using a number of techniques, including sequential ionic layer adsorption and reaction (SILAR), chemical bath deposition (CBD), electrophoretic deposition, direct adsorption, and linker-assisted assembly. Depending on the size, shape, surface of QDs, and morphology of the host semiconductors, QDs may adsorb on the surface of nanotubes, nanorods, or the pores of mesoporous thin films. The host semiconductor serves as both a scaffold and a charge acceptor for the QDs, with the band shift between the QDs and the host material acting as a driving force for efficient charge separation [76]. By expanding the absorption toward the visible (400 nm - 700 nm) or NIR regions, QDs-sensitization overcomes the light absorbance limit exhibited by the host materials. Due to the narrow band gap of QDs, the photons with the energies corresponding to visible light may be responsible for the charge carriers production injected then into host semiconductors [77]. It is very important to understand properties and charge dynamics of QDs in order to improve the STH efficiency

and to make the system practically viable. Among others, chalcogenide quantum dots have attracted much attention as building blocks for light-harvesting devices because of its exceptional optical features, such as a wide light absorption range throughout the near-IR regions and a high extinction coefficient. Chalcogenide quantum dots have been actively utilized for generation of hydrogen in PEC water splitting cells due to its superior water durability compared to other light harvesting materials such as halide perovskites and metal organic dyes. It should be underlined here that titania based materials have excellent stability in strong base electrolytes due to which TiO_2/chalcogenide QD photoanodes are found to be more attractive for PEC hydrogen generation in comparison to ZnO/chalcogenide QD photoanodes. Consequently, development of TiO_2/chalcogenide QD photoanodes in PV cells with high photocurrent density greater than 30 $mAcm^{-2}$ is extremely attractive for PEC applications [78]. Kim *et al.* [77] developed quantum dot-sensitized TiO_2 photoelectrodes, by applying TiO_2 surface passivation layer (SPL) on mesophorous TiO_2 film surface to prepare TiO_2/QD (PbS and CdS) photoanodes and achieved photocurrent density of 34.59 $mAcm^{-2}$ for PV cells, one of the highest value reported. The SPL enhanced the surface states density and facilitated multiple trapping and detrapping transport. This combined with reduced electron recombination, resulted in high collective efficiency. SPL treated TiO_2/QD photoanodes showed excellent photocurrent density of 14.43 $mAcm^{-2}$ at 0.82 V vs RHE in PEC water splitting cell.

11.7 Photocathodes

This section has been devoted to discuss some of the important materials to fabricate efficient photocathodes such as p-type NiO semiconductor material as it is highly stable in acidic and under photoelectrosynthesis conditions, materials based on semiconductor quantum dots like lead sulfides and copper indium sulfide QDs as they act as good light absorbers and they exhibit good carrier multiplication efficiency and Cu and Cu_2O catalysts as they exhibit favorable elctronic and chemical properties and high elemental abund-

ance for CO_2 reduction.

11.7.1 NiO

p-NiO is a p-type semiconductor material with bandgap ranging from 3.3 eV-4.3 eV. The conductivity of p-NiO crystallites can be enhanced by increasing the concentration of Ni3+ via lithium doping, nickel vacancies, or interstitial oxygen. The broad bandgap of the p-NiO semiconductor limits its ability to use a wide range of the solar spectrum, limiting generated photocurrent density. As a result, research on the use of p-NiO for PEC water splitting and CO_2 reduction is hampered [65]. DuChene *et al.* [79] demonstrated light-induced modulation of catalytic selectivity for photoelectrochemical CO_2 reduction in aqueous mediums using copper nanoparticles dispersed onto p-type nickel oxide (p-NiO) photocathodes. Hu *et al.* [80] fabricated NiO photocathodes by anodizing Ni foil in an organic based electrolyte for PEC system. Al_2O_3 was used as blocking layer to modify the surface of NiO in order to minimize the surface charge recombination and to enhance IPCE. On the other hand, Frehan *et al.* [81] demonstrated that Cu-doping of NiO has the potential to reduce surface charge recombination and to provide more effective photocathodes.

11.7.2 Cu_2O

Cuprous oxide (Cu_2O) is the most widely studied p-type semiconductor in the PEC CO_2 reduction systems, due to its lower toxicity, low cost, and abundance on the earth. More notably, the Cu_2O CB is lower than the CO_2 reduction potential, with a relatively narrow bandgap (1.9-2.2 eV). Cu_2O due to its sufficiently more negative CB than CO_2 reduction potential, it can act as catalyst for CO_2 reduction. According to the theoretical studies, undersolar irradiation, the maximum solar-to-fuel conversion efficiency of Cu_2O may reach even 18%, and the maximum photocurrent density is 14.7 $mAcm^{-2}$ [57]. Despite this, many studies show substantially lower photocurrent densities and solar-to-fuel conversion efficiencies than predicted values, due to severe photocorrosion of Cu_2O in aqueous solution. Photocorrosion of the photocathode will signific-

antly decrease the overall reduction efficiency because most photogenerated holes are absorbed by Cu_2O [34]. Surface coating is one of the technique to overcome the photocorrosion problem and to increase the life time of electrode. $CuO_2/CuO/Al_2O_3$ composite produced by coating with CuO and Al_2O_3 layer by layer demonstrated improved stability and PEC performance. Under AM 1.5 light irradiation $Cu_2O/CuO/Al_2O_3$ composite exhibited high photocurrent density of 1.8 $mAcm^{-2}$ when compared to 0.25 $mAcm^{-2}$ exhibited by bare CuO_2 flim. CuO cover layer over Cu_2O facilitated electron transfer process and increase in Al_2O_3 layers led to increased stability of electrodes [82]. Paracchino *et al.* [83] demonstrated a highly effective electrodeposited p-Cu_2O electrode that was protected by atomic layer deposition (ALD) of Al-doped ZnO and TiO_2 nanolayers and covered with electrodeposited platinum. In 1 M Na_2SO_4 solution, the electrode produced photocurrents as high as 7.6 $mAcm^{-2}$ at 0 V vs RHE. The relationship between the products and various Cu-based photocathodes are summarized in Tab. 11.2.

Tab. 11.2: Relationship between the products and different Cu-based photocathodes. Reprinted with the permission from reference [84].

Photocathodes	**Condition**	**Main products**	**Faradaic efficiency**
Cu_2O/Al:ZnO/TiO_3/ Re(tBu-bipy)(CO_3)Cl	2 mM Re(tBu-bipy) $(CO)_3Cl$ + 0.1M MeOH, -1.73 V vs Fc+/Fc	CO	100%
Cu/Cu_2O	0.1 M Na_2SO_4 -0.2 V vs SHE	CH_3OH	95%
$CuInS_2$/pyridine	0.1 M $NaHCO_3$ -0.54 V vs SHE	CH_3OH	97%
Al:ZnO/CdS/ (Cu(In,Ga)Se_2)	0.1 M $KHCO_3$ -0.2 V vs RHE	CO	99.30%
CuO/Fe_2O_3	0.1 M $KHCO_3$ -1.1 V vs SCE	CH_3OH C_2H_5OH	91.20%

11.7.3 Quantum Dots on Photocathodes

The development of QDs-sensitizers for photocathodes is confined to graphene QDs, PbS QDs, CdSe QDs, CdSeTe QDs, CdTe QDs, carbon QDs, and CdSe/CdS QDs, and other materials. Mismatch in photocurrent density and catalytic efficiency of photoanode and photocathode may restrict the development of bias-free tandem PEC system. When it comes to the PEC process, the electron/hole transfer/transport process has a significant impact on charge separation and, as a result, PEC performance. In photocathodes, holes are extracted and transported to the host materials. In order to improve the PEC system, it is critical to select and engineer the host materials [72]. For the first time in 2010, Nann *et al.* [85] developed gold as host material for the design of QDs sensitized PEC photocathode, on which lnP QDs and $Fe_2S_2(CO)_6$ co-catalysts were assembled. Despite the fact that these systems are inefficient in practical applications, the simple PEC design highlighted the active involvement of QDs and established the basis for the construction of QD-based photocathodes. Meng *et al.* [86] developed bi-functional molecular linker called mercaptoacetic acid (MAA) to anchor colloidal CdSe QDs on to NiO film, inorder to satisfy energy requirements for consecutive hole transfer and interact with QDs to enable ultrafast hole transfer. The state-of-the-art of QDs-based PEC photocathodes are summarized in Tab. 11.3.

Tab. 11.3: Summary of QD based photocathodes. Reprinted with the permission from reference [76].

Photocathodes	**Onset potential**	**Condition**	**Current density (μAcm^{-2})**	**Faradaic efficiency**
NiO/CdSe	0 V vs RHE	0.1 M phosphate (pH = 6.8)	-20	85%
NiO/CdSe/NiS	0 V vs Ag/AgCl	0.5 M Na_2SO_4 (pH = 6.8)	-130	95%
NiO/CdSe/CoP	0 V vs NHE	0.1 M Na_2SO_4 (pH = 6.8)	-110	81%
NiO/CdTe/NiS	-0.222 V vs Ag/AgCl	0.1 M H_2PO_4 (pH = 6.0)	-110	100%
NiO/CdSe-CdTe assembly	-0.27 V vs Ag/AgCl	0.0284 M ascorbic acid (pH = 4.56)	-140	92%

11.8 Conclusion

In this chapter, a brief description about the basic operation principles of PEC water splitting and CO_2 reduction processes with the recent progresses in the design of photoelectrodes for PEC system are given. Researchers have produced various types of semiconductors and alternative cell configurations in the previous several decades since the first report on the PEC energy conversion process, compared to the single large-bandgap TiO_2 photoelectrode at the beginning. Among various cell configurations PEC/PEC tandem cell system using earth-abundant materials are developed to achieve unbiased PEC solar-to-fuel generation. However, there many challeneges mentioned in the earlier Sec. 11.3.5 and Sec. 11.4.6 have to be addressed in order to achieve high solar energy conversion efficiency. To achieve this task development of novel photoelec-

trodes and further efficiently integrating them in PEC tandem cell are required. Good water stability, adequate band levels, excellent optical absorption in the visible region, and good electrical characteristics are all requirements for PEC photoelectrode materials. Among many n-type semiconductor materials TiO_2 is commonly studied as photoanode for water oxidation and p-type metal oxide such as NiO and Cu_2O are applied as photocathodes for CO_2 reduction and water reduction. However, metal-oxides based photoelectrodes, on the other hand, have low photocurrent densities due to their poor light absorption and charge transfer properties. In order to control the onset potential and to increase the photocurrent of photoelectrodes, its important to improve the preparation methods of photoelectrodes and to employ more effective catalysts and photoabsorbers (QDs) on photoelectrodes. Furthurmore, its important to understand the complicated mechanism in PEC CO_2 reduction on photocathode and to further commercially develop PEC tandem cell on industrial scale.

Bibliography

[1] AG Olabi and Mohammad Ali Abdelkareem. Renewable energy and climate change. *Renewable and Sustainable Energy Reviews*, 158:112111, 2022.

[2] AG Olabi, Khaled Obaideen, Khaled Elsaid, Tabbi Wilberforce, Enas Taha Sayed, Hussein M Maghrabie, and Mohammad Ali Abdelkareem. Assessment of the pre-combustion carbon capture contribution into sustainable development goals sdgs using novel indicators. *Renewable and Sustainable Energy Reviews*, 153:111710, 2022.

[3] Anam Kalair, Naeem Abas, Muhammad Shoaib Saleem, Ali Raza Kalair, and Nasrullah Khan. Role of energy storage systems in energy transition from fossil fuels to renewables. *Energy Storage*, 3(1):e135, 2021.

[4] Somnath C Roy, Oomman K Varghese, Maggie Paulose, and Craig A Grimes. Toward solar fuels: photocatalytic conversion of carbon dioxide to hydrocarbons. *ACS nano*, 4(3):1259–1278, 2010.

[5] No new oil, gas or coal development if world is to reach net zero by 2050, says world energy body. www.theguardian.com/environment/2021/may/18/no-new-investment-in-fossil-fuels-demands-top-energy-economist, The Guardian.

[6] Ralph L House, Neyde Yukie Murakami Iha, Rodolfo L Coppo, Leila Alibabaei, Benjamin D Sherman, Peng Kang, M Kyle

Brennaman, Paul G Hoertz, and Thomas J Meyer. Artificial photosynthesis: Where are we now? where can we go? *Journal of Photochemistry and Photobiology C: Photochemistry Reviews*, 25:32–45, 2015.

[7] Dolf Gielen, Francisco Boshell, Deger Saygin, Morgan D Bazilian, Nicholas Wagner, and Ricardo Gorini. The role of renewable energy in the global energy transformation. *Energy strategy reviews*, 24:38–50, 2019.

[8] Isolda Roger, Michael A Shipman, and Mark D Symes. Earth-abundant catalysts for electrochemical and photoelectrochemical water splitting. *Nature Reviews Chemistry*, 1(1):1–13, 2017.

[9] Niklas V von der Assen, Ana Maria Lorente Lafuente, Martina Peters, and André Bardow. Environmental assessment of co2 capture and utilisation. In *Carbon dioxide utilisation*, pages 45–56. Elsevier, 2015.

[10] Phong D Tran, Lydia H Wong, James Barber, and Joachim SC Loo. Recent advances in hybrid photocatalysts for solar fuel production. *Energy & Environmental Science*, 5(3):5902–5918, 2012.

[11] Daniel G Nocera. Solar fuels and solar chemicals industry. *Accounts of chemical research*, 50(3):616–619, 2017.

[12] Marcos de Oliveira Melo and Luciana Almeida Silva. Photocatalytic production of hydrogen: an innovative use for biomass derivatives. *Journal of the Brazilian Chemical Society*, 22:1399–1406, 2011.

[13] Hong Pang, Takuya Masuda, and Jinhua Ye. Semiconductor-based photoelectrochemical conversion of carbon dioxide: Stepping towards artificial photosynthesis. *Chemistry–An Asian Journal*, 13(2):127–142, 2018.

[14] Fei Cao, Gerko Oskam, and Peter C Searson. A solid state, dye sensitized photoelectrochemical cell. *The journal of physical chemistry*, 99(47):17071–17073, 1995.

[15] Arthur J Nozik. Photoelectrochemical devices for solar energy conversion. In *Photovoltaic and Photoelectrochemical Solar Energy Conversion*, pages 263–312. Springer, 1981.

[16] Bruce Parkinson and John Turner. The potential contribution of photoelectrochemistry in the global energy future. 2013.

[17] H Gerischer. Electrochemical behavior of semiconductors under illumination. *Journal of the electrochemical society*, 113(11):1174, 1966.

[18] Michael Grätzel. *Photoelectrochemical cells*, pages 26–32.

[19] Akira Fujishima and Kenichi Honda. Electrochemical evidence for the mechanism of the primary stage of photosynthesis. *Bulletin of the chemical society of Japan*, 44(4):1148–1150, 1971.

[20] Akira Fujishima and Kenichi Honda. Electrochemical photolysis of water at a semiconductor electrode. *nature*, 238(5358):37–38, 1972.

[21] Jian Zhao, Xin Wang, Zhichuan Xu, and Joachim SC Loo. Hybrid catalysts for photoelectrochemical reduction of carbon dioxide: a prospective review on semiconductor/metal complex co-catalyst systems. *Journal of Materials Chemistry A*, 2(37):15228–15233, 2014.

[22] Di Wei and Gehan Amaratunga. Photoelectrochemical cell and its applications in optoelectronics. *International Journal of Electrochemical Science*, 2(12):897–912, 2007.

[23] LJ Paterno. 1–nanomaterials: Solar energy conversion. *Nanoscience and its Applications*, pages 1–33, 2017.

[24] A Hamnett and S Dennison. Bright future for liquid electrolyte solar cells? *Nature*, 300(5894):687–688, 1982.

[25] Maheshwar Sharon. The photoelectrochemistry of semiconductor/electrolyte solar cells. *Encyclopedia of Electrochemistry: Online*, 2007.

[26] W Kreuter and HJIJoHE Hofmann. Electrolysis: the important energy transformer in a world of sustainable energy. *International Journal of Hydrogen Energy*, 23(8):661–666, 1998.

[27] Antonio Vilanova, Tania Lopes, Carsten Spenke, Michael Wullenkord, and Adelio Mendes. Optimized photoelectrochemical tandem cell for solar water splitting. *Energy Storage Materials*, 13:175–188, 2018.

[28] KGU Wijayantha and DH Auty. Twin cell technology for hydrogen generation. 2005.

[29] Yaping Zhang, Yuyu Bu, Lin Wang, and Jin-Ping Ao. Regulation of the photogenerated carrier transfer process during photoelectrochemical water splitting: A review. *Green Energy & Environment*, 6(4):479–495, 2021.

[30] Kan Zhang, Ming Ma, Ping Li, Dong Hwan Wang, and Jong Hyeok Park. Water splitting progress in tandem devices: moving photolysis beyond electrolysis. *Advanced Energy Materials*, 6(15):1600602, 2016.

[31] Kevin Sivula and Michael Grätzel. Tandem photoelectrochemical cells for water splitting. 2013.

[32] Quan Gu, Ziwei Gao, Sijia Yu, and Can Xue. Constructing ru/tio2 heteronanostructures toward enhanced photocatalytic water splitting via a ruo2/tio2 heterojunction and ru/tio2 schottky junction. *Advanced Materials Interfaces*, 3(4):1500631, 2016.

[33] Yerkin Shabdan, Aiymkul Markhabayeva, Nurlan Bakranov, and Nurxat Nuraje. Photoactive tungsten-oxide nanomaterials for water-splitting. *Nanomaterials*, 10(9):1871, 2020.

[34] Guru Dayal. Current trends and future roadmap for solar fuels. *Advances in Solar Energy Research*, pages 445–484, 2019.

[35] None Gurudayal, Joel Ager, and Nripan Mathews. Perovskite solar cells for photoelectrochemical water splitting and

co2 reduction. Technical report, Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), 2018.

[36] Qi Chen, Guozheng Fan, Hongwei Fu, Zhaosheng Li, and Zhigang Zou. Tandem photoelectrochemical cells for solar water splitting. *Advances in Physics: X*, 3(1):1487267, 2018.

[37] Yudai Kawase, Tomohiro Higashi, Kazunari Domen, and Kazuhiro Takanabe. Recent developments in visible-light-absorbing semitransparent photoanodes for tandem cells driving solar water splitting. *Advanced Energy and Sustainability Research*, 2(7):2100023, 2021.

[38] Zichen Gong, Bin Liu, Peng Zhang, Shujie Wang, Shaokun Jiang, Tuo Wang, and Jinlong Gong. Performance prediction of multiple photoanodes systems for unbiased photoelectrochemical water splitting. *ACS Materials Letters*, 3(7):939–946, 2021.

[39] H Yoneyama, H Sakamoto, and H Tamura. A photoelectrochemical cell with production of hydrogen and oxygen by a cell reaction. *Electrochimica Acta*, 20(5):341–345, 1975.

[40] Dongqi Xue, Miao Kan, Xufang Qian, and Yixin Zhao. A tandem water splitting cell based on nanoporous bivo4 photoanode cocatalyzed by ultrasmall cobalt borate sandwiched with conformal tio2 layers. *ACS sustainable chemistry & engineering*, 6(12):16228–16234, 2018.

[41] Xiang Yin, Qiong Liu, Yahui Yang, Yang Liu, Keke Wang, Yaomin Li, Dongwei Li, Xiaoqing Qiu, Wenzhang Li, and Jie Li. An efficient tandem photoelectrochemical cell composed of feooh/tio2/bivo4 and cu2o for self-driven solar water splitting. *International journal of hydrogen energy*, 44(2):594–604, 2019.

[42] Ch Leygraf, M Hendewerk, and Gabor A Somorjai. Photocatalytic production of hydrogen from water by a p-and n-type polycrystalline iron oxide assembly. *The Journal of Physical Chemistry*, 86(23):4484–4485, 1982.

[43] Gopal K Mor, Oomman K Varghese, Rudeger HT Wilke, Sanjeev Sharma, Karthik Shankar, Thomas J Latempa, Kyoung-Shin Choi, and Craig A Grimes. p-type cu- ti- o nanotube arrays and their use in self-biased heterojunction photoelectrochemical diodes for hydrogen generation. *Nano Letters*, 8(7):1906–1911, 2008.

[44] Gyan Prakash, Sharma Arun, Prakash Upadhyay, Dilip Kumar Behara, Sri Sivakumar, and Raj Ganesh S Pala. Fundamental aspect of photoelectrochemical water splitting. In *The Water–Food–Energy Nexus: Processes, Technologies, and Challenges*, pages 667–690. CRC Press, 2017.

[45] Anupma Thakur, Dibyendu Ghosh, Pooja Devi, Ki-Hyun Kim, and Praveen Kumar. Current progress and challenges in photoelectrode materials for the production of hydrogen. *Chemical Engineering Journal*, 397:125415, 2020.

[46] Chaoran Jiang, Savio JA Moniz, Aiqin Wang, Tao Zhang, and Junwang Tang. Photoelectrochemical devices for solar water splitting–materials and challenges. *Chemical Society Reviews*, 46(15):4645–4660, 2017.

[47] Bhupendra Kumar, Mark Llorente, Jesse Froehlich, Tram Dang, Aaron Sathrum, and Clifford P Kubiak. Photochemical and photoelectrochemical reduction of co2. *Annual review of physical chemistry*, 63:541–569, 2012.

[48] P Ding, T Jiang, N Han, and Y Li. Photocathode engineering for efficient photoelectrochemical co2 reduction. *Materials Today Nano*, 10:100077, 2020.

[49] M Halmann. Photoelectrochemical reduction of aqueous carbon dioxide on p-type gallium phosphide in liquid junction solar cells. *Nature*, 275(5676):115–116, 1978.

[50] Xiaoxia Chang, Tuo Wang, and Jinlong Gong. Co 2 photoreduction: insights into co 2 activation and reaction on surfaces of photocatalysts. *Energy & Environmental Science*, 9(7):2177–2196, 2016.

[51] Ning Zhang, Ran Long, Chao Gao, and Yujie Xiong. Recent progress on advanced design for photoelectrochemical reduction of co2 to fuels. *Sci. China Mater*, 61(6):771–805, 2018.

[52] Zhonghui Zhu, Xianglei Liu, Chuang Bao, Kai Zhang, Chao Song, and Yimin Xuan. How efficient could photocatalytic co2 reduction with h2o into solar fuels be? *Energy conversion and management*, 222:113236, 2020.

[53] Evangelos Kalamaras, Huizhi Wang, M Mercedes Maroto-Valer, John M Andresen, and Jin Xuan. Theoretical efficiency limits of photoelectrochemical co2 reduction: A route-dependent thermodynamic analysis. *ChemPhysChem*, 21(3):232–239, 2020.

[54] Vignesh Kumaravel, John Bartlett, and Suresh C Pillai. Photoelectrochemical conversion of carbon dioxide (co2) into fuels and value-added products. *ACS Energy Letters*, 5(2):486–519, 2020.

[55] Emanuele Moioli, Robin Mutschler, and Andreas Züttel. Renewable energy storage via co2 and h2 conversion to methane and methanol: Assessment for small scale applications. *Renewable and Sustainable Energy Reviews*, 107:497–506, 2019.

[56] Yang Yang, Saira Ajmal, Xiuzhen Zheng, and Liwu Zhang. Efficient nanomaterials for harvesting clean fuels from electrochemical and photoelectrochemical co 2 reduction. *Sustainable Energy & Fuels*, 2(3):510–537, 2018.

[57] Penglu Wang, Songcan Wang, Haiqiang Wang, Zhongbiao Wu, and Lianzhou Wang. Recent progress on photo-electrocatalytic reduction of carbon dioxide. *Particle & Particle Systems Characterization*, 35(1):1700371, 2018.

[58] Shunsuke Sato, Takeo Arai, Takeshi Morikawa, Keiko Uemura, Tomiko M Suzuki, Hiromitsu Tanaka, and Tsutomu Kajino. Selective co2 conversion to formate conjugated with h2o oxidation utilizing semiconductor/complex hybrid photocatalysts.

Journal of the American Chemical Society, 133(39):15240–15243, 2011.

[59] Takeo Arai, Shunsuke Sato, Tsutomu Kajino, and Takeshi Morikawa. Solar co 2 reduction using h2o by a semiconductor/metal-complex hybrid photocatalyst: enhanced efficiency and demonstration of a wireless system using srtio 3 photoanodes. *Energy & Environmental Science*, 6(4):1274–1282, 2013.

[60] Haiying He, Peter Zapol, and Larry A Curtiss. A theoretical study of co2 anions on anatase (101) surface. *The Journal of Physical Chemistry C*, 114(49):21474–21481, 2010.

[61] Qi Liu, Di Wu, Yong Zhou, Haibin Su, R Wang, Chunfeng Zhang, Shicheng Yan, Min Xiao, and Zhigang Zou. Single-crystalline, ultrathin znga2o4 nanosheet scaffolds to promote photocatalytic activity in co2 reduction into methane. *ACS applied materials & interfaces*, 6(4):2356–2361, 2014.

[62] Masakazu Anpo, Hiromi Yamashita, Yuichi Ichihashi, Yo Fujii, and Miwa Honda. Photocatalytic reduction of co2 with h2o on titanium oxides anchored within micropores of zeolites: effects of the structure of the active sites and the addition of pt. *The Journal of Physical Chemistry B*, 101(14):2632–2636, 1997.

[63] Junseok Lee, Dan C Sorescu, and Xingyi Deng. Electron-induced dissociation of co2 on tio2 (110). *Journal of the American Chemical Society*, 133(26):10066–10069, 2011.

[64] Xiaoxia Chang, Tuo Wang, Piaoping Yang, Gong Zhang, and Jinlong Gong. The development of cocatalysts for photoelectrochemical co2 reduction. *Advanced Materials*, 31(31):1804710, 2019.

[65] Yu Hui Lui, Bowei Zhang, and Shan Hu. Rational design of photoelectrodes for photoelectrochemical water splitting and co2 reduction. *Frontiers of Physics*, 14(5):53402, 2019.

[66] Gopal K Mor, Karthik Shankar, Maggie Paulose, Oomman K Varghese, and Craig A Grimes. Enhanced photocleavage of water using titania nanotube arrays. *Nano letters*, 5(1):191–195, 2005.

[67] Jin Un Kim, Hyun Soo Han, Joonsuk Park, Woosung Park, Ji Hyun Baek, Jae Myeong Lee, Hyun Suk Jung, and In Sun Cho. Facile and controllable surface-functionalization of tio2 nanotubes array for highly-efficient photoelectrochemical water-oxidation. *Journal of catalysis*, 365:138–144, 2018.

[68] Teera Butburee, Yang Bai, Huanjun Wang, Hongjun Chen, Zhiliang Wang, Gang Liu, Jin Zou, Pongtanawat Khemthong, Gao Qing Max Lu, and Lianzhou Wang. 2d porous tio2 single-crystalline nanostructure demonstrating high photoelectrochemical water splitting performance. *Advanced Materials*, 30(21):1705666, 2018.

[69] Palyam Subramanyam, Bhagatram Meena, Duvvuri Suryakala, and Challapalli Subrahmanyam. Tio2 photoanodes sensitized with bi2se3 nanoflowers for visible–near-infrared photoelectrochemical water splitting. *ACS Applied Nano Materials*, 4(1):739–745, 2021.

[70] S David Tilley, Maurin Cornuz, Kevin Sivula, and Michael Grätzel. Light-induced water splitting with hematite: improved nanostructure and iridium oxide catalysis. *Angewandte Chemie International Edition*, 49(36):6405–6408, 2010.

[71] Jae Young Kim, Ganesan Magesh, Duck Hyun Youn, Ji-Wook Jang, Jun Kubota, Kazunari Domen, and Jae Sung Lee. Single-crystalline, wormlike hematite photoanodes for efficient solar water splitting. *Scientific reports*, 3(1):2681, 2013.

[72] Fatwa F Abdi, Lihao Han, Arno HM Smets, Miro Zeman, Bernard Dam, and Roel Van De Krol. Efficient solar water splitting by enhanced charge separation in a bismuth vanadate-silicon tandem photoelectrode. *Nature communications*, 4(1):2195, 2013.

[73] Il Song Park, Min Ho Lee, Tae Sung Bae, and Kyeong Won Seol. Effects of anodic oxidation parameters on a modified titanium surface. *Journal of Biomedical Materials Research Part B: Applied Biomaterials: An Official Journal of The Society for Biomaterials, The Japanese Society for Biomaterials, and The Australian Society for Biomaterials and the Korean Society for Biomaterials*, 84(2):422–429, 2008.

[74] Hyeonseok Yoo, Moonsu Kim, Yong-Tae Kim, Kiyoung Lee, and Jinsub Choi. Catalyst-doped anodic tio2 nanotubes: binder-free electrodes for (photo) electrochemical reactions. *Catalysts*, 8(11):555, 2018.

[75] MV Diamanti, M Pedeferri, et al. The anodic oxidation of titanium and its alloys. In *Encyclopedia of Interfacial Chemistry: Surface Science and Electrochemistry*, pages 41–54. Elsevier, 2018.

[76] Lei Jin, Haiguang Zhao, Zhiming M Wang, and Federico Rosei. Quantum dots-based photoelectrochemical hydrogen evolution from water splitting. *Advanced Energy Materials*, 11(12):2003233, 2021.

[77] Tea-Yon Kim, Byung Su Kim, Jong Gyu Oh, Seul Chan Park, Jaeyoung Jang, Thomas W Hamann, Young Soo Kang, Jin Ho Bang, Sixto Giménez, and Yong Soo Kang. Interfacial engineering at quantum dot-sensitized tio2 photoelectrodes for ultrahigh photocurrent generation. *ACS Applied Materials & Interfaces*, 13(5):6208–6218, 2021.

[78] Hong Bin Yang, Jianwei Miao, Sung-Fu Hung, Fengwei Huo, Hao Ming Chen, and Bin Liu. Stable quantum dot photoelectrolysis cell for unassisted visible light solar water splitting. *ACS nano*, 8(10):10403–10413, 2014.

[79] Joseph S DuChene, Giulia Tagliabue, Alex J Welch, Xueqian Li, Wen-Hui Cheng, and Harry A Atwater. Optical excitation of a nanoparticle cu/p-nio photocathode improves reaction selectivity for co2 reduction in aqueous electrolytes. *Nano letters*, 20(4):2348–2358, 2020.

[80] Chenyan Hu, Kenneth Chu, Yihua Zhao, and Wey Yang Teoh. Efficient photoelectrochemical water splitting over anodized p-type nio porous films. *ACS Applied Materials & Interfaces*, 6(21):18558–18568, 2014.

[81] Kaijian Zhu, Sean K Frehan, Anna M Jaros, Devin B O'Neill, Jeroen P Korterik, Kasper Wenderich, Guido Mul, and Annemarie Huijser. Unraveling the mechanisms of beneficial cu-doping of nio-based photocathodes. *The Journal of Physical Chemistry C*, 125(29):16049–16058, 2021.

[82] SocMan Ho-Kimura, Savio JA Moniz, Junwang Tang, and Ivan P Parkin. A method for synthesis of renewable cu2o junction composite electrodes and their photoelectrochemical properties. *ACS Sustainable Chemistry & Engineering*, 3(4):710–717, 2015.

[83] Adriana Paracchino, Nripan Mathews, Takashi Hisatomi, Morgan Stefik, S David Tilley, and Michael Grätzel. Ultrathin films on copper (i) oxide water splitting photocathodes: a study on performance and stability. *Energy & Environmental Science*, 5(9):8673–8681, 2012.

[84] Keke Wang, Yanfang Ma, Yang Liu, Weixin Qiu, Qingmei Wang, Xuetao Yang, Min Liu, Xiaoqing Qiu, Wenzhang Li, and Jie Li. Insights into the development of cu-based photocathodes for carbon dioxide (co 2) conversion. *Green Chemistry*, 23(9):3207–3240, 2021.

[85] Thomas Nann, Saad K Ibrahim, Pei-Meng Woi, Shu Xu, Jan Ziegler, and Christopher J Pickett. Water splitting by visible light: a nanophotocathode for hydrogen production. *Angew. Chem. Int. Ed*, 49(9):1574–1577, 2010.

[86] Bin Liu, Xu-Bing Li, Yu-Ji Gao, Zhi-Jun Li, Qing-Yuan Meng, Chen-Ho Tung, and Li-Zhu Wu. A solution-processed, mercaptoacetic acid-engineered cdse quantum dot photocathode for efficient hydrogen production under visible light irradiation. *Energy & Environmental Science*, 8(5):1443–1449, 2015.

Rozdział 12

Przemysł garbarski – producent materiału luksusowego i źródło globalnych problemów z odpadami produkcyjnymi po procesie przetwarzania skór bydlęcych

Paulina Bandrów

Institute of Fluid Flow Machinery, Polish Academy of Sciences, Fiszera 14, 80-231 Gdansk, Poland

12.1 Wstęp literaturowy

12.1.1 Zarys historyczny

Skóra zwierzęca od wieków służyła człowiekowi i była przez niego wykorzystywana w codziennym życiu. Już w czasach prehistorycznych używana do ochrony ciała, wymusiła na człowieczeństwie potrzebę zabezpieczenia tego surowca przed gniciem i rozkładem. Szukanie nowych sposobów i ulepszanie starych spowodował silny rozwój dziedziny „garbarstwa" – technologicznego procesu, który prowadzi do zabezpieczenia surowej skóry przed gniciem powodując wzrost jej odporności termicznej i właściwości wytrzymałościowych. Początkowo przetwarzanie skóry opierało się jedynie na jej suszeniu. Jednak skóra była wtedy twarda i łamliwa. Z czasem rozpoczęto, podczas procesu ugniatania, dodawanie tłuszczy zwierzęcych, które pozwoliły na uzyskanie odpowiedniej miękkości. Tak powstała pierwsza metoda garbowania – garbowanie tłuszczowe. Następnym krokiem stało się wykorzystywanie kory czy liści jako garbników – garbowanie roślinne. Wciąż jednak skóry garbowane pokryte były włosem. Dopiero po pewnym czasie zauważono, że proces gnicia zachodzący na skórze powoduje usuwanie z niej włosa. Skóra bez włosa stała się bardziej użyteczna, co umożliwiło produkcję większej liczby przydatnych artykułów. Pierwsze zapisy o garbowaniu i barwieniu skór pochodzą sprzed 4500 lat z papirusów egipskich. Później spotykane są informacje o europejskich próbach przetwarzania skór, które spowodowały silny rozwój tej gałęzi przemysłu i chęci poszukiwania coraz to lepszych metod, a tym samym coraz szerszym wykorzystaniu skór zwierzęcych. Rozwój chemii w XVIII i XIX wieku spowodował wiele prób, aby wykorzystać tę naukę również w garbarstwie. Poszukiwanie nowych, efektywniejszych garbników doprowadziło do powstania tzw. garbowania chromowego, którego twórcą był F. Knapp. Lata 20. były przełomowym okresem. Naukowcy z całego świata angażowali się w poszukiwanie i ulepszanie procesu garbowania skór. Skóry zamorskie sprowadzane do Europy wymagały innych sposobów na ich przetwarzanie co sprzyjało rozwojowi. Już nie tylko przemysł odzieżowy czy obuwniczy wywierał zapotrzebowanie na skórę, ale również tapicerski czy samochodowy [1].

12.1.2 Wprowadzenie

Garbowanie jak zostało wyjaśnione powyżej jest procesem przetwarzania skór surowych, ulegających szybkiemu rozkładowi i gniciu, w materiał trwały oraz łatwy do dalszego przerobu, o odpowiednich właściwościach fizycznych. Składa się z szeregu złożonych reakcji chemicznych nadających wysoką trwałość oraz wyjątkowo specyficzny charakter. Dodatkowo skóra przetwarzana ulega wielu mechanicznym procesom właściwego przygotowania a także wykończenia [2]. Dzięki tak rozwiniętym metodom przerobu skór możliwe jest uzyskanie końcowego produktu o wysokiej jakości i właściwościach jak na przykład, trwałość, wygląd, odporność na temperaturę czy wodę, miękkość, ścisłość, przepuszczalność potu jak i powietrza [3].

Naturalna skóra jest najczęściej wyjściowym materiałem służącym do wytwarzania wielu niezbędnych produktów (odzież, obuwie) jak i produktów luksusowych, mając na celu podwyższenie wartości przedmiotu czy jego poprawę jakościową i wizualna (tapicerka samochodowa, meble).

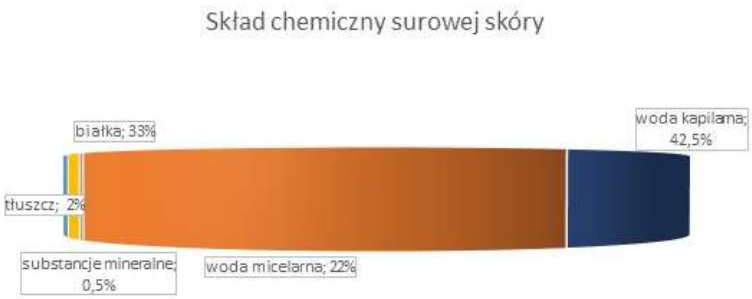
Przetwarzanie naturalnego surowca jakim jest skóra surowa do wysokiej jakości wyrobu gotowego w postaci siedzeń samochodowych, galanterii czy obuwia stanowi ogromne wyzwanie jakościowe, technologiczne i środowiskowe. Powodem tego jest podatność skóry na gnicie i uszkodzenia, nie tylko w procesie przetwarzania, ale również już podczas hodowli zwierząt czy ich uboju [4].

Skuteczność osiągnięcia celów jakościowych i środowiskowych przy produkcji skóry zależy od jej rodzaju i zawartości związków chemicznych. Skład surowej skóry zależy od wielu czynników, m.in. od rasy, płci czy gatunku zwierzęcia. Można jednak określić przeciętny chemiczny skład, taki też został przedstawiony na Rys. 12.1.

Białka występujące w surowej skórze dzielą się na włókniste (strukturalne) oraz bezpostaciowe (globularne). Wśród białek włóknistych można wymienić kolagen, elastynę i keratynę. Zawartość kolagenu rośnie równocześnie z ubytkiem wody podczas chromowego garbowania. W przypadku garbowania chromowego zawartość kolagenu może wzrosnąć nawet do 72%, natomiast podczas garbowania roślinnego tylko do 41%, czego powodem jest wprowadzanie dużych

ilości roślinnych garbników. Białka bezpostaciowe powinny zostać usunięte ze skóry podczas procesów przygotowawczych.

Woda kapilarna jest tą, która utrzymuje się siłami włoskowatości pomiędzy włóknami. Usunięcie jej jest procesem odwracalnym. Woda micelarna (1/3 całkowitej wody) jest chemicznie związana z treścią komórki- wewnątrzkomórkowa. Po jej usunięciu nie ma możliwości odwrócenia tego procesu. Skóry, w których pojawił się ubytek wody micelarnej w wyniku suchej konserwacji lub intensywnego suszenia, są gorszej jakości i bardzo trudno je rozmoczyć. Tłuszcze i substancje mineralne występują w niewielkich ilościach. Najbogatsza w tłuszcze jest warstwa podskórna. Skład tłuszczu skórnego zwierzęcia przedstawia się następująco: C – 76,5%, O – 11,5%, H – 12%. Substancje mineralne nie wpływają na jakość skóry jako garbarskiego surowca [1].



Rys. 12.1: Skład chemiczny surowej skóry.

Silny wzrost zapotrzebowania na skóry spowodował powstawaniem potężnych liczby produktów ubocznych, które stało się problemem zakładów garbarskich i przemysłu skórzanego. Część z nich znalazła jednak zastosowanie w innych gałęziach przemysłu, takich jak wytwarzanie karmy dla zwierząt, poprawa właściwości gleby czy przemysł chemiczny. Dodatkowo tego typu zakłady określane są jako *truciciele*, nie tylko z powodu generowania odpadów. Powodują wydzielanie się nieprzyjemnego zapachu oraz ogromną produkcję ścieków, zawierających różnego typu związki jak choćby chrom. Nieodpowiednia utylizacja odpadów może powodować silne zanieczyszczenie środowiska [5].

Podczas przerobu skóry surowej o wadze 1 Mg powstaje jedynie około 200 kg produktu, pozostałe 800 kg stanowi odpad pochodzący ze skór niegarbowanych, garbowanych a także osad pochodzący z oczyszczalni ścieków [6].

Współczesne metody utylizacji tego typu odpadu nie są wystarczające, dodatkowo generują wysokie koszty. Powoduje to, że podstawowym sposobem na ich zagospodarowanie jest po prostu składowanie [7]. Niestety tego typu rozwiązanie jest bardzo niebezpieczne, ponieważ produkty rozpadu mogą przenikać do gleby, wód a także wydzielać się do atmosfery. Dodatkowo istnieje ryzyko samozapłonu, co zwiększy zanieczyszczenie środowiska [8].

Cały proces przetwarzania skór od lat ulega niewielkim zmianom, pomimo tego wprowadzono pewne udoskonalenia w przypadku ochrony środowiska dla tej gałęzi przemysłu. Podstawowymi elementami mającymi na celu minimalizację negatywnego wpływu na wodę, powietrze czy glebę są zmiany przepisów o kontroli zanieczyszczeń, jak największa minimalizacja powstających odpadów czy ich racjonalne unieszkodliwianie a także rodzaj stosowanych chemikaliów. Bardzo istotna jest analiza ryzyka oraz związane z tym rozwiązania pozwalające w razie ewentualnej awarii ograniczyć niebezpieczny wpływ dla środowiska.

Ocena rozwiązań mających służyć lepszej ochronie środowiska powinna być przeprowadzana z perspektywy kosztów i ogólnych korzyści. Jakakolwiek zmiana techniki może wpływać na pozostałe moduły technologiczne. Stąd też powinny być wybierane techniki, które pozwolą osiągnąć w procesie jak najmniejszy wpływ na środowisko naturalne jako całości. Pod względem tych kryteriów oceniane zostają najlepsze dostępne techniki, mogą one więc dotyczyć zmian ograniczania zanieczyszczeń w modułach procesowych jak również na końcu łańcucha.

Równie istotne stają się regularne przeglądy maszyn i procesu jak i ich odpowiednia eksploatacja, w tym także odpowiednie szkolenie pracowników [9].

12.1.3 Produkcja skór w Europie i na świecie

Zarówno w Europie jak i na świecie całkowita produkcja skór uzależniona jest od kilku aspektów. W dużej mierze wpływ ma wielkość populacji zwierząt, konsumpcja mięsa a tym samym tempo ich uboju. Jedne z największych skupisk bydła występują w Stanach Zjednoczonych, Chinach, Argentynie, Brazylii, Indiach, Unii Europejskiej czy w Rosji.

Różnice w jakości skór są widoczne pomiędzy bydłem pochodzących z Europy i tym z systemów ranczerskich i pasterskich, gdzie zwierzęta są znacznie bardziej narażone na uszkodzenia skóry czy oznaczane za pomocą wypaleń.

Powoduje to, że możliwe jest uzyskanie znacznie większej powierzchni użytkowej ze skór bydła europejskiego, a tym samym ograniczenie ilość odpadu. Dodatkowo skóry te są wolne od zakazanych pestycydów [8].

Na przestrzeni lat obserwuje się wzrost importu skór do Unii Europejskiej jednak nadal zakłady garbarskie mają problemy ze sprowadzaniem wystarczającej ilości surowca również z krajów trzeciego świata. Jedną z największych barier handlowych stanowią cła, a także pewne ograniczenia wywozu surowych skór i półproduktu *wet white/ wet blue*. Na wolny rynek międzynarodowy trafia tylko średnio ok. 40% dostępnych na skalę światową skór. Powoduje to wysokie i niestabilne ceny surowca, a także jego niedobór [4].

Dane pokazują, że w krajach rozwijających się, pomimo iż posiadają 78% populacji światowej bydła, tylko 64% skór jest przetwarzanych dalej, z czego odnosząc to do wagi wynosi jedynie 57% [8].

Handel skórami bydlęcymi na świecie ulega silnej transformacji i tak też kraje należące do rozwijających się z eksporterów netto coraz częściej zostają importerami netto. Powodem jest wzrost możliwości produkcyjnych garbarni na Dalekim Wschodzie czy też w Ameryce Łacińskiej.

Import skór bydlęcych na początku bieżącego wieku znacznie zmalał w porównaniu do poprzedniego stulecia. W przeciągu 7 lat spadek wyniósł prawie 65%, z 353 575 ton w roku 2000 na 124 000 ton w 2007 roku [8].

Tab. 12.1: Główni dostawcy skóry dla Unii Europejskiej. [Źródło: Eurostat 2012].

	Udział (%)					**Wzrost (%)**
	2006	**2007**	**2008**	**2009**	**2010**	**[2006-2010]**
USA	17,7	15,9	17,6	15,1	17,4	-13,4
Szwajcaria	6,9	6,8	7,8	8,7	9	14,9
Bośnia i Hercegowina	7,8	7	6,8	6	8,5	-3,4
Iran	6,8	9,1	11,1	11,9	8,4	9,9
RPA	6	7,5	8,2	8,1	7,1	4,2
Australia	7,1	7,8	7,3	7,6	6,7	-17,2
Nowa Zelandia	8,9	6,7	8,2	9,2	6,6	-35,2
Norwegia	3,1	2,8	3,8	3,1	3,7	5,1
Białoruś	0,8	0,6	0,2	2	3,2	232,6
Serbia	1,3	1,3	1,3	2	2,6	79,9
Razem	100	100	100	100	100	-12

Tab. 12.1 zawiera zestawienie głównych dostawców skór, z której wynika, że największym dostawcą skóry surowej do Unii Europejskiej wciąż pozostają Stany Zjednoczone, ich całkowity udział importu spoza UE w dostarczaniu surowca wynosi ponad 15%. Bilans całkowity dla dostawców skór dla UE w latach 2006-2010 utrzymuje się na poziomie -12%.

Unia Europejska nie jest kluczowym producentem skór na świecie, inne główne centra mają miejsce w Meksyku, Brazylii, Argentynie, Korei Płd., Pakistanie, Chinach i Indiach.

Porównując Azję czy obie Ameryki do Unii Europejskiej łatwo da się zauważyć, iż na światowym rynku udział Europy ma wyraźną tendencję do zmniejszania się, zaś w przypadku pozostałych dwóch obszarów wzrasta.

Obserwując przetwórstwo skór surowych na świecie opartych na mokrym zasoleniu, gdzie całkowita produkcja wyniosła 6,0 mln ton, widać, że możliwe jest uzyskanie 522 600 ton skóry grubej, a także 1 185 mln m^2 skóry lekkie, w którą wliczamy dwoiny. Dla Europy liczby te wynoszą 71 700 ton skóry grubej i 230 mln m^2 skóry lekkiej.

Tab. 12.2: Najważniejsze dla garbarzy europejskich rynki zbytu wykończonych skór [Źródło: 164, Eurostat 2012].

	2006	**2007**	**2008**	**2009**	**2010**	**Wzrost (%) [2006-2010]**
Hong Kong	25,7	22,5	21,5	21,8	22,6	-16,8
Chiny	11,1	13,5	12,8	14,5	14,5	23,6
Tunezja	5,7	6,6	7,4	8,3	7,4	23,5
USA	10,6	8,7	8,2	6	6,8	-39,2
Chorwacja	2,3	5	5,9	5,3	3,9	59
Indie	3,4	3,5	3,8	4,1	3,8	8,2
Turcja	4,8	4,5	3,3	2,9	3,2	-37,7
Wietnam	1,9	2,2	2,7	3	3,2	60,5
Korea Płd.	3	2,4	2,2	3	2,9	-7,6
Maroko	2	2	2,4	3	2,9	40,4
Razem	100	100	100	100	100	-5,3

Unia Europejska wciąż pozostaje jednym z największych dostawców skór na świecie na rynek międzynarodowy. 15% światowej produkcji skór i 60% produkcji europejskiej przypada na Włochy. W Tab. 12.2. przedstawione zostały najważniejsze rynki przemysłu skórzanego krajów, do których Europa eksportuje skóry wykończone. W przeciągu czterech lat sytuacja wielokrotnie się zmieniała, jednak to państwa wschodnie przeważają w odbiorze skór. Całkowity bilans wzrostu zbytu skór dla lat 2006-2010 utrzymuje się na -5,3% [8].

12.1.4 Unia Europejska – rozmieszczenie przemysłu skórzanego

Włochy odgrywają najważniejszą rolę w europejskim przemyśle skórzanym. Ich pozycja uwarunkowana jest wielkością zatrudnienia, liczbą przedsiębiorstw, wielkością obrotu, sprzedaży oraz produkcji. Na drugim miejscu znajduje się Hiszpania i odpowiada wraz z Niemcami, Wielką Brytanią, Francją i Portugalią za większą część pozostałego rynku skórzanego w Europie. We Włoszech działało 1415 garbarni, chociaż jeżeli chodzi o ich wielkość to są znacznie

mniejsze w porównaniu do pozostałych krajów, Hiszpania posiada ich około 140, Francja 62, a w Niemczech znajdziemy ich 30. Kraje skandynawskie choć w przeszłości posiadały silnie rozwiniętą tę gałąź przemysłu specjalizującą się w produkcji skór, w obecnym czasie posiadają jedynie kilka garbarni.

Garbarnie bardzo często występują w pewnych skupiskach, a w rejonach gdzie te zakłady zostały zlokalizowane, życie społeczności zależy od tej gałęzi przemysłu. Taka sytuacja występuje we Włoszech. Tab. 12.3 prezentuje największe klastry przemysłu skórzane we Włoszech [8].

Tab. 12.3: Liczba garbarni – region we Włoszech.

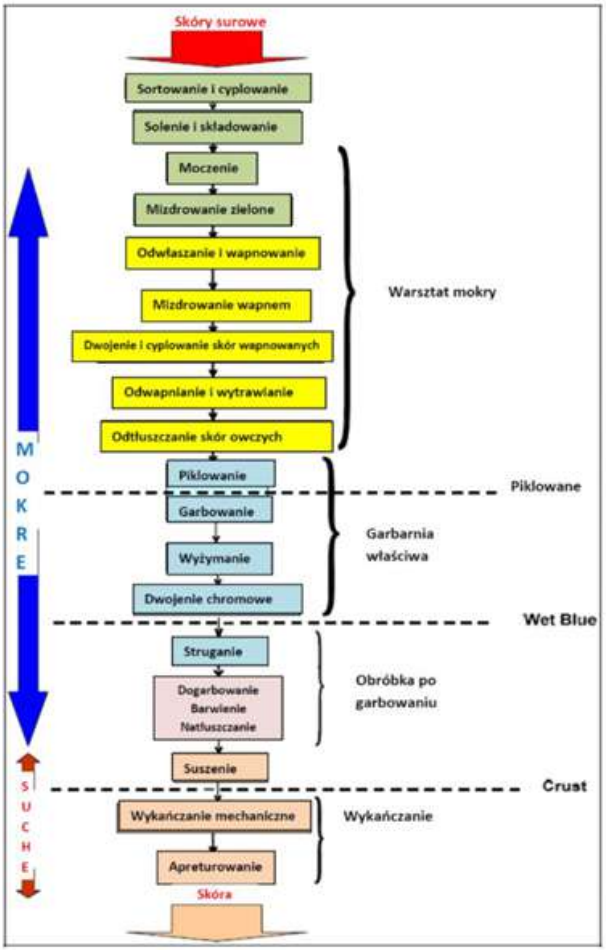
Region	**Miejscowość/ Gmina**	**Liczba garbarni**
Toskania	Santa Croce sull'Arno Ponte Egol	615
Vicenza	Arzignano Zermeghed Montebello Vicentino	465
Kampania	Neapol	193
Lombardia	Turbigo Castano Primo	70

Santa Croce i Arzignano należą do głównych ośrodków specjalizujących się w przetwarzaniu skór bydlęcych. Większość produkowanych skór w Santa Croce trafia do branży obuwniczej, natomiast skóry z Arzignano przeznaczane zostają na odzież, obuwie ale również na tapicerkę skórzaną. W Arzignano produkcja odbywa się w sposób bardziej przemysłowy i jest znacznie większa w porównaniu do Santa Croce.

Obserwując przemysł skórzany w Hiszpanii widoczne jest, że 60% garbarni zlokalizowanych jest w Katalonii, zaś 35% w Madrycie, Walencji i Murcji. Garbarnie portugalskie głównie umieszczone są w Lizbonie i Dolinie Tag, około 85 zakładów, oraz w Porto 15 [8].

12.1.5 Produkcja skór

Proces produkcyjny przetwarzania skór jest bardzo skomplikowany. Opiera się zarówno na procesach chemicznych jak i mechanicznych. Pomimo, że ogólny zarys procesów garbarskich jest stały, to każda z garbarni posiada indywidualny proces technologiczny różniący się od innych. Wynika to zarówno z surowca jak i branży, dla której produkowane są skóry.



Rys. 12.2: Skład chemiczny surowej skóry.

Ogólny schemat przetwarzania skór przedstawia Rys. 12.2 i opiera się na trzech podstawowych etapach: warsztat mokry, dogarbowanie skór oraz wykończanie skór [1].

Warsztat mokry

Skóry zdjęte ze zwierząt w pierwszym etapie są solone, a następnie magazynowane w chłodnych, ciemnych pomieszczeniach z niebieskim światłem, mające na celu ograniczenie procesów gnicia i rozwoju organizmów rozkładających materiał organiczny. Skóry są sortowane pod względem wagi.

Pierwszą operacją przetwarzania skór jest moczenie, które ma na celu zwiększenie podatności skóry na kolejne chemiczne jak i fizykochemiczne procesy. Polega ono na przywróceniu stanu skóry, który miała ona w momencie zdjęcia z danego zwierzęcia, a tym samym na odzysku wody kapilarnej. Zawartość wody w świeżej skórze wynosi około 65%.

Samo nawodnienie znacznie lepiej przebiega od strony mizdry (odmięsna strona skóry).

Skóry solone na sucho bądź wręcz wysuszone są znacznie trudniejsze do namoczenia, wielokrotnie okazuje się, że jest to wręcz niemożliwe.

Najistotniejszymi elementami mającymi wpływ na proces moczenia są:

a) metody konserwacji – mokre solenie powoduje, że stan namoczenia można osiągnąć znacznie szybciej, gdyż woda micelarna nie została naruszona

b) środki, które przyspieszają namok – najlepiej wpływają środki o zasadowym odczynie, ilość zależy od stopnia odwodnienia skóry

c) czas moczenia – zależy od temperatury, użycia antyseptyków

d) skład chemiczny wody – woda miękka ułatwia namoczenie skór

e) temperatura kąpieli – im wyższa temperatura tym większa prędkość nawodnienia i mniejszy stopień nawodnienia. Zbyt

niska temperatura powoduje znacznie wolniejsze nawodnienia i silne nabrzmiewanie lica

f) współczynnik kąpielowy – zależy od niego rozwój bakterii i częstotliwość wymiany kąpieli. Stosunek ilości wody do zielonego ciężaru surowca.

g) wpływ mechaniczny – obrotowe bębny ułatwiają namok skór

h) środki odkażające – stosowanie środków zwilżających i podnoszenie temperatury bez obawy przed rozwojem bakterii.

Tak rozmoczone skóry mogą zostać przekazane do procesu odwłaszania. Jest to pierwszy technologiczny proces powodujący zmiany w tkance skórnej. Celem tej operacji jest usunięcie zarówno włosa jak i naskórka od skóry właściwej.

Odmięśnianie (mizdrowanie) polega na usunięciu tkanki podskórnej z połączonymi z nią resztkami tkanki tłuszczowej, mięsa.

Wapnienie przeprowadza się najczęściej za pomocą wapna gaszonego, ale możliwe jest użycie go z siarczkami metali alkalicznych. Jego celem jest:

a) zmydlenie częściowe naturalnego tłuszczu

b) rozluźnienie włókien

c) osłabienie wiązania naskórka i włosa ze skórą właściwą.

Dwojenie polega na ścinaniu strony odmięsnej skóry, jeżeli ze skór ciężkich chcemy uzyskać skóry lżejsze.

Odwapnianie polega na usunięciu ze skóry stosowanych we wcześniejszych etapów zwapnionej golizny wapna gaszonego, siarczku sodowego i innych związków zasadowych. Obecność ich w kolejnych procesach przetwarzania skór działa szkodliwie na surowiec [10]. Przebiega zarówno poprzez wymywanie jak i użycie specjalnych środków chemicznych, które w reakcji z wapnem powodują tworzenie związków rozpuszczalnych w wodzie [10, 11].

Wytrawianie oparte jest na pracy enzymów, głównie proteazy mające na celu usunięcie wapna wciąż związanego ze skórą, poprawę lica, nadanie goliźnie prężności oraz spowodowanie, że stanie się miękka i śliska, rozluźnienie cebulek włosa, naskórka jak i pigmentu,

zmydlenie tłuszczy a także zwiększenie ciągliwości i przepuszczalności powietrza tkanki skórnej [1].

Garbowanie roślinne

Naturalne garbniki roślinne są wytwarzanymi substancjami organicznymi w poszczególnych częściach roślin. Najczęściej wytwarzane garbiki pochodzą z kory. Natomiast obecne są również te z drewna, liści i owoców a także korzeni. Na rynku znajduje się spora ilość tego typu środków, a także wciąż produkowane i testowane są nowe z tropikalnych roślin. Niestety nie wszystkie mogą być wykorzystywane w garbarstwie z powodu zbyt niskiej zawartości garbników i słabej jakości. Dodatkowo istotne jest rozłożenie garbników w roślinie, a także łatwość eksploatacji roślin.

Garbniki występują w postaci koloidów w wodnych roztworach, o różnym stopniu dyspersji. Pojawiają się także tzw. półgarbniki, ulegają one w odpowiednich warunkach pewnym przemianom chemicznym, dzięki czemu zyskują właściwości garbujące. Główną cechą półgarbników jest to, iż występują one na granicy roztworów właściwych i koloidalnych.

Jedną z głównych cech garbników roślinnych jest łatwa rozpuszczalność w wodzie. Dodatkowo powinny mieć wysoką zdolność reagowania z białkowymi substancjami.

Surowce, z których możliwe jest pozyskiwanie garbników dzieli się na 3 grupy:

a) Ubogie w garbniki: zawartość garbnika ok. 6%

b) Średnio bogate w garbniki: zawartość garbników 10-20%

c) Bogate w garbnik: ponad 20%

Garbowanie skór tą metodą przeprowadza się w roztworach wodnych danego garbnika roślinnego, otrzymywanego w wyniku ekstrakcji. Środki garbujące są ługowane z wykorzystaniem wody. Produkt może występować w postaci stałej bądź płynnej.

Sam proces garbowania za pomocą roślinnych garbników opiera się na dyfuzji danego garbnika w głąb tkanki golizny. Taki garbnik wiąże się z kolagenem. Kolor skóry zmienia się na brązowy stąd możliwe jest obserwowanie wnikania garbnika. Proces wnikania garb-

nika w skórę wiąże się z natychmiastowym jego wiązaniem z grupami aktywnymi białek włóknistych. Prawidłowy proces garbowania przebiega tak, że cząsteczki garbnika nie są zatrzymywane na powierzchni skóry a równomiernie przechodzą w głąb skóry. Skóry po garbowaniu roślinnym noszą nazwę „wet-white" oraz posiadają kolor żółty.

Odpady po takim ługowaniu, na przykład jeżeli jest to kora bądź drewno, mają wysoką zawartość wody, około 80%. Dlatego tak ważne jest jego wysuszenie do około 50%, a następnie możliwe jest połączenie z miałem i wykorzystanie w kotłowniach.

Garbowanie chromowe *Właściwości chromu*

Chrom – pierwiastek należący do grupy metali przejściowych z bloku d układu okresowego. Posiada liczbę atomową 24. Występuje w postaci twardego, kruchego metalu o srebrzysto -niebieskim kolorze, połyskliwy. Pozyskiwany jest z rud chromowo-żelazowych, zw. chromitem. Po wyprażeniu wcześniej już rozdrobnionego chromitu w połączeniu z węglanem sodowym w temp. 1200°C powstaje chromian sodowy. Po oczyszczeniu chromianu wykorzystuje się kwas siarkowy, po czym powstaje dwuchromian sodowy. Jest on podstawowym produktem do produkcji wszystkich związków chromu. Wykorzystywany jest do produkcji stali nierdzewnych, kwasoodpornych w postaci stopu.

Tworzy 3 rodzaje związków:

a) Chrom dwuwartościowy – nie posiadający większego znaczenia

b) Chrom trójwartościowy – tlenek słabo zasadowy i wykazujący amfoteryczne właściwości. Sole chromowe są związkami podstawowymi w garbarstwie.

c) Chrom sześciowartościowy – po ich zredukowaniu do chromu trójwartościowego już staje się wartościowym surowcem podczas garbowania. Posiada właściwości mutagenne jak i kancerogenne.

Garbowanie chromowe – proces

Ekstrakty chromowe mają postać zielonych proszków o zawartości

Cr_2O_3 25-35% bądź cieczy z 18% Cr_2O_3. Brzeczki chromowe tworzy się poprzez rozcieńczenie w wodzie ekstraktu. Następnie roztwór jest zagotowywany i trzymany w takich warunkach 12-24h.

Pikiel reguluje w początkowej fazie wnikanie garbnika do tkanki skórnej. Po zakwaszeniu skóry za pomocą roztworu chlorku sodu oraz kwasu siarkowego, następuje właściwe garbowanie. Sól obojętna zapobiega kwasowemu pęcznieniu skór. Istnieją dwie drogi garbowania chromowego. Pierwszą z nich jest jednokąpielowe garbowanie. Wykorzystuje się tutaj zasadowe siarczany chromu (III) z modyfikującymi związkami posiadającymi zasadowość 33%. Związki chromu łączą się z golizną. Metoda ta polega na początkowym dodawaniu do bębna brzeczki, która posiada niską zasadowość. Następnie podczas wnikania garbnika w skórę zwiększa się zasadowość. Pozwala to na równomierne wnikanie garbnika, dzięki czemu lico na skórze jest gładsze i delikatniejsze. Proces garbowania jest utrwalany za pomocą sody, która zwiększa stopień wiązania tkanki skórnej i garbnika.

Drugą metodą garbowania jest garbowanie dwukąpielowe. Proces ten przeprowadza się w dwóch kąpielach: impregnująca i redukująca. Podczas kąpieli impregnującej golizna jest nasycana dwuchromianem z kwasem mineralnym. W drugiej kwas dwuchromowy (obecny na kolagenowym włóknie) jest redukowany tiosiarczanem sodowym z użycie kwasu do połączeń chromowych zasadowych. Metoda ta ze względu na długotrwały proces i nieekonomiczne wykorzystanie chromu została wyparta przez metodę jednokąpielową.

Garbnik chromowy wiązany jest przez wolne grupy karboksylowe kolagenu. Kompleksy kationowe wbudowuje się między główne łańcuchy skórnego białka. Jeden chromowy kompleks łączy szereg cząsteczek kolagenu. Sole chromowe, które posiadają w swojej cząsteczce grupy zasadowe OH mają zdolności garbujące. Takie roztwory chromowe o właściwościach zasadowych ulegają kondensacji olowej i mogą tworzyć na tyle duże cząsteczki, iż umożliwia to wiązanie wielu łańcuchów polipeptydowych. Dzięki temu skóra zyskuje wysoką temperaturę skurczu, odporność na pęcznienie zarówno w kwasach jak i zasadach oraz na rozpad hydrolityczny i enzymatyczny.

Czynniki mające największy wpływ na proces garbowania chromowego to:

a) Dyfuzja garbnika

b) Wiązanie garbnika z białkiem

Produktem są skóry o kolorze niebieskim pochodzącym od garbnika, którym jest chrom i nazywane są *wet-blue*.

Regeneracja chromu

Zużyte brzeczki mogą zostać ponownie wykorzystane w początkowej fazie garbowania przy produkcji podszewek. Niestety ilości te są znaczne stąd znaleziono sposób na regenerację związków chromu poprzez wytrącanie go w postaci osadu. Tak powstały osad ulega dekantacji i rozpuszcza się w kwasie siarkowym zgodnie ze stechiometrią reakcji. Do samego wytrącania wykorzystuję się wodorotlenek sodu bądź węglan sodu [5].

Wykończanie skór

Jak zostało to przedstawienie na Rys. 12.2 warsztat mokry produkcji skór zarówno w garbarni właściwej jak i podczas dogarbowywania opiera się na *kąpielach* i operacjach w wodzie. Po tej obróbce skóry są suszone do oczekiwanej wilgotności określonej dla każdego rodzaju skór [9].

W przeszłości dział mokry przeprowadzany był w specjalnych zagłębieniach lub przymocowanych pojemnikach. Skóry zaś były przenoszone pomiędzy kolejnymi pojemnikami. Tego typu metoda wciąż jest stosowana w pojedynczych garbarniach oraz na Dalekim Wschodzie. Wspomniane pojemniki nie opróżnia się już między kolejnymi partiami skór, a podczas stosowania kilku zbiorników w każdym z nich znajduje się ten sam roztwór o innym stężeniu (wzrastającym) [12, 13].

Jednak dzisiaj wszystkie garbarnie przemysłowe używają bębnów obrotowych, pozwala to na przeprowadzanie kilku etapów procesu w jednym zbiorniku bez konieczności wyciągania skór i przenoszenia. Płyn jest wymieniany i dostarczany automatycznie bezpośrednio do bębna. Tego typu procesy trwają od kilkunastu do kilku-

dziesięciu godzin, w zależności od etapu i rodzaju skór. Procesy te działają na strukturę skóry oraz penetrują ją na wskroś.

Podczas warsztatu mokrego przetwarzania skór są zużywane różne ilości środków chemicznych mających na celu:

- otworzenie skóry i jej porów dla łatwiejszego wnikania środków pozostałych (środki neutralizacyjne),
- zabarwienie skór (barwniki),
- mające na celu dogarbowanie i wzmocnienie struktury włókien (środki dogarbowujące)
- zamknięcie skóry, w celu zatrzymania dozowanych wcześniej środków chemicznych pomiędzy włóknami kolagenowymi (kwas mrówkowy) [3].

Po opróżnieniu bębnów skóry są wyżymane i suszone, a następnie poddawane kolejnej obróbce mechanicznej poprawiającej jakość skór. Po wybarwieniu i uzyskaniu oczekiwanych kryteriów dla skór Crust (półprodukt, skóra wybarwiona i dogarbowana, niewykończona), następuje wykańczanie.

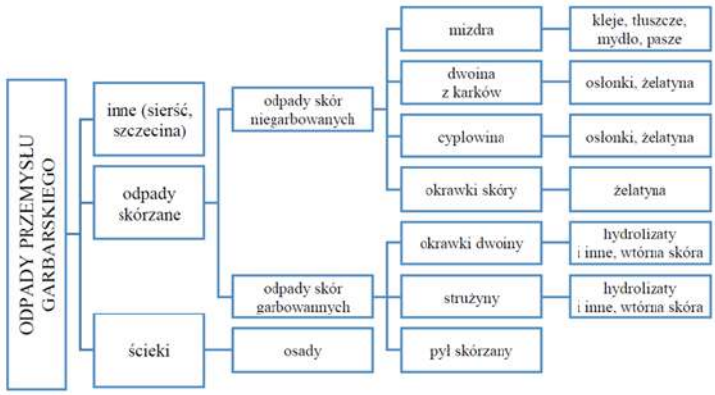
Ten etap przetwarzania skór działa powierzchniowo na skóry i głównie skupia się na poprawę właściwości fizycznych skór i walorach estetycznych. Na tym etapie używane są środki wykończeniowe w postaci różnego typu past i lakierów.

12.1.6 Przemysł garbarski a ochrona środowiska – źródła zanieczyszczeń w garbarni

Informacje wstępne

Garbarstwo jest intensywnie produkującym zanieczyszczenia przemysłem, który wytwarza ogromne ilości odpadów stałych po produkcyjnych, zapachowych czy gazowych do atmosfery, a także szlamów pochodzących z oczyszczalni ścieków. Rozwiązanie problemów związanych z utylizacją odpadów jest kwestią istotną ekonomicznie jak i społecznie nie tylko dla Polski, ale również w skali Europy czy świata [11]. Odpowiednie wyniki przeprowadzonych badań każdego rodzaju garbarskich odpadów daje szansę na minimalizację negatywnych skutków na środowisko czy zdrowie ludzi. Możliwe jest to

poprzez ograniczenie lub wyeliminowanie składowania odpadów czy skuteczną i oszczędną neutralizację gazowych i zapachowych emisji [14] [15]. Najważniejsze elementy ochrony środowiska w garbarniach dotyczą emisji (ładunku i stężenia) zanieczyszczeń do gleby, wody i powietrza, a także ich kontroli. Podczas procesów technologicznych produkcji skór wykorzystywane są duże ilości środków chemicznych, o różnych właściwościach, z czego wiele z nich może powodować specjalne wymogi dotyczące oczyszczania powietrza czy ścieków. Środki te powodują, że skóra uzyskuje odpowiednie właściwości chemiczne i fizyczne, na przykład nie gnije oraz staje się znacznie odporniejsza na rozdzieranie. Dodatkowo należy brać pod uwagę zanieczyszczenia wynikające z niekontrolowanych wycieków do gleby i wód gruntowych powstające z awarii a także podczas obróbki odpadów i ścieków [5].



Rys. 12.3: Schemat wykorzystania odpadów powstających w produkcji skór. [Źródło: Joanna Alvarez, Magdalena Panek: Ocena możliwości zastosowania odpadowego włókna kolagenowego w technologii wytwarzania materiałów skóropodobnych. Technologia i Jakość Wyrobów 62, 2017].

Rys. 12.3 obrazuje rozkład odpadów wytwarzanych podczas garbowania skór. Po analizie widać, że większość z nich to odpady stałe, powstające po procesach mechanicznych obróbki skór. Niektóre z nich znalazły zastosowanie i stanowią produkty uboczne

procesu garbowania. Sytuacja taka ma miejsce w przypadku odpadów skór niegarbowanych, które można przetworzyć w osłonki czy żelatynę używane w przemyśle spożywczym [16]. Niestety odpady powstające już po wygarbowaniu skór są znacznie trudniejsze w wykorzystaniu i dopiero prowadzone są badania, osady natomiast wciąż stanowią jedynie odpad bez dalszego przetwarzania [14, 15].

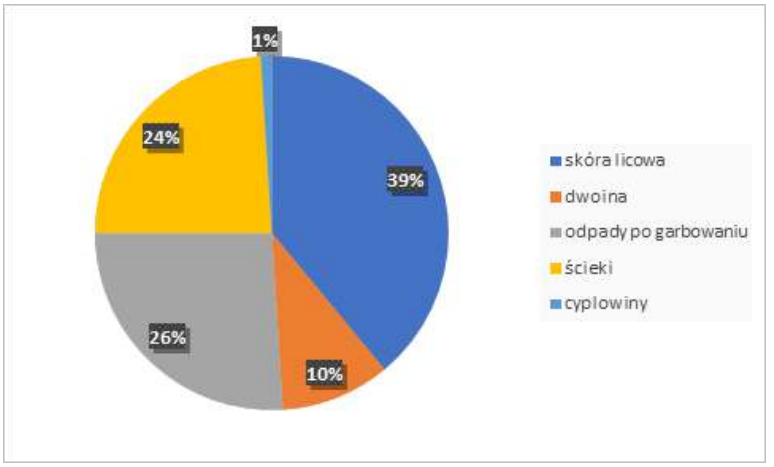
Podstawową częścią każdego audytu dotyczącego odpadów garbarskich jest ocena wydajności operacji będących składowymi procesu przetwarzania skór. Pracownicy garbarni posiadają szczegółowe dane na temat odpadów powstających podczas poszczególnych operacji takich jak cyplowanie, dwojenie, struganie, jednak rzadko posiadają przegląd całej gamy powstających odpadów.

Tylko 20-25% solonego surowca bydlęcego w odniesieniu do wagi zostaje przerobione na skóry. Natomiast nieprzerobiony materiał staje się produktem ubocznym, wciąż użytecznym i wykorzystywanym dalej, pozostałość to odpad. Ocenia się, że garbarski przemysł światowy generuje od 600 tys./rok do 4 mln ton/ rok stałych odpadów, na garbarnie polskie przypada około 250 ton/ rok. Tak więc z 1 tony surowca – surowej skóry – powstaje około 800 kg stałych odpadów [17], a także 15-50 m3 ścieków.

Skóra po przetworzeniu zyskuje właściwości niegnilne, a to wszystko za pomocą procesów mokrych (garbowania i dogarbowania), które doprowadzają do powstania skór Crust- półprodukty. Sama jakość jak i ilość odpadów i emisji produkowanych podczas procesów w garbarni jest silnie uzależniona od rodzaju skóry, która jest przetwarzana, pochodzenia skór, a także stosowanych metod. Przebieg całego procesu technologicznego powoduje wysokie zużycie energii [18]. Wynika to z konieczności ogrzewania powietrza podczas procesów suszenia, wody do pewnych operacji wymagających dostarczenia pary jak i po prostu do pracy maszyn służących obróbce mechanicznej skór. Ograniczenie zużycia energii stanowi jeden z głównych celów dla tego przemysłu, gdyż świat zmierza do jak największej redukcji występujących gazów cieplarnianych [4].

Analiza bilansu chromu przedstawiona na Rys. 12.4. udowadnia, że przy zastosowaniu w garbarni konwencjonalnej technologii niespełna 50% użytego chromu zostaje związane w skórze, pozostała

część zostaje usunięta w postaci odpadów. Dopiero nowoczesne metody garbowania jak i odzysk umożliwia zwiększenie wydajność chromu [19].



Rys. 12.4: Zawartość chromu w skórze gotowej, stałych odpadach i ściekach.

Odpady warsztatu mokrego

Zgodnie z Rys. 12.3 widać, ze podczas samego warsztatu mokrego powstaje wiele grup odpadowych. Część z nich znajduje zastosowanie w innych przemysłowych gałęziach. Najważniejsze na samym początku jest odpowiednie przesortowanie i podzielenie ich, a następnie odpowiednie ich zakonserwowanie. Pominięcie, któregoś z etapów może spowodować obniżenie jakości lub zniszczenie, któregoś z rodzajów odpadów. Odpady klejowe występują w postaci tkanki przymięsnej, dwoiny, łebki. Są one dzielone na dwie podgrupy: klejówki -zawierające tkankę przymięsną i niezawierające tkanki przymięsnej, dwoiny oraz odzierki – łebki. Klejówki zabezpieczane są za pomocą środków (na przykład mleko wapienne), które nie naruszają substancji skórnej, a produkcja kleju nie jest niczym utrudniona.

Odpady te wysyłane mogą być do fabryk produkujących kleje. Odzierki po odpowiednim wypłukaniu, krojeniu a następnie wap-

nieniu, zobojętnia się kwasem, dzięki czemu powstaje tzw. zupa klejowa, która już po zagęszczeniu daje klej. Dwoiny stosowane są do produkcji skór podszewkowych, pewnych imitacji weluru lub nakłada się na nie specjalne sztuczne lico, dzięki czemu zyskują znacznie większą wartość.

Odpady w postaci sierści i szczeciny są dość cenne zarówno w kraju jak i za granicą. Ich zastosowanie jest bardzo szerokie, ponieważ możliwe jest produkowanie z nich różnego typu szczotek i pędzli, ale również pilśniu jak i filcu. Dodatkowo może powstawać gorszej jakości wełna. Odpady tłuszczowe powstają jako tłuszcz naturalny. Taki półprodukt wysyłany jest do odpowiednich zakładów, gdzie przetwarzany jest na łój czy wykorzystywany podczas produkcji mydła i gliceryny.

Odpady procesu wykończania skór

Wyselekcjonowane grupy odpadów można podzielić ze względu na:

I. Postać, w której występują:
   a) Odpady stałe
   b) Ścieki
   c) Gazowe zanieczyszczenia

II. Sposób garbowania:
   a) Chromowe
   b) Niechromowe
   c) Mieszane

III. Pochodzenie- operację, podczas której powstają:
   a) Strużyny-struganie
   b) Cyplowiny surowca- cyplowanie
   c) Dwoiny – dwojenie
   d) Pył – międlenie, odpylanie
   e) Cyplowiny suche – cyplowanie
   f) Odsiek – pozostałość po wykończeniu skór

g) Ścieki – dogarbowanie i barwienie

h) Muł – oczyszczanie ścieków

Odpady stałe to w większości organiczne materiały, które składają się z tłuszczy, białek, a także środków chemicznych, które zostały wprowadzone do skóry podczas wcześniejszych etapów. Objętość i skład tych odpadów opiera się zwykle na rodzaju procesów, które zostały wykorzystane. Strużyny jako jeden z rodzajów odpadów stałych powstają już na samym początku procesów przetwarzania skór, poprzez struganie za pomocą noży strony odmięsnej skóry zwanej potocznie mizdrą, mają postać wiór przedstawionych na Rys. 12.5. Dokładniej operacja ta polega na wprowadzaniu skóry do strugarki, w której znajdują się wały nożowe, gdzie następuje mechaniczne usunięcie nadmiaru organicznej materii. Operacja ta, jako jedna z najważniejszych, ma celu pozyskanie skór o odpowiedniej grubości, aby jakość produktu końcowego była na oczekiwanym poziomie. Strużyny powstające z tego etapu stanowią nawet 2-20% wagi skóry wygarbowanej [10]. Dodatkowym zanieczyszczeniem generowanym z tej operacji jest pył powstający podczas strugania i unoszący się w powietrzu, co powoduje konieczność zabezpieczania dróg oddechowym operatorów maszyn [20]. Podział strużyn przedstawia się następująco: strużyny chromowe (posiadają kolor niebieski, pochodzą po zestruganiu skór garbowanych za pomocą chromu) oraz niechromowe (w kolorze żółtym, pochodzą ze strugania skór po garbowaniu roślinnym -wolnym od chromu). Ilość poszczególnych strużyn jest silnie zależna od zapotrzebowania na rynku na dany rodzaj skór, jednak to skóry chromowe wciąż przeważają i stanowią większość produkcji (dochodzi nawet do 70% w stosunku do całej produkcji). Powodem jest wysoka jakość skór w wyrobie gotowym. Dzięki zastosowanemu siarczanowi chromu (III), podstawowemu garbnikowi garbowania chromowego, skóra osiąga odpowiednie właściwości, takie jak mechaniczna i termiczna odporność oraz miękkość. Ilość użytego chromu podczas procesu wynosi około 1-5% masy skóry solonej.

Strużyny są niepotrzebnym produktem ubocznym. Na chwilę obecną ich utylizacja jak i zagospodarowanie jest bardzo uciążliwe, nie ma zadowalającego rozwiązania na tą problematyczną kwestię. Są jedy-

Tab. 12.4: Przykładowe parametry strużyn.

Parametry	Wartości
Roczny uzysk [ton]	2400
Miesięczny uzysk [ton]	200
Stosunek chromowe/ niechromowe (wartości przykładowe)	70/30
Rozpuszczony węgiel organiczny (OC) [mg/kg]	1300
Zawartość chromu (dla strużyn chromowych) [mg/kg]	31 600
Zawartość chromu (dla strużyn chromowych)	65,20%
Sucha masa w 105°C	58,40%
Chlorki [mg/kg]	18 400
Fluorki [mg/kg]	5,78
Siarczany [mg/kg]	6 780
Gęstość nasypowa [kg/m^3]	285

nie składowane na dostosowanych do tego celów wysypiskach, jednak z powodu ogromnych ilości staje się to coraz większym utrudnieniem dla garbarni jak i firm wyspecjalizowanych w odpadach. Dodatkowo są stopniowo prowadzone badania nad ich kompostowaniem, tworzeniem hydrolizatu białkowego [21], a w sytuacji intensywnego zmielenia na drobny produkt dodawane są do tworzyw sztucznych [22] jako wypełniacz [9].

Podczas samego strugania skór jak to zostało przedstawione w Tab. 12.4 powstaje średnio około 200 Mg/ miesiąc strużyn, 9-10 ton/dobę. Może to stanowić około 40-50% całkowitej ilości odpadów. Rodzaj odpadów mizdrowych odgrywa istotną rolę, ponieważ skóry chromowe są znacznie trudniejsze do przerobu ze względu na zawarty w nich chrom, który może się wydzielać podczas obróbki. Jego zawartość w skórach wynosi około 31 600 mg/kg. Dodatkowo w składzie znajdują się chlorki około 18 400 mg/kg, fluorki 5,778 mg/ kg oraz siarczki 6 780 mg/kg. Wilgotność strużyn zawiera się przedziale w 50-60%, co powoduje konieczność ich suszenia przed przerobem. Gęstość nasypowa tego typu odpadów wynosi 285 kg/m^3.

Rys. 12.5: Fotografia odpadów w postaci strużyn: chromowe (lewa) i niechromowe (prawa).

Rys. 12.6: Fotografia odpadów w postaci dwoin.

Dwojenie jest etapem, podczas którego również generowane są odpady stałe przedstawione na Rys. 12.6, jednak w tym przypadku skóry są już zabarwione i dogarbowane, przez co znajduje się w nich znacznie bogatsza kompozycja środków chemicznych. Proces ten ma

na celu tylko minimalne usunięcie strony mizdrowej, aby wyrównać grubość skór na gotowo. Dwojenie może odbywać się zarówno od strony mizdrowej, ale również pewien rodzaj skór wymaga dwojenia strony licowej. Ilość tych odpadów waha się w granicach około 7-10 ton/miesiąc. Co może stanowić około 3-5% całkowitej ilości odpadów. Ilość tego rodzaju odpadów uzależniony jest zarówno od procesu technologicznego jak i wymagań klienta. Dwoina skóry surowej w przypadku, gdy jest wystarczająco gruba, może zostać wykorzystana do produkcji zamszu. W przypadku gdy po dwojeniu jest zbyt cienka, bardzo często przeznacza się taki odpad na żelatynę, osłonki wędlin. Jeżeli skóra zdwojona po garbowaniu da nam dużą ilość odpadów zostają one przekazane producentom szewskiej tektury, jednak w większości przypadków są wyrzucane jako odpad [5]. Dodatkowo powstają odpady w postaci kawałków skór zwane cyplowinami, postać ich została pokazana na Rys. 12.7. Odpad ten występuje w dwóch postaciach. Pierwszą z nich są cyplowiny surowca, czyli powstałe jeszcze przed dogarbowaniem i wybarwienie skór. Stąd są one uboższe w związaną we włókna chemię. Podzielone są one na cyplowiny niechromowe jak i chromowe. Ilość ich wynosi około 2% całkowitej ilości odpadów. Zależne jest to od jakości skór, przeprowadzonego wcześniej procesu strugania jak i kwalifikacji pracowników wykonujących daną operację.

Rys. 12.7: Fotografia odpadów w postaci cyplowin surowca niechromowego (lewa) i chromowego (prawa).

Drugi rodzaj cyplowin to tzw. cyplowiny suche – Rys. 12.8. Powstają one już na samym końcu procesu technologicznego, mające na celu jedynie wyrównanie obrzeży skóry i pozbycie się zbędnych skrawków skór. Średnio miesięcznie w zależności od garbarni, jakości wcześniej przeprowadzonego procesu technologicznej (mechaniczna obróbka skór) jak i umiejętności pracowników ilość ta waha się w granicach 2% całkowitej ilości odpadów.

Rys. 12.8: Fotografia odpadów w postaci cyplowin suchych.

W znacznie mniejszej ilości występują odpady w postaci pyłów (ok.1%), przedstawionych na Rys. 12.9. Odpad ten powstaje podczas międlenia bębnowego skór jak i odpylania ich. Ma on bardzo duże rozdrobnienie i jest suchy (około 10% wilgotności skór). Składowanie tego typu odpadów na chwilę obecną jest w znacznie większym stopniu ograniczone niż miało to miejsce w przeszłości, powodem mogą być wprowadzone przepisy prawne dotyczące tego typu procederu. Spowodowało to poszukiwanie rozwiązań mających na celu ponowne wykorzystanie tych odpadów za pomocą kompostowania, produkcji biogazu czy energii odnawialnej.

Rys. 12.9: Fotografia odpadu w postaci pyłu.

Osady dzisiaj już nie mogą być składowane w takim stopniu jak miało to miejsce jeszcze 30 lat temu. Jednym z powodów jest coraz mniejsza dostępność składowisk, a także dyrektywa 99/31/WE określająca przepisy dotyczące składowania odpadów. Członkowskie państwa mają prawo zadecydować o limicie w kwestii zawartości całkowitego organicznego węgla w odpadach skierowanych na składowisko. Niektóre państwa już postawiły granice i tak w Austrii wynosi ona 5%, zaś w Niemczech 1-6%.

Możliwa jest opcja składowania osadów z podczyszczalni ścieków już po pewnym wstępnym ich oczyszczeniu za pomocą fermentacji beztlenowej. Niestety, metoda jest dość problematyczna ze względu na przejście siarczanów w siarczki poprzez pewne gatunki bakterii. Może to doprowadzić do wydzielania siarkowodoru. Wytwarzany siarkowodór jest trujący, rozbudowanie składowiska o instalację do jego oczyszczania generuje kolejne koszty, a tym samym dla garbarni również koszty oddania odpadów uległyby zwiększeniu. Dlatego na początku oczyszcza się je wstępnie, aby usunąć siarkę przed procesem beztlenowej fermentacji. Celem składowisk jako najprost-

szej metody utylizacji odpadów z garbarni jest jedynie składowanie i zabezpieczenie ich bez nakładu dodatkowych kosztów [5].

Pomimo tylu etapów oczyszczania odpadów, zawartość całkowitego węgla organicznego wciąż może pozostać za wysoka, a tym samym mogą zostać przekroczone limity wprowadzone w wymogach niektórych państw członkowskich Unii Europejskiej dotyczące zawartości całkowitego węgla organicznego odpadów przeznaczonych do unieszkodliwiania na składowiskach [5].

Kompostowanie jako metoda wspomagająca i przygotowująca odpady do składowania, pozwala na redukcję organicznej materii. Polega na układaniu w cienkie pasy odpadów, co pozwala przyspieszyć rozpad tlenowy. Przeszkodą jest silny zapach, który niektóre odpady garbarskie wydzielają, co powoduje, że ten sposób przerobu nie jest praktykowany często [4].

W Europie dość powszechnie stosuje się osady, a także produkty beztlenowej fermentacji na grunt rolny. Jednak coraz częściej pojawiają się głosy sprzeciwu wynikające ze strachu, iż może zawarty w nich być chrom, patogeny czy inne zanieczyszczenia, które spowodują wysycenie gleb odżywczymi składnikami. Określa się maksymalne stężenie zawartości metali ciężkich w tym osadzie oraz ich ilość, które pozwala jeszcze na wykorzystanie go na gleby. Dlatego wstępne oczyszczenie osadu staje się niezbędne [9].

Dzisiaj prowadzi się coraz więcej badań nad spalaniem, pirolizą [23] czy gazyfikacją odpadów z jednoczesnym odzyskiem energii [24]. Produktem ubocznym tych procesów jest popiół, który jest znacznie łatwiejszy do utylizacji niż wstępne surowce procesów termicznych. Jeżeli jednak zawarty jest chrom w odpadach pojawia się konieczność prowadzenia spalania w dwóch etapach. Pierwszy ma miejsce we wstępnej komorze, do której dopływ tlenu jest ograniczony,aby zapobiec utlenienia chromu (III) do chromu (VI). Zgodnie z IV rozdziałem dyrektywy 2010/75/WE podczas spalania wymaganejest aby, w komorze wtórnego spalania, temperatura gazów utrzymana była przez min.2 s powyżej 1123 K. Pozwoli to na zniszczenie furanów i dioksyn. Ważne podczas tego jest stałe monitorowanie emisji zanieczyszczeń. Koszty, które zapewniają zgodność powodują, że próg opłacalności dla spalania zostaje znacznie podniesiony [8].

Zużycie wody i ścieki po procesie garbowania

Ścieki garbarskie

Ścieki garbarskie są specyficznym odpadem produkowanym w przemyśle ze względu na wysokie zanieczyszczenie i bogaty skład [25]. Powoduje to, że oczyszczenie ich do poziomu akceptowalnego przez oddziały kanalizacji i wodociągów jest bardzo trudne. Problem ten można rozwiązać poprzez dobór odpowiedniej technologii oczyszczania, co doprowadzi do uzyskania oczekiwanej jakości ścieków [26].

Ilość użytej podczas wyprawy skór wody określa ilość produkowanych ścieków. W garbarniach średnie zużycie wody kształtuje się na poziomie około 60-80 litrów wody/1 kg przetwarzanej skóry. Oczywiście ilość ta jest zależna także od procesu technologicznego. Możliwa jest redukcja ilości wody dostarczanej do procesu, natomiast ilość ścieków odprowadzanych nie ulega zmniejszeniu, ponieważ wzrasta wtedy stężenie zanieczyszczeń w wodach odprowadzanych z procesu [9]. Dlatego też ścieki pochodzące z garbarni są najtrudniejszymi z przemysłowych ścieków, a skład ich zmienia się w zależności od rodzaju wyprawy oraz ilości i rodzaju stosowanych środków chemicznych [27].

Ponad 80% garbarni zlokalizowanych w Europie swoje ścieki odprowadza do publicznej kanalizacji. W przypadku Włoch oraz Hiszpanii potworzone są klastry branży skórzanej, które podłączone zostały do jednych oczyszczalni ścieków [9].

Ścieki garbarskie dzielą się na:

a) Ścieki warsztatu mokrego

b) Ścieki z garbowania

c) Ścieki z barwienia, natłuszczania

d) Ścieki z wykończalni [5]

Garbarnie przeprowadzają wstępne oczyszczanie ścieków w przyzakładowych podczyszczalniach, aby uzyskać założone w normach wartości poszczególnych składników. Podstawowe metody oczyszczania ścieków przez garbarnie zostały przedstawione w Tab. 12.5. Garbarnie w celu zmniejszenia produkowanej objętości osadu wykorzystuje odwodnienie. Wykorzystywane są do tego celu głównie

urządzenia mechaniczne typu prasy filtracyjne czy wirówki, czasami dodatkowo stosuje się po tym suszenie [20].

Nośnikiem zanieczyszczeń w ściekach pochodzących z garbarni są przede wszystkim:

- kwas mrówkowy;
- mrówczan sodu;
- soda (węglan);
- chrom +3 , jako siarczan chromu;
- barwniki organiczne nie zawierające metali ciężkich;
- detergenty;
- emulsje tłuszczów naturalnych.

Tab. 12.5: Metody oczyszczania ścieków wraz ze składnikami usuwanymi podczas danego oczyszczania.

Metoda	Przeznaczenie
Wstępne mechaniczne oczyszczanie	Ma na celu usunięcie naturalnych tłuszczów, łoju i oleju, a także ciał stałych z wykorzystaniem sedymentacji.
Oczyszczanie fizykochemiczne	Przebiega z wykorzystaniem procesów utleniania, sedymentacji, flotacji, neutralizacji, wyrównania przepływów. Polega na usunięciu organicznej materii i siarczków, a także chromu (III) w procesach garbowania.
Oczyszczanie biologiczne	Wykorzystuje filtry zraszane i bioaerację. Polega na obniżaniu zawartości organicznych związków, a także obejmuje nitryfikację i denitryfikację.
Sedymentacja	Umożliwia oddzielenie osadu od ścieków oczyszczonych. Osad pierwotny jest pobierany i oczyszczany w tzw. buforowym zbiorniku.

Ścieki warsztatu mokrego mają podobny, a wręcz taki sam skład w każdej garbarni, ponieważ technologia przetwarzania niewiele się różni. Dodatkowo mają wysokie stężenia i są alkaliczne. Głównymi składnikami tego typu ścieków są: sól kuchenna, białka rozpuszczalne i produkty ich rozkładu, związki siarki, związki wapniowe, bakterie, kawałki skór i sierści. Największe ilości ścieków odprowadzane są z etapu moczenia, gdzie do ścieków trafia sól kuchenna, białka rozpuszczalne i bakterie oraz z procesu wapnienia, z którego odprowadzane są wodorotlenki i węglany wapnia, związki organiczne, produkty pochodzące z rozpadu skóry i włosa oraz siarczki [5].

Analiza ścieków

Etapy podstawowego dla garbarni procesu technologicznego, które w głównej mierze odbywają się w roztworach powodują, że ścieki zawierają duże ilości zanieczyszczeń organicznych takich jak chemiczne zapotrzebowanie na tlen (ChZT) i biologiczne zapotrzebowanie na tlen (BZT) oraz składniki mineralne (garbniki). W laboratoriach przy oczyszczalniach podczas badania ścieków określa się: ChZT, BZT, zawieszoną materię, azot amonowy i całkowity, zawartość chromu, pH, fosfor całkowity, siarczki, temperaturę oraz zawartość tłuszczu [18, 28]. Średnio do procesu technologicznego produkcji skór dodawane jest ok. 500 kg środków chemicznych wpływających na właściwości chemiczne i fizyczne skór, a tym samym na właściwości fizyko-chemiczne odpadów. Pod hasłem wody użytej znajduje się woda procesowa oraz woda wykorzystywana do technologicznych procesów w celu czyszczenia, wytwarzania energii. Sama ilość wody procesowej jest w dużej mierze zależna od rodzaju procesów, surowca czy produktów. Średnio przyjmuje się, że na 1 tonę przetwarzanej skóry przypada 15-50m^3 ścieków [4]. Ogólnym miernikiem stężenia ścieków jest chemiczne zapotrzebowanie na tlen. Stąd też zakłady garbarskie, które zużywają mniej wody produkują ścieki o większym wskaźniku ChZT. Barwniki dodawane do procesu barwienia i dogarbowania wpływają za zwiększenie ChTZ i BZT, a także uwalnianie przyswajalnych chlorowców organicznych. Jeżeli poziom zastosowanego barwnika przekroczy 10 ppm w ściekach to jego obecność staje się widoczna nawet dla ludzkiego oka i potrafią zmienić ich kolor. Środki te są bardzo drogie, dlatego garbarnie dążą do maksymalnego wyczerpania kąpieli z barwnika, udaje się to w ponad 90%. Stosowane podczas tej operacji barwniki łączą się za pomocą jonowych oddziaływań pomiędzy anionowo-sulfonową grupą barwnika i grupą kolagenu kationowo-aminową. Ilość stosowanego barwnika w dużej mierze zależy od produkcji zakładu i obecnych trendów na rynku [19]. Na etapie wykończalniczym pigmenty zawierające ołów lub chromiany są zakazane przez poprawki dla dyrektywy 67/548/EWG i 76/769/EWG, określone jako niebezpieczne [8]. W Tab. 12.6 i Tab. 12.7. znajdują się przykładowe dane dotyczące ilości wody pobieranej do procesu dogarbowania skór oraz powstających z tego ścieków w przeliczeniu na rok i miesiąc, a także

składu ścieków surowych pochodzących z procesu chemicznej obróbki skór (dogarbowanie i barwienie) oraz tych wstępie oczyszczonych w przyzakładowej podczyszczalni z podaniem przykładowych zawartości poszczególnych związków. Ilość ścieków produkowanych w garbarni może dochodzić nawet do 60 000-80 000 m^3/rok, przy ilości wody wprowadzonej do garbarni na poziomie około 80 000-90 000 m^3/rok, jak to zostało przedstawione w Tab. 12.6. Trudno jednoznacznie określić skład ścieków garbarskich. Wyniki poszczególnie badanych parametrów zmieniają się w czasie i zależne są od kilku składowych. Bardzo duży wpływ ma rodzaj i ilość użytej chemii, ilość i rodzaj przetwarzanych skór (skóry chromowe i niechromowe), a także jakość skór i ich zdolność wchłaniania chemii. Wykorzystując przykładowe wyniki składu ścieków surowych jak i oczyszczonych poziom ChZT jest bardzo wysoki. Zgodnie z wynikami podanymi w Tab. 12.6. w ściekach surowych może dochodzić nawet do 8000-14000, a po wstępnym oczyszczeniu udaje się go zredukować nawet do 2000-4000. Niestety ten poziom jest wciąż za wysoki i pojawia się potrzeba redukcji w większym stopniu ChZT-ów. Zwłaszcza, że produkcja skór jest zmienna w czasie, a także wykorzystywane są nowe środki chemiczne. Ilość ich znacznie wzrasta w porównaniu do garbowania skór jeszcze kilka lat temu. pH ścieków oczyszczonych utrzymuje się na poziomie zasadowym i wynosi 7,5-8. Ścieki surowe ze względu na zawartość różnych związków chemicznych są w tolerancji kwasowej 4-4,5. Chrom w ściekach surowych ma znacznie przekroczoną wartość dopuszczalną. Oczyszczanie w podczyszczalni pozwala na silną redukcję. Taki chrom jest odbierany ze ścieków i trafia do osadu po procesie oczyszczania wraz z innymi składnikami. Ilość osadu w porównaniu do ścieków może wynosić około 5-6%. Każdy z parametrów ma tolerancję, która nie może być przekroczona. Ścieki wstępnie oczyszczone i zawierające poszczególne parametry w tolerancji trafiają do miejskiej oczyszczalni, gdzie są ponownie badane, a także łączone z innymi ściekami i w kolejnych stopniach oczyszczane. Przykładowe maksymalne wartości, które są dopuszczalne dla ścieków dla zakładów wodociągów i kanalizacji zostały zawarte w Tab. 12.8. Wynika z niej, że ilość ChZT nie może przekraczać 8000, natomiast pH powinno się mieścić w tolerancji 6,5-9,5.

Poza procesami działu mokrego, ścieki pojawiają podczas wykończania skór, które ma na celu nadanie odpowiednich właściwości wizualnych i haptycznych skór. Etap ten powoduje wysokie zużycie wody, a także obecność w ściekach pigmentów, rozpuszczalników czy lakierów. Z powodu wysokiej zawartości związków chemicznych, pojawia się duży problem dla środowiska, co wymaga specjalnego zarządzania nimi [8]. Oszacowano, że podczas procesu efektywnym zużyciem środków chemicznych określa się, gdy zawartość ich w produkcie gotowym wynosi 15%. Wynika z tego, że aż 85% trafia do ścieków [20]. Mechaniczne oczyszczanie ścieków prowadzi do wytworzenia ogromnych ilości osadów i szlamów, które są trudnym i poważnym problemem dla środowiska. Osad powstały po oczyszczaniu ścieków jest bogaty w różnego rodzaju związki chemiczne, w tym chrom. Na chwilę obecną nie znaleziono racjonalnego sposobu na utylizację tego rodzaju odpad, dlatego też jest on jedynie składowany na specjalnych wysypiskach. Bydlęce garbarnie zlokalizowane w Europie wytwarzają nawet około 300 ton osadu/ miesięcznie, których wilgotność waha się w granicach 40-80%. Wstępna obróbka ścieków generuje około 5-10% stałych substancji, w odniesieniu do objętości ścieków. Osad, który jest produktem końcowym tych procesów jest w postaci stałej.

Sposoby jakimi unieszkodliwiane są tego typu osady w dużej mierze uzależnione zostało od lokalnych możliwości składowisk.

Tab. 12.6: Średnie przykładowe parametry wody wykorzystywane podczas procesów garbowania skór.

Parametry wody	**Wartości**
Roczne użycie wody [m^3]	80 000-90 000
Miesięczne zużycie wody [m^3]	6 000-9 000
Roczna produkcja ścieków [m^3]	60 000-80 000
Miesięczna produkcja ścieków [m^3]	5 000-7 000

Tab. 12.7: Średnie przykładowe parametry ścieków powstających podczas procesów garbowania skór.

Parametry	**Ścieki surowe**	**Ścieki oczyszczone**
pH	4-4,5	7,5-8,5
Przewodność mS/ml	9-11	12-,19
Chlorki mg/l	500-2000	2000-5500
NH4 mg/l	2-9	5-14
Siarczany mg/l	2000-4000	2000-4000
ChZT	8000-14000	4000-7000
BZT	–	1500-2000
Cr^{3+} [mg/l]	12-40	0,2-1,0
N og [mg/l]	70-100	40-60
P og [mg/l]	30-60	2-6
Zawiesina ogólna	–	30-70

Tab. 12.8: Przykładowe maksymalne dopuszczalne wartości.

Parametr	**Maksymalna dopuszczalna wartość**
pH	6,5-9,5
Przewodność mS/ml	–
NH4 mg/l	200
Chlorki mg/l	6000
Siarczany mg/l	5000
ChZT	8000
BZT	3000
Cr^{3+} [mg/l]	1
N og [mg/l]	300
P og [mg/l]	15
Zawiesina ogólna	1500

Garbarnie skór bydlęcych mają zaproponowane przez Komisję do spraw Środowiska Międzynarodowego Związku Inżynierów i Techników Przemysłu Skórzanego wartości średnie zanieczyszczeń, które zostały przedstawione w Tab 12.9.

Tab. 12.9: Wartości określone przez Komisję ds. Środowiska Międzynarodowego Związku Inżynierów i Techników Przemysłu Skórzanego.

Wartości namoczonej surowej skóry [na tonę]	**woda [m3/t]**	**ChZT (kg/t)**	**BZT (kg/t)**	**zawiesiny (kg/t)**	**Cr^{3+} (kg/t)**	**$S2_-$ (kg/t)**
Solone bydlęce skóry						
mokry warsztat	7-25	120-160	40-60	70-120	-	2-9
garbowanie	1-3	10-20	3-7	5-10	2-5	-
wyprawa (dogarbowanie, barwienie)	4-8	15-40	5-15	10-20	1-2	-
wykańczanie	0-1	0-10	0-4	0-5	-	-
Razem	12-37	145-230	48-86	85-155	3-7	2-9

Aspekty prawne dotyczące odpadów garbarskich

Podział garbowania jest zależny od stosowanych środków garbujących. Pomimo coraz większej świadomości dotyczącej zanieczyszczeń środowiska, wciąż w 80-90% światowej produkcji skór, proces technologiczny opiera się na garbowaniu skór za pomocą soli chromu (III). Załącznik X Ramowej Dyrektywy Wodnej 2000/60/WE po uaktualnieniu poprzez dyrektywę 2008/105/WE nie posiada w kwestii substancji priorytetowych zawartego w swoim spisie chromu (III) . Wykaz europejski odpadów niebezpiecznych [KATALOG ODPADÓW NIEBEZPIECZNYCH] nie zawiera w swoim zestawieniu odpadów garbarskich z zawartością chromu (III), są one klasyfikowane jako ” Odpady skóry wygarbowanej zawierające chrom (wióry; obcinki; pył ze szlifowania skór)„ pod kodem 04 01 08- jak przedstawia to Tab. 12.10 [8]. Powodem tego jest brak charakterystyk będących podstawą do podporządkowania ich jako niebezpieczny odpad [13]. Chrom (VI) nie jest stosowany podczas procesu garbowania skór, natomiast może powstawać na drodze utleniania chromu (III) [9, 15]. Reakcjami doprowadzającymi do tworzenia się chromu

(VI) jest utlenianie pośrednie, wykorzystanie w reakcji chemicznej produktów przejściowych , jak i procesy starzeniowe materiału i promieniowanie UV [3]. Czyni go to rakotwórczym [15] i wysoko toksycznym dla człowieka [16], ponieważ posiada zdolność penetracji komórek ludzkiego ciała [21]. Szybko wchłaniany poprzez system transportu siarczanów może spowodować wiele zmian wewnątrzkomórkowych jak również mutacje DNA [22]. Przepisy Unii Europejskiej wg Rozporządzenia Komisji EU Nr. 301/2014 z 25 marca 2014 pozwalają na próg zawartości chromu (VI) dla artykułów skórzanych mających kontakt ze skórą ludzką w tolerancji ¡ 3 mg/ kg, w przeliczeniu dla całkowitej suchej masy 0,0003% [29]. Dlatego też garbarnie stosują specjalne środki ostrożności, aby zapobiegać reakcjom utleniania chromu (III) do chromu (VI) [23]. Wg normy ABNT/ NBR 10004 wszystkie odpady stałe wytwarzane w garbarniach zawierające chrom (III) zostają sklasyfikowane jako odpad „klasy I" [22].

Tab. 12.10: Europejski katalog odpadów.

Odpady z przemysłu skórzanego i futrzarskiego	
Kod odpadu	**Rodzaj odpadu**
04 01 01	Odpady z mizdrowania (odzierki i dwoiny wapniowe)
04 01 02	Odpady z wapnienia
04 01 03*	Odpady z odtłuszczania zawierające rozpuszczalniki (bez fazy ciekłej)
04 01 04	Brzeczka garbująca zawierająca chrom
04 01 05	Brzeczka garbująca nie zawierająca chromu
04 01 06	Osady zawierające chrom, zwłaszcza z zakładowych oczyszczalni ścieków
04 01 07	Osady nie zawierające chromu, zwłaszcza z zakładowych oczyszczalni ścieków
04 01 08	Odpady skóry wygarbowanej zawierające chrom (wióry, obcinki, pył ze szlifowania skór)
04 01 09	Odpady z polerowania i wykańczania
04 01 99	Inne nie wymienione odpady

Emisja do atmosfery zanieczyszczeń

W przypadku emisji do powietrza pojawia się podział na emisje pochodzące:

1. z procesu technologicznego (węglowodory alifatyczne, pył),
2. ze spalania paliw: dwutlenek siarki, dwutlenek azotu i pył
3. ze zbiorników uśredniających: aceton, amoniak, siarkowodór

Tab. 12.11: Przykładowe liczby źródeł emisji i ich skład.

Emisje	
Zidentyfikowana łączna ilość źródeł zanieczyszczeń	18
Źródła zanieczyszczeń z procesu technologicznego [liczba]	8
WWA [kg/rok]	1 - 47
Pył PM 2,5 [kg/rok]	0,84-1,27
Źródła zanieczyszczeń ze spalania [liczba]	1
NO2[kg/rok]	1299
Źródła zanieczyszczeń ze zbiornika uśredniającego [liczba]	1
WWA [kg/rok]	0,05-1,8

Ilość przykładowa źródeł zanieczyszczeń oraz ich skład znajduje się w postaci zestawienia w Tab. 12.11. Dodatkowo pewna emisja pochodzi z mechanicznych procesów podczas odpowietrzania zbiorników oraz ostrzenia noży. Niektóre z emisji mogą być toksyczne, a także zawierają substancje zapachowe [30]. Dyrektywa dotycząca emisji przemysłowych, rozdział V, reguluje kwestie prawne o minimalizacji lotnych związków organicznych w wydzielanej do powietrza emisji. W przypadku zbyt wysokiego poziomy tego typu substancji wymaga się wprowadzenia wentylacji, która umożliwia oczyszczanie powietrza z pyłów powstających w wyniki różnego typu operacji mechanicznych oraz z używania środków chemicznych sypkich. Zawartość pyłów jest uzależniona od przebiegu procesu oraz produkcji skór, może wynosić średnio około 0,84-1,27 kg/rok.

Emisje z procesów spalania dla celów energetycznych, lotnych związków lotnych (LZO) i siarczków są bardzo ważne ze względu na jakość powietrza atmosferycznego. Spalanie odpadów garbarskich staje się problematyczne z powodu wydzielania toksycznych związków takich jak chrom (VI), wyższych węglowodorów alifatycznych.

Emisje powstające dla celów energetycznych nie mają negatywnego wpływu na pobliskie okolice, natomiast wpływają na wzrastający efekt cieplarniany.

Emisje LZO wynikają głównie z wykorzystania w procesie wykończania skór rozpuszczalników. Ich ilość w dużej mierze uzależniona

jest od produkcji oraz rodzaju użytego środka, a może wynosić do 300 gr/m_3 gotowej już skóry. Część jednak powstaje w wyniku szlifowania skór. Podczas ich ograniczania stosuje się filtry z węgla aktywnego, jednak wśród garbarni nie jest to zbyt powszechny sposób [8].

Miejsce wydzielania amoniaku głównie występuję przy operacji odwapniania i barwienia. Jednak już poprzez płukanie jak i kontrole procesową garbarnie są w stanie zminimalizować jego ilość.

Sama wentylacja nie jest dobrym sposobem, ponieważ może powodować uwolnienie emisji do środowiska.

Garbarnie zobowiązują się do przestrzegania przepisów prawnych rozporządzenia (WE) nr 1069/2009, ponieważ poprzez niekontrolowane wycieki chemii płynnej i odpady mogą w dużej mierze wpłynąć na glebę jak i wody gruntowe [4, 17].

Energia wykorzystywana w procesach garbarskich

Wielkość zużycia energii podczas przetwarzania skór zależy od kilku istotnych czynników. Pierwszym z nich są metody produkcji oraz jej wielkość. Większa produkcja powoduje, że zużycie wzrasta i jest konieczność użytkowania większej ilości maszyn. Sama metoda produkcji jest zależna od rodzaju skór, na które pojawił się popyt. Mogą się one różnić ilością przeprowadzanych operacji. Bardzo istotne jest zaawansowanie silników elektrycznych i ich wiek.

Suszenie jako operacja wymagająca dostarczenia ilości energii jest procesem silnie energochłonnym, pochłania nawet 45% całkowitego zużycia energii zakładu. Ilość energii wykorzystywanej do tej operacji zależy w dużej mierze od wilgotności skór i warunków panujących na hali, jeżeli w pomieszczeniu znajduje się wysoka wilgotność pojawia się konieczność zwiększenia energii podczas suszenia lub powtórzenie go dwukrotnie.

Poziom rozwinięcia zakładu wpływa na jakość i wiek maszyn, jeżeli ich stan nie jest prawidłowy, może dojść do strat cieplnej energii ze zbiorników procesowych i hal produkcyjnych. Utratę tego ciepła można ograniczyć za pomocą termoizolacji. Zwiększenie jej spowodowane zaś może być poprzez spadek zewnętrznej temperatury.

Porównywanie energii pomiędzy garbarniami jest dość trudne, ponieważ występują różnice w procesie technologicznym jak i pod względem zmechanizowania zakładu. Dlatego najlepszym sposobem jest monitorowanie energii i jej zużycia dla każdego z etapów technologicznych z osobna.

Podsumowując największe zanieczyszczenia pochodzące z garbarni pojawiają się bardzo problematyczne i trudne do ograniczenia aspekty:

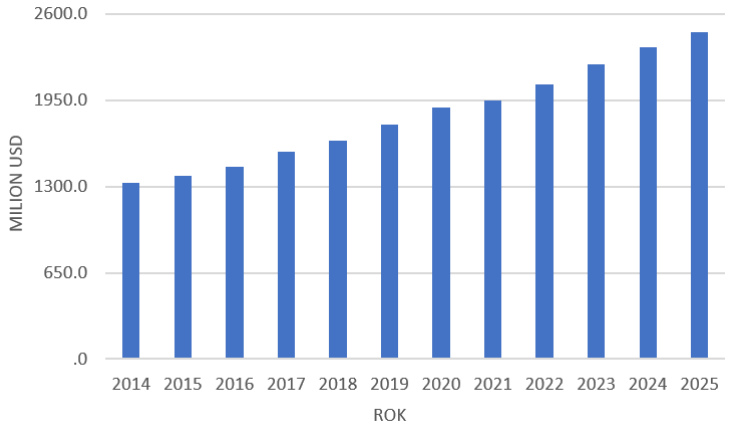
- zużycie wody
- oczyszczanie ścieków
- zużycie energii (ograniczenie lub wykorzystanie alternatywnych źródeł energii)
- odpady stałe (chromowe i niechromowe)
- zanieczyszczenia powietrza [4].

12.2 Produkcja skór w branży Automotive

12.2.1 Rynek motoryzacyjny

Skóra samochodowa pojawiła się na rynku od momentu powstania pojazdów napędzanych silnikiem. Jest ubocznym produktem przemysłu mięsnego i ma wysoki potencjał, zwłaszcza w pojazdach luksusowych, pomimo dostępności znacznie tańszych wariantów. Pomimo, że zawsze była bardzo ważną częścią światowego przemysłu skórzanego, w ciągu ostatnich lat stała się również jednym z najbardziej zaawansowanych sektorów. Stało się to za sprawą wysokich standardów, które są wymagane, zarówno w zakresie specyfikacji jak i wydajności. Skóra naturalna zdominowała rynek światowy skórzanych wnętrz samochodowych w 2017r. pod względem wartości. View Research, będący sektorem wyposażenia samochodów, w 2017 roku został wyceniony na 28 320 mln USD. Przewiduje się, że w latach 2018-2025 uda się osiągnąć poziom 6,5% rocznie. Porównując ogólną wartość wnętrza samochodu rozkłada się ona następująco pomiędzy dwie składowe: skóra 55,3% i podłoże sztuczne 44,7%. Popyt na skóry samochodowe w 2017r. wyceniono na 15 667,2 mld USD, a prognozy wskazują, że w 2025 r. osiągnie 24 652,7 mln USD,

co zostało przedstawione na Rys. 12.10 z dokładnym podziałem na poszczególne lata .



Rys. 12.10: Przychód roczny ze skór samochodowych w latach 2014-2025.

W roku 2016 sektor motoryzacyjny przejął 15-18% wszystkich skór produkcji światowej. W 2020 nastąpił wzrost do około 25%. Wzrost ten determinowany jest poprzez rozwój gospodarki wschodniej Azji i Pacyfiku, jak na przykład Chin czy Indii, co wiąże się z przeniesieniem części produkcji motoryzacyjnej do tych państw.

Rozwój rynku wnętrz samochodowych jest uzależniony od zapotrzebowania na sztuczne i syntetyczne podłoża, które jest coraz większe. Powodem jest konkurencyjna cena jak i znacznie łatwiejsza produkcja tego typu materiałów w porównaniu do skóry.

Biorąc pod uwagę całą branżę skórzaną to artykuły do tapicerki samochodowej są najbardziej wymagającymi produktami ze wszystkich produkowanych artykułów skórzanych [24].

12.2.2 Właściwości skór Automotive

Skóra wykorzystywana we wnętrzach samochodów określana jest mianem produktu luksusowego, wszystkie wymagania dotyczące produkcji skór stawiane firmom produkującym je są bardzo rygory-

styczne. Dotyczy to jakości skór jak i aspektów środowiskowych i podejścia do ekologii.

Klienci odbierają ją jako naturalny i trwały produkt, który musi spełniać wymagające kryteria określone przez producentów oryginalnego wyposażenia (OEM). Stworzyli oni profile wydajności, w których zawarte zostało zużycie mechaniczne, niska emisja, ochrona przed warunkami atmosferycznymi (między innymi odporność na działania światła, możliwość odbijania światła, dzięki czemu siedzenia nie nagrzewają się w lecie) czy zrównoważona produkcja. Specyfikacje i wymagania stawiane artykułom motoryzacyjnym są najsurowsze w branży.

Dodatkowo skóra użyta w branży Automotive musi charakteryzować się walorami estetycznymi takimi jak odpowiednia miękkość, ścisłość, kolor, dotyk czy zapach. Bardzo istotna jest niska zawartość lotnych związków organicznych (LZO). Ze względu na bardzo zaostrzone normy, przetworzone skóry są powtarzalne. Skóra jest produktem znacznie bardziej zrównoważonym niż materiały sztuczne. Dodatkowo cechuje się znacznie większą trwałością.

Zarówno wysoka ochrona jak i wytrzymałość są osiągalne dzięki polimerowym powłokom. Niestety ulega ona podczas eksploatacja uszkodzeniom i zużyciu, przez co pęka i się ściera. Możliwe jest sprawdzenie zachowania się skóry po wielu latach użytkowania. Takie testy przeprowadzane są w laboratoriach, gdzie skóra poddawana jest badaniom starzeniowym. Polega to na wystawianie po wykończeniu skóry na kilka cykli światła przy wcześniej określonej wilgotności i temperaturze. Daje to możliwość sprawdzenia skóry pod względem mechanicznych właściwości, głównie jak elastyczna jest nałożona powłoka [31]. W wykończaniu skór tapicerki samochodowej wykorzystuje się PUD, które wspomagają stabilność długoterminową [25].

W ostatnich latach zaobserwowano na rynku motoryzacyjnym rozpowszechnianie skór garbowania „wet-white" , co oznacza, że chrom został zastąpiony garbnikami roślinnymi, a także użyto do procesu aldehyd glutarowy. . Dzięki temu takie skóry są biodegradowalne, gdyż nie został zastosowany w nich poza chromem żaden metal utrudniający ten proces [26].

12.2.3 Generowanie odpadów podczas produkcji tapicerki samochodo

W ostatnich dziesięcioleciach nastąpił gwałtowny rozwój branży motoryzacyjnej co pociągnęło za sobą wzmożoną produkcję odpadów poprodukcyjnych i poeksploatacyjnych. Dlatego tak ważne stało się utworzenie i utrzymanie zasad prawnych i technicznych dotyczących zarządzania i unieszkodliwiania powstałych odpadów. Koszty odbioru odpadów garbarskich jak i odpadów po oczyszczaniu ścieków w formie osadu wzrosły w przeciągu roku nawet dwukrotnie. Dodatkowo staje się problematyczne znalezienie odbiorcy tych odpadów, gdyż samo składowanie zajmuje zbyt wiele powierzchni a także powoduje zanieczyszczanie środowiska, natomiast przetwarzanie ich w cementowniach stanowi niewielki procent całkowitej ilości generowanych odpadów.

12.2.4 Emisja lotnych związków organicznych w branży motoryzacyjnej

Zgodnie z europejskimi przepisami LZO są związkami o prężności pary 0,01 kPa lub większej w temperaturze 293,15K. Oznacza to, że już w temperaturze pokojowej mają wysoka prężność pary, a dodatkowo część z nich może być toksyczna dla ludzi i środowiska. Dlatego tak ważna staje się eliminacja bądź przynajmniej minimalizacja.

Ponad 20 lat temu do przemysłu motoryzacyjnego została wprowadzona kontrola emisji lotnych związków organicznych. Test Fogging był jednym z pierwszych kontrolujących LZO. Opisuje on parowanie substancji chemicznych z danego materiału stosowanego do wnętrz pojazdu pod wpływem ciepła. Dochodzi do kondensacji tych substancji na szybach samochodowych, co powoduje powstawanie mętnej warstwy zwanej zamgleniem.

Sam test polega na ogrzewaniu wcześniej wysuszonej skóry w szklanej zlewce. Substancje odparowane skraplają się na szklanej płycie, zw. zamglenie reflektometryczne bądź na aluminiowej folii – zamglenie grawimetryczne.

Zamglenie refleksometryczne określa w określonych warunkach czasu

i temperatury, stopień zamglenia na płytce szklanej przez odparowujące substancje. Wyrażone w %. Zamglenie grawimetryczne jest określeniem substancji skondensowanych i odparowanych na aluminiowej folii w określonych warunkach czasu i temperatury. Wyrażone w mg.

Każdy z testów mających na celu określenie LZO ma swoje zalety i wady, dlatego też nie można porównywać otrzymanych różnymi metodami wyników. Dla minimalizacji LZO proponuje się pracę z produktami, które zawierają śladowe ilości takich związków. W celu zatwierdzenia należy dokładnie zbadać każdy produkt. W lotnych związkach organicznych obecne są dwa związki: aldehyd octowy i formaldehyd. Dla każde zostały przygotowane normy. W przypadku formaldehydu wartości wymagane wynoszą ≤ 3 ppm (ISO 17226-3) lub ≤ 10 ppm (ISO 17226-1), natomiast dla aldehydu octowego $\leq 0{,}6$ mg [27].

12.3 Proces rozkładu skór niegarbowanych

12.3.1 Drobnoustroje w garbarstwie

Drobnoustroje należą do grupy jedno- lub wielokomórkowych organizmów, w postaci zarodników występują w dużej ilości w przyrodzie. Ze względu na to, iż skóra surowa jest doskonałą pożywką dla nich, są bardzo obficie rozsiane na jej powierzchni.

Podział drobnoustrojów przedstawia się następująco:

1. Bakterie
2. Drożdże
3. Pleśniaki
4. Grzybki skórne
5. Wirusy

Bakterie są jednokomórkowcami, które w korzystnych dla siebie warunkach mogą wydawać pokolenie nawet co 30 minut. Po przeprowadzeniu prostych obliczeń pojawia się wynik, pokazujący iż z jednej bakterii po dwóch dniach powstaje ok. 280mln bakterii. Z powodu tak szybkiej rozrodczości bakterie są bardzo niebezpieczne

dla surowej skóry. Dochodzi jednak do pewnego hamowania rozwoju bakterii, w wyniku antagonizmów pomiędzy różnymi rodzajami drobnoustrojów. Rozrodczość bakterii jest również zależna od temperatury otoczenia, odczynu środowiska, ilości pokarmu czy środków konserwacji skóry. Jednak te drobnoustroje mają zdolność przechodzenia w spory, przetrwalniki. Dzięki temu potrafią dostosowywać się do panujących warunków i uodparniać na niekorzystne środowisko, jak na przykład wysoka temperatura czy obecność środków chemicznych. W takiej postaci bakteria może przetrwać wiele lat, w momencie poprawy warunków przechodzi w wegetatywną postać [1].

Wyróżniamy dwa rodzaje bakterii:

1. Aerobowe- tlenowe; potrzebujące tlenu pochodzącego z powietrza do rozwoju
2. Anaerobowe- beztlenowe; rozwijające się bez powietrza, a tlen pobierają z rozkładanej substancji
3. Bakterie rozwijające się w warunkach tlenowych i beztlenowych

Skóra żywych organizmów posiada na swojej powierzchni wiele bakterii, jednak nie są one aktywne. Dopiero podczas zranienia czy uszkodzenia staje się ona miejscem, gdzie bakterie mogą się rozwijać. Na martwiej skórze, gdzie pojawiają się ognojenia i pozostałości krwi zaczynają uaktywniać się bakterie proteolityczne, czyli zdolne do rozkładu białek. Tego typu bakterie zaczynają atakować martwą skórę i powodują jej gnicie. Proces ten jest bardzo skomplikowany pod względem chemicznym, ale ostatecznym produktem jest amoniak. Dodatkowo istnieją inne rodzaje bakterii, które wydzielają pigmenty, stąd na martwej skórze często pojawiają się plamy od strony mizdry (strona nieowłosiona).

Drożdże nie maja wpływu na skórę. Natomiast wprowadzane ze skórą do kąpieli procesowych mogą wywoływać w nich zmiany. Mają zdolność fermentowania cukrów. Rodzaje drożdży możemy podzielić na hodowlane, które stosowane są między innymi w piwowarstwie oraz dzikie. Różnią się one kształtem [21].

Pleśniaki należą już grzybów, czyli do grup wyżej zorganizowanych niż drożdże czy bakterie. Występują w postaci nitek. Możliwe jest ich zaobserwowanie na skórze w postaci nalotów. Mogą rozwijać się powierzchniowo tylko bądź wewnętrznie, czyli w całej masie skóry. Powoduje to, że potrafią one wnikać w głąb komórek a nie tylko je otaczać. Bardzo często występują w rozworach wodnych roślinnych garbników, co powoduje ich rozkładanie. Zużywają także kwasy, co doprowadza do rozwoju bakterii. Te drobnoustroje pojawiają się na skórach surowych już zakonserwowanych sposobem suchym, a także wyprawionych magazynowanych w wilgotnych pomieszczeniach [1].

Grzybki skórne są drobnoustrojami wielokomórkowymi, budową podobne do pleśniaków. Różnią się jednak tym, iż te żyją na zwierzętach żywych, stąd też mogą powodować choroby skóry, co później przekłada się na surowiec i jego jakość. Ten rodzaj grzybów po zdjęciu skóry ze zwierzęcia i zakonserwowaniu jej wymiera.

Wirusy należą do najmniejszych znanych mikroorganizmów. Powodują różnego rodzaju choroby, z których znaną jest tzw. pryszczyca. Ich odporność na środki chemiczne jest bardzo wysoka, a najskuteczniejsze w ich eliminacji jest formalina lub silne zasady.

12.3.2 Drobnoustroje i ich rozwój

Do prawidłowego rozwoju drobnoustrojów bardzo ważne są następujące czynniki:

1. Wilgotność
2. Temperatura
3. Światło
4. Odczyn środowiska
5. Środki chemiczne

Każdy rodzaj drobnoustrojów wymaga optymalnej temperatury do swojego prawidłowego wzrostu. Tolerancja dla różnych grup waha się od 20 do 70 °. Dodatkowo bardzo istotna jest wilgotność w pożywce. Dla drożdży czy bakterii minimalna wilgotność wynosi 30%, natomiast dla pleśniaków wystarczy już 12%. Jeżeli ta wartość nie zostanie osiągnięta rozwój drobnoustrojów zostaje zatrzymana.

W przypadku odczynu środowiska bakterie najlepiej rozwijają się przy odczynie obojętnym lub słabo zasadowych, zaś drożdże i pleśniaki potrzebują środowiska kwaśnego. Bakterie gnilne występujące na skórze surowej rozwijają się przy pH 6,5-8,4. Czerwone plamy na skórze surowej od strony mizdry tworzą bakterie, które są odporniejsze na środowisko zasadowe. Sama surowa skóra zdjęta ze zwierzęcia ma pH obojętne bądź słabo kwaśne, a dodanie sody podczas procesu konserwacji z solą, powoduje, że bakterie nie mogą się rozwijać.

Światło rozproszone nie jest szkodliwe dla bakterii, a znane są bakterie proteolityczne, gdzie w takich warunkach dochodzi do wzmożonego ich wzrostu. Natomiast światło nadfioletowe działa już hamująco na rozwój bakterii, stąd lampy z takim światłem są wykorzystywane w magazynach skór surowych. Środki chemiczne działają na drobnoustroje bardzo negatywnie, hamując ich wzrost i niszcząc je. Sól jest jednym z nich, powodując odwadnianie ich poprzez osmozę i doprowadzając do postaci nieaktywnej bądź wręcz niszcząc je całkowicie [21].

Kwasy i zasady należące do mocnych niszczą wszystkie drobnoustroje, tak samo jak etanol 50-cio procentowy, który powoduje odwadnianie ich. Etanol 100-procentowy jest już słabszy i nie ma tak niszczącego działania. Powodem jest to, że ścina białko znajdujące się na powierzchni surowej skóry, przez co nie przedostaje się do wnętrza. Fenole i ich pochodne czy formaldehyd jest równie niszczący ze względu na reakcje białkiem plazmy komórki [1].

Drobnoustroje antagonistyczne to takie, które niszczą się wzajemnie lub przynajmniej hamują swój rozwój. Wynika to z wydzielin, które reagują z innym rodzajem drobnoustrojów i ich białkiem [1].

12.3.3 Enzymy

Bakterie powodują hydrolityczny rozpad białek, węglowodanów oraz tłuszczów. Nie działają jednak na nie bezpośrednio a wykorzystują do tego enzymy, tzw. fermenty. Enzymy są białkowymi substancjami produkowanymi przez drobnoustroje, o działaniu katalizującym, ale w wąskim zakresie.

Pierwszym rodzajem enzymów są hydrolazy, które odpowiedzialne są za hydrolizę poszczególnych substancji. Wśród nich możemy wymienić: proteazy, lipazy, karbohydrazy. Proteazy są enzymami przyspieszającymi poszczególne stadia hydrolizy białka. Lipazy powodują rozpad hydrolityczny tłuszczów, natomiast karbohydrazy są odpowiedzialne za hydrolizę złożonych węglowodanów [1]. Desmolazy są to enzymy, które przyspieszają reakcje chemiczne typu utleniania, redukcji.

W przypadku skór surowych istotne są proteazy, które swoją rolę odgrywają podczas początkowych etapów hydrolitycznego rozpadu białka. Rozkładają one białka do peptonów, a następnie te do wielopeptydów. Proteazy możemy podzielić w zależności od środowiska, w którym funkcjonują na:

1. Trypsyna – środowisko zasadowe
2. Katepsyna – środowisko obojętne (występuje w dużej ilości w skórze)
3. Pepsyna – środowisko kwaśne.

Za kolejny etap rozkładu odpowiedzialne są peptydazy, powodują one redukcję wielopeptydów do peptydów prostych, a następnie do aminokwasów. W skórze surowej można znaleźć lipazy, a w naskórku czy warstwie podskórnej amylazę, która rozkłada krochmal [1].

W wodzie enzymy tworzą koloidalne roztwory, dzięki czemu dają się wysalać. Oznacza to, że koagulują z konkretnymi solami, głównie z NaCl, po czym przestają działać. Enzymy mają jednak zdolność uszkadzania skóry, bez względu na to czy pochodzą ze środowiska zewnętrznego czy z wnętrza. Skóra jeżeli nie jest zakonserwowana lub tylko w niewielkim stopniu zabezpieczona może ulegać działaniu drobnoustrojów, których na skórze żywej nie było. Ich działanie degradacyjne na skórę opiera się na wydzielaniu przez nie właśnie enzymów. Enzymy, które były wytwarzane za życia zwierzęcia również mogą przyczyniać się do jej rozkłady po zdjęciu na zasadzie samotrawienia, czyli tak zwanej autolizy. Polega ona na tym, iż enzymy, które już nie otrzymują tlenu dostarczanego przez krew atakują włókniste białka skóry i doprowadzają do ich rozkładu. Speptyzowany kolagen zostaje rozpuszczony koloidalnie w wodzie.

Najlepsze warunki dla samotrawienia to temperatura 40°, pH słabo kwaśne i brak tlenu. Schemat hydrolitycznego rozpadu białek skóry przebiegający przy udziale enzymów wygląda następująco:

Białka → Peptony → Wielopeptydy → Dwupeptydy → Aminokwasy → Amoniak

Peptony maja zdolność rozpuszczania się koloidalnego w wodzie, zaś aminokwasy tworzą roztwory właściwe. Jednak aby enzym miał właściwości rozkładające musi występować przynajmniej w minimalnym stężeniu dającym mu tą zdolność. Stężenie to nazywane jest stężeniem progowym. Natomiast cechą wspólną wszystkich enzymów jest zakres temperatur 0-70 °, poza którym nie działają w ogóle, a jego przekroczenie powoduje, że są nieaktywne. Przekroczenie górnej granicy temperatury doprowadza do nieodwracalnego zniszczenia enzymu [1].

12.3.4 Powody konserwacji skór surowych

Skóra surowa przebywa długą drogę od ściągnięcia jej ze zwierzęcia do garbarni. Biorąc pod uwagę transport, magazynowanie, szczególnie w porze letniej, konserwacja jest jedną z najistotniejszych etapów w przygotowaniu skóry do dalszej obróbki. Dodatkowo ważny jest czas, po którym skóra została zakonserwowana po uboju, opóźnienie może spowodować, że mikroflora, głównie w postaci proteolitycznych bakterii, uszkodzi surowiec. Doprowadzi to do gnicia skóry co spowoduje, że nie będzie możliwe jej przetwarzanie. Bardzo często surowa skóra po zdjęciu musi poczekać w magazynie na wzięcie jej do produkcji, ze względu na ograniczone zdolności produkcyjne garbarni [32].

Skóra surowa zaraz po zdjęciu ze zwierzęcia jest idealną pożywką dla bakterii. Powodem jest to, iż wilgotność jej jest w granicach 50-70%, a temperatura jest zbliżona do ciała, inaczej tracą czynności życiowe, a tym samym nie wydzielają enzymów. Dodatkowo nie dochodzi do autolizy, gdyż enzymy własne również są w fazie nieaktywnej. Skórę po uboju należy schłodzić, co zapobiega tworzeniu plam solnych. Następnie taką skórę się oczyszcza się, po czym przechodzi do konserwacji [1].

12.4 Wnioski

Przemysł garbarski wytwarza odpady będące jednymi z najtrudniejszych do zagospodarowania co stanowi problem na skalę globalną. Ich złożony charakter, a także ilość strumieni odpadów, różniących się od siebie silnie właściwościami, spowodowało, że na chwilę obecną nie ma racjonalnego sposobu ich utylizacji. Powyższa praca pokazuje również, że ilość poszczególnych odpadów w przeliczeniu miesięcznym dla jednej garbarni jest bardzo duża- dochodzi nawet do kilkudziesięciu ton. Niestety dotychczas prowadzone badania nad ich przetwarzaniem nie przyniosły zadowalających wyników dla zakładów garbarskich. Ilość odpadów, a także ich szkodliwość powoduje, ze garbarnie nazywane „trucicielami" środowiska mają coraz większy problem do rozwiązania. Powodem jest zawartość rakotwórczego chromu (+6) a także innych związków chemicznych, dodawanych w poszczególnych etapach przetwarzania. Skład oraz postać użytych środków chemicznych powoduje, że każdy z odpadów wymaga indywidualnego podejścia. Obecnie stosowany sposób radzenia sobie z odpadami poprzez „składowanie" jest najbardziej niekorzystną formą dla środowiska, natomiast wciąż najbardziej powszechną. Dodatkowo zauważamy ciągły i dość silny wzrost zapotrzebowania na skórę samochodową, który będzie w najbliższych latach miał miejsce. Skutkiem tego będzie dynamiczna eskalacja problemów z odpadami garbarskimi. Powyższy przegląd literaturowy ma na celu pokazanie skali problemu i zwrócenie uwagi, że skóra na swojej drodze od ściągnięcia ze zwierzęcia do gotowych siedzeń w samochodzie pozostawia znaczący ślad w środowisku, którym są grupy ogromnych ilości odpadów na każdym etapie jej przetwarzania. Dlatego tak ważna jest cyrkularna gospodarka związana z odpadami, która umożliwi pozyskanie cennych składników, a także wytworzenia wartościowych produktów z jednoczesną redukcją produkowanych odpadów. Na chwilę obecną bardzo ważne jest znalezienie sposobu ich utylizacji, który pozwoli rozwiązać problemy środowiskowe w przemyśle garbarskim. Jest to istotne w równym stopniu dla ekologii jak i dla garbarni, które same nie potrafią poradzić sobie z narastającym problemem.

Bibliografia

[1] J Iwanowski and T Persz. *Garbarstwo cz. I.* Wydawnictwo Przemysłu Lekkiego i Spozywczego: Warszawa, Poland, 1966.

[2] Huiyan Jiang, Junsheng Liu, and Wei Han. The status and developments of leather solid waste treatment: A mini-review. *Waste Management & Research*, 34(5):399–408, 2016.

[3] Praca zbiorowa. *Poradnik garbarza.* PWT, Warszawa, Poland, 1953.

[4] F Duperriez, T Poncet, R Lety, I Kulińska, K Kosińska, M Mikulska, H Prygiel, and T Sadowski. Przewodnik wprowadzania systemu zarządzania środowiskowego w garbarniach. *AFNOR: Paris, France*, pages 11–13, 2010.

[5] J Iwanowski and T Persz. *Garbarstwo cz. II.* Wydawnictwo Przemysłu Lekkiego i Spozywczego: Warszawa, Poland, 1966.

[6] Jing Hu, Zuobing Xiao, Rujun Zhou, Weijun Deng, Mingxi Wang, and Shuangshuang Ma. Ecological utilization of leather tannery waste with circular economy model. *Journal of Cleaner Production*, 19(2-3):221–228, 2011.

[7] RO Oruko, R Selvarajan, HJO Ogola, JN Edokpayi, and JO Odiyo. Contemporary and future direction of chromium tanning and management in sub saharan africa tanneries. *Process Safety and Environmental Protection*, 133:369–386, 2020.

[8] Katarzyna Fela, Krystyna Wieczorek-Ciurowa, Michał Konopka, and Zenon Woźny. Present and prospective leather

industry waste disposal. *Polish Journal of Chemical Technology*, 13(3):53–55, 2011.

[9] EC-IPPC. Integrated pollution prevention control, 2013.

[10] M Bosnic, J Buljan, and RP Daniels. Pollutants in tannery effluents-definitions and environmental impact-limits for discharge into water bodies and sewers. *United Nations Industrial Development Organization*, 2000.

[11] O Rodziewicz. *Podstawy technologii garbarstwa*. WSI, Radom, Poland, 1984.

[12] KJ Sienkiewicz. *Fizykochemia wyprawy skór*. WNT, Warszawa, Poland, 1986.

[13] J Kanagaraj, T Senthilvelan, RC Panda, and S Kavitha. Eco-friendly waste management strategies for greener environment towards sustainable development in leather industry: a comprehensive review. *Journal of Cleaner Production*, 89:1–17, 2015.

[14] David Blowes. Tracking hexavalent cr in groundwater.(perspectives: environmental chemistry). *Science*, 295(5562):2024–2026, 2002.

[15] Piotr Celary and Jolanta Sobik-Szołtysek. Vitrification as an alternative to landfilling of tannery sewage sludge. *Waste Management*, 34(12):2520–2527, 2014.

[16] Mozhiarasi Velusamy, Bhagiratha Chakali, Sathish Ganesan, Farha Tinwala, and Srinivasan Shanmugham Venkatachalam. Investigation on pyrolysis and incineration of chrome-tanned solid waste from tanneries for effective treatment and disposal: an experimental study. *Environmental Science and Pollution Research*, 27:29778–29790, 2020.

[17] J Kanagaraj, KC Velappan, NK Babu, and S Sadulla. Solid wastes generation in the leather industry and its utilization for cleaner environment-a review. 2006.

[18] Diego Navarro, Jianhui Wu, Wei Lin, Pere Fullana-i Palmer, and Rita Puig. Life cycle assessment and leather production. *Journal of Leather Science and Engineering*, 2:1–13, 2020.

[19] Donald G Barceloux and Donald Barceloux. Vanadium. *Journal of Toxicology: Clinical Toxicology*, 37(2):265–278, 1999.

[20] Golam Rabbani, Baki Billah, Anil Giri, Sarder M Hossain, Ahmmad I Ibne Mahmud, Bilkis Banu, Ulfat Ara, and Sheikh M Alif. Factors associated with health complaints among leather tannery workers in bangladesh. *Workplace Health & Safety*, 69(1):22–31, 2021.

[21] Karel Kolomaznik, Milan Adamek, I Andel, and M Uhlirova. Leather waste? potential threat to human health, and a new technology of its treatment. *Journal of hazardous materials*, 160(2-3):514–520, 2008.

[22] Semih Şaşmaz, Bağdagül Karaağaç, and Nurseli Uyanık. Utilization of chrome-tanned leather wastes in natural rubber and styrene-butadiene rubber blends. *Journal of Material Cycles and Waste Management*, 21(1):166–175, 2019.

[23] Elena Bañón, Antonio Marcilla, Angela N García, Pascual Martínez, and Milagros León. Kinetic model of the thermal pyrolysis of chrome tanned leather treated with naoh under different conditions using thermogravimetric analysis. *Waste management*, 48:285–299, 2016.

[24] Suelem Daiane Ferreira, Janaina Junges, Bianca Scopel, Christian Manera, Eduardo Osório, Ivan Pedro Lazzarotto, and Marcelo Godinho. Steam gasification of biochar derived from the pyrolysis of chrome-tanned leather shavings. *Chemical Engineering & Technology*, 42(12):2530–2538, 2019.

[25] G Durai and M Rajasimmam. Biological treatment of tannery waterwaste: a review. *Environmental Science & Technology*, 4, 2011.

[26] Abdelazeem Sh Sallam, Adel RA Usman, Hassan A Al-Makrami, Mohammad I Al-Wabel, and Abdulrasoul Al-Omran. Environmental assessment of tannery wastes in relation to dumpsite soil: a case study from riyadh, saudi arabia. *Arabian Journal of Geosciences*, 8:11019–11029, 2015.

[27] Olcay Tünay, Isik Kabdasli, Derin Orhon, and Esra Ates. Characterization and pollution profile of leather tanning industry in turkey. *Water Science and Technology*, 32(12):1–9, 1995.

[28] LS Clesceri, CG Greenberg, and AD Eaton. *Standard methods for the examination of water and wastewater*, volume 20. American Public Health Association., 1999.

[29] Andrea Luca Tasca and Monica Puccini. Leather tanning: Life cycle assessment of retanning, fatliquoring and dyeing. *Journal of Cleaner Production*, 226:720–729, 2019.

[30] W Domanski and J Surgiewicz. Chemical hazards in the tanning process. *Bezpieczenstwo pracy – Nauka i Praktyka*, 4, 2001.

[31] Allen Skaja, Dilhan Fernando, and Stuart Croll. Mechanical property changes and degradation during accelerated weathering of polyester-urethane coatings. *JCT research*, 3:41–51, 2006.

[32] Alois Orlita. Microbial biodeterioration of leather and its control: a review. *International Biodeterioration & Biodegradation*, 53(3):157–163, 2004.

Rozdział 13

Dobór i optymalizacja płaszczowo-rurowego parownika typu NKN w mikrosiłowni ORC o mocy 30kWe

Artur Grajewski

Institute of Fluid Flow Machinery, Polish Academy of Sciences,
Fiszera 14, 80-231 Gdansk, Poland

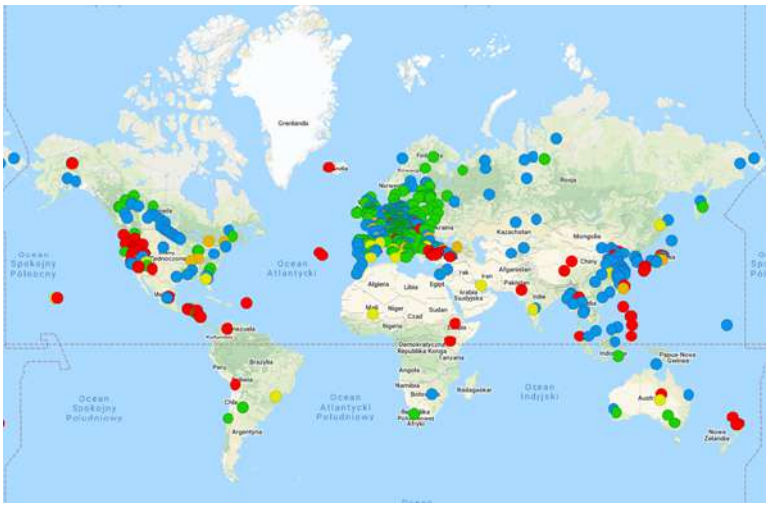
Streszczenie

W różnych gałęziach przemysłu zarówno w kraju jak i na świecie ogromne ilości ciepła odpadowego w postaci gorących spalin emitowane są do atmosfery. Jednym z możliwych wariantów zagospodarowania tego typu odpadowych źródeł ciepła są instalacje ORC, które cieszą się coraz większą popularnością. Tania energia odpadowa z jednej strony, i wciąż rosnące koszty wytwarzania energii z jednostek opalanych paliwami kopalnymi z drugiej strony, przyczyniają się do intensyfikacji prac nad zaawansowanymi siłowniami ORC, w których jednym z najważniejszych elementów są wymienniki ciepła. Prawidłowy dobór konstrukcji wymienników ciepła i ich późniejsza optymalizacja jest niezmiernie ważna w kontekście poprawnej pracy instalacji, ale także kosztów budowy oraz opłacalności całego przedsięwzięcia. W pracy zaproponowano nowy typ parownika do mikrosiłowni ORC z bezpośrednim grzaniem czynnika roboczego. Parownik ten został zaprojektowany w układzie płaszczowo-rurowym typu NKN, w którym spaliny płyną rurami a czynnik roboczy odpowiednio ukształtowanym płaszczem. Dla tak zaprojektowanego parownika przeprowadzono szereg obliczeń, w których określono parametry cieplno-przepływowe wymiennika w zależności od liczby rurek. Optymalnym rozwiązaniem okazał się wymiennik NKN składający się z 344 gładkich rur procesowych w rozmiarze ϕ20x1 mm i długości 3000 mm. Całkowita powierzchnia wymiany ciepła wyniosła 63 m^2 a uzyskana moc cieplna 217 W.

13.1 Wprowadzenie i cel pracy

Instalacje ORC są coraz częściej stosowanym rozwiązaniem zagospodarowania wysoko i średniotemperaturowych źródeł energii, nad czym pracuje wiele firm z całego świata. Instalacje te cieszą się powodzeniem w skali zarówno mikro jak i makro, a znajdują swoje zastosowanie w układach gdzie spalana jest biomasa [1–3], w układach geotermalnych [4–6], w układach odzysku ciepła z gorących spalin [7–10] często o temperaturze przekraczającej 400°C, a także w układach współpracujących z turbiną gazową lub innych aplikacjach charakteryzujących się wysoką temperaturą ciepła odpa-

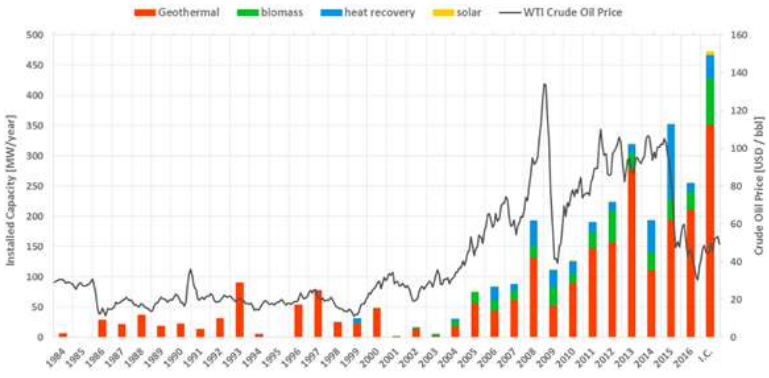
dowego [11, 12]. Mapę układów ORC zainstalowanych na świecie przedstawiono na Rys. 13.1. Jak można zauważyć dominują instalacje wykorzystujące ciepło pochodzące z biomasy oraz ciepło odpadowe. Całkowita liczba układów zainstalowanych na świecie wynosi 1149 [13]. Najwięcej instalacji powstało kolejno w Europie, Ameryce Północnej i Azji. Pojedyncze jednostki pracują w Ameryce Południowej, Australii oraz Afryce.



Rys. 13.1: Mapa instalacji ORC na świecie wraz z zaznaczonym rodzajem źródła ciepła: kolor zielony – biomasa, kolor czerwony – geotermia, kolor żółty – energia słoneczna, kolor niebieski i pomarańczowy – ciepło odpadowe [13].

W ostatnim czasie obserwuje się duże zainteresowanie budową instalacji ORC, co przedstawiono na wykresie słupkowym przyrostu mocy zainstalowanej, Rys. 13.2 Ostatni słupek obrazuje moc instalacji w budowie. Wartość przyrostu mocy od roku 2011 nie spadła poniżej 180 MW rocznie. Największy udział mocy przypada na układy wykorzystujące ciepło geotermalne, a więc układy charakteryzujące się stosunkowo dużą mocą zainstalowaną w pojedynczej jednostce.

W układach ORC o mocy przekraczającej kilkanaście kWe bardzo często stosowane są wymienniki płaszczowo-rurowe [9, 15–20]. Od-



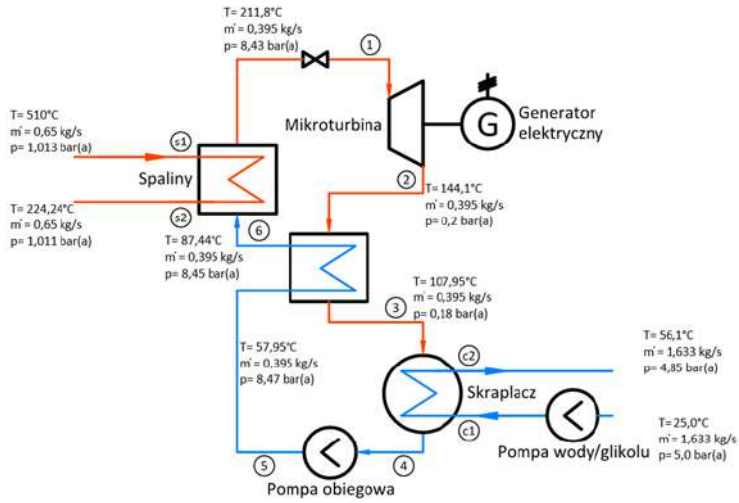
Rys. 13.2: Roczny przyrost łącznej mocy zainstalowanej instalacji ORC w latach 1984-2016 [14].

powiednio zaprojektowane tego typu urządzenia, charakteryzują się solidną, spawaną budową, pozwalającą na pracę w warunkach zarówno wysokiej temperatury, jak i wysokiego ciśnienia roboczego. Tego samego nie można powiedzieć o wymiennikach płytowych, gdzie grubości chociażby płyt grzewczych zazwyczaj nie przekraczają wartości 0,5 mm. Wymienniki płaszczowo-rurowe, w odróżnieniu od wymienników płytowych, zazwyczaj projektuje się indywidualnie, dla konkretnych parametrów pracy. W przypadku wymienników płytowych, dobiera się urządzenie z danego typoszeregu, gdzie geometria płyty jest już narzucona i istnieje jedynie możliwość doboru odpowiedniej ilości płyt grzewczych, co stanowi pewnego rodzaju ograniczenie.

13.2 Analizowana mikrosiłownia

Przedmiotem analizy jest mikrosiłownia ORC przedstawiona na Rys. 13.3, której celem jest wykorzystanie ciepła odpadowego spalin, o temperaturze 510°C.

Układ ten został zaprojektowany przy założeniu, że gorące spaliny ogrzewają czynnik roboczy, tym samym powodując odparowanie z jednoczesnym przegrzaniem par do temperatury 211,8°C (1). Spaliny na wylocie z wymiennika wychładzają się do temperatury około



Rys. 13.3: Schemat układu ORC: s1 – spaliny gorące, s2 – spaliny wychłodzone, 1 – świeża para czynnika roboczego, 2 – przepracowana para czynnika roboczego, 3 – przepracowana para po schłodzeniu w regeneratorze, 4 – skropliny (niskie ciśnienie), 5 – chłodna ciecz opuszczająca pompę (wysokie ciśnienie), 6 – ogrzana ciecz zasilająca parownik (wysokie ciśnienie), c1 – ciecz chłodząca (woda/glikol), c2 – ogrzana woda/glikol.

224,24°C (s2) i następnie odprowadzane są do atmosfery. Czynnik roboczy w postaci par toluenu o temperaturze 211,8°C (1) pozbawiony w zupełności fazy ciekłej, kierowany jest na łopatki mikroturbiny napędzającej generator elektryczny o mocy znamionowej 30 kWe.

Następnie, rozprężone pary czynnika roboczego o temperaturze 144,1°C (2) kierowane są do regeneratora, gdzie ulegają wychłodzeniu do temperatury 107,95°C (3) podgrzewając płynący po drugiej stronie czynnik roboczy do temperatury 87,44°C (6). Wychłodzony w regeneratorze toluen trafia do skraplacza, gdzie ulega kondensacji. Przemiana fazowa następuje na skutek odbioru ciepła przez roztwór glikolu, który zasila obieg pomocniczy. Wykroplony w skraplaczu toluen przy pomocy głównej pompy obiegowej tłoczony jest po-

przez regenerator do parownika, tym samym rozpoczynając kolejny cykl obiegu zamkniętego. Parametry pracy parownika zamieszczono w Tab. 13.1.

Tab. 13.1: Parametry robocze parownika.

	Strona Rury	**Strona Płaszcz**
Medium	Spaliny (CO_2, O_2, N_2, H_2O)	Toluen ($C_6H_5CH_3$)
Temperatura wlot/wylot [°C]	510/ 222	87,44/211,8
Przepływ [kg/h]	2340	1422
Ciśnienie robocze [bar (abs)]	1,04	8,45
Dopuszczalny spadek ciśnienia [kPa]	0,4	3,0
Moc cieplna [kW]	∼217	

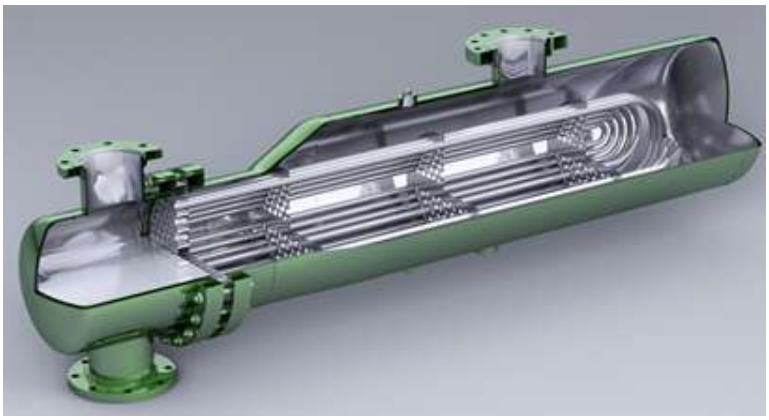
Światowi producenci wymienników ciepła posługują się na co dzień zazwyczaj jednym z dwóch najbardziej zaawansowanych programów służących do doboru wymienników płaszczowo- rurowych. Pierwszym z nich jest Aspen Exchanger Design&Rating [21], drugim HTRI [22]. W obu tych programach można zasymulować niemal każdy proces wymiany ciepła pomiędzy dwoma dowolnie wybranymi płynami, począwszy od wody, pary wodnej, glikolu, czynników chłodniczych, a skończywszy na różnego rodzaju węglowodorach czy też mieszaninach gazów i cieczy. Obszerne bazy płynów dają szerokie możliwości symulacji procesów wymiany ciepła, znajdujących zastosowanie w energetyce zawodowej, przemyśle rafineryjnym, przemyśle chemicznym, przemyśle papierniczym, przemyśle farmaceutycznym, chłodnictwie, i wielu innych. Oba wspomniane programy bazują na konstrukcjach i wytycznych norm TEMA [23], uznanych i stosowanych przez inżynierów na całym świecie.

Celem niniejszej pracy jest dobór parownika o konstrukcji NKN wg. norm TEMA [23]. W celu wykonania niezbędnych obliczeń cieplnych, posłużono się specjalistycznym oprogramowaniem Aspen Exchanger Design&Rating.

13.3 Propozycja nowej konstrukcji

Na podstawie kompletu danych termodynamicznych obiegu, w tym parametrów parownika, zostaną przeprowadzone obliczenia mające

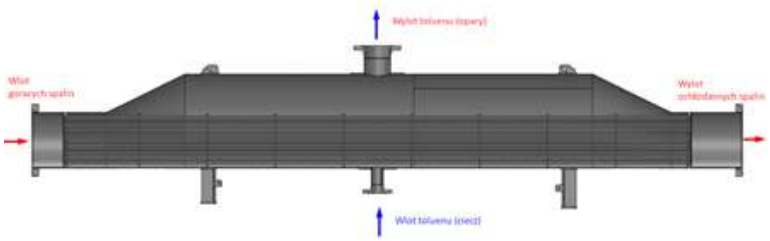
na celu wytypowanie najlepszego rozwiązania konstrukcyjnego według standardu TEMA [23]. Biorąc pod uwagę że spaliny są zapylone i zachodzi konieczność czyszczenia powierzchni wymiany ciepła, zdecydowano się na konstrukcję w której spaliny zostaną umieszczone w przestrzeni wewnątrz rurowej. Wymiennik pełni funkcję parownika, dlatego po analizie wielu różnych rozwiązań wybrano aparat wyparny typu „kettle reboiler" [24] w nieco zmodyfikowanej wersji. Typowy wymiennik ciepła typu „kettle reboiler" to urządzenie o konstrukcji BKU wg. norm TEMA [23], gdzie B oznacza zastosowanie komory przedniej demontowalnej, K to typ płaszcza kettle o średnicy większej aniżeli średnica komory przedniej, zaś U oznacza, iż zastosowano pęk rur w postaci u-rurek i wymiennik pozbawiony jest komory tylnej, a płaszcz zamknięty jest dennicą elipsoidalną. Typową konstrukcję wymiennika typu „kettle reboiler" przedstawia Rys. 13.4.



Rys. 13.4: Wymiennik ciepła typu BKU, źródło: https://rpadams.com/aku-series-heat-exchangers.

Z uwagi na wspomniany problem zanieczyszczenia spalin, niemożliwe jest zastosowanie typowej konstrukcji BKU. W urządzeniach tego typu, zgodnie z założeniami TEMA [23], czynnik odparowywany zawsze znajduje się w przestrzeni międzyrurowej czyli w płaszczu. Spaliny, z uwagi na zawarte w nich zanieczyszczenia powodują konieczność wprowadzenia zmian w typowej konstrukcji.

Wymóg okresowego wykonywania czynności serwisowych, polegających na mechanicznym czyszczeniu wewnętrznej powierzchni rur procesowych eliminuje z dalszych rozważań zastosowanie konstrukcji z u-rurami, gdzie takiej możliwości nie ma. Dodatkowo, spaliny jako nośnik ciepła odpadowego, charakteryzujący się ciśnieniem roboczym niewiele wyższym od ciśnienia atmosferycznego, wymusza poszukiwania rozwiązania charakteryzującego się niskim oporem przepływu, nie przekraczającym 0,4 kPa. Dlatego też, zaproponowano zastosowanie wymiennika jednoprzepływowego po stronie spalin z rurami prostymi o średnicy wewnętrznej umożliwiającej czyszczenie. Przekrój nowej konstrukcji parownika przedstawiono na Rys. 13.5.



Rys. 13.5: Poglądowy rysunek wymiennika typu NKN w przekroju wzdłużnym.

Przedstawiona na Rys. 13.5 konstrukcja to wymiennik typu NKN [23], który umożliwia realizację zadanego procesu wymiany ciepła zachodzącego pomiędzy gorącymi spalinami, a toluenem. Dla tej wybranej konstrukcji zostały przeprowadzone obliczenia cieplne.

13.4 Obliczenia projektowe

Opracowanie koncepcji wymiennika to pierwszy etap projektowania. Kolejnym jest wykonanie obliczeń cieplno-przepływowych, których celem jest wytypowanie optymalnej powierzchni wymiany ciepła, niezbędnej do osiągnięcia zadanej mocy cieplnej oraz parametrów towarzyszących, takich jak temperatury na wylocie oraz dopuszczalne maksymalne wartości spadków ciśnienia.

Wszystkie przedstawione wartości i wyniki w postaci tabel oraz rysunków, zostały obliczone przy pomocy programu Aspen Exchanger Design&Rating [21].

Wymiennik typu NKN (Rys. 13.5) jest urządzeniem jednoprzepływowym, dzięki czemu przy odpowiedniej ilości rurek zapewnia się niskie wartości oporów przepływu, a także gwarantuje stały i stabilny rozkład temperatury na całej długości rur procesowych. Z uwagi na fakt, iż projektowany jest nowy parownik w już istniejącej instalacji w zabudowie kontenerowej, należy uwzględnić ograniczenie co długości urządzenia, które nie może przekroczyć 4100 mm. Przy założeniu że komora wlotowa i wylotowa w sumie będą miały długość w zakresie 1000-1100 mm, zakłada się, iż maksymalna długość rur wymiennika powinna wynieść nie więcej aniżeli 3000 mm. Przy pomocy programu Aspen Exchanger Design&Rating wykonano obliczenia mające na celu wytypowanie najbardziej optymalnej średnicy rur procesowych oraz ich ilości, przy założeniu maksymalnej długości rur równej 3000 mm. W celu wykonania obliczeń zakłada się dla każdego rozważanego przypadku identyczną konstrukcję zgodnie z (Rys. 13.5) i formułuje bilans cieplny (Rys. 13.6) przy zachowaniu wymaganych parametrów roboczych (patrz Tab. 13.1).

		Hot Stream (1) Tube Side		Cold Stream (2) Shell Side	
Fluid name		Flue gas		Toluene	
		In	Out	In	Out
Mass flow rate	kg/h	2340		1422	
Mass flow rate multiplier		1		1	
Temperature	°C	510		87,44	211,8
Vapor mass fraction		1	1	0	1
Pressure (absolute)	bar	1,04	1,036	8,45	8,42
Pressure at liquid surface in column					
Heat exchanged	kW				
Heat exchanged multiplier			1		
Exchanger effectiveness					
Adjust if over-specified		Flowrate		Outlet temperature	
Estimated pressure drop	bar	0,004		0,03	
Allowable pressure drop	kPa	0,4		3	
Fouling resistance	m²-K/W	0,00176		9E-05	

Rys. 13.6: Bilans cieplny w programie Aspen Exchanger Design&Rating.

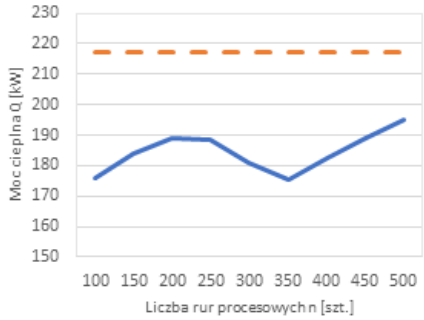
Spaliny o temperaturze 510°C, przy ciśnieniu na wlocie do wymiennika równym 1,04 kPa i przepływie 2340 kg/h zasilają wymiennik po stronie rur. Celem jest uzyskanie mocy cieplnej 217 kW. W tym celu formułuje się kompletny bilans cieplny po stronie czynnika roboczego, przy założeniu temperatury na wlocie 87,44°C, zaś na wylocie przy uwzględnieniu przegrzania 211,8°C. Toluen na wlocie do wymiennika ma ciśnienie 8,45bar(a), a wartość strumienia przepływu wynosi 1422 kg/h.

Tab. 13.2: Tabela porównawcza parametrów cieplno-przepływowych dla rury procesowej o wymiarze ϕ35x1 i liczbie rur n=100-500.

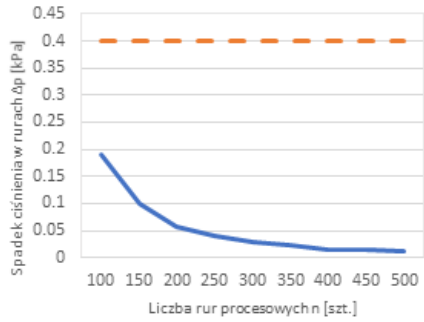
	n = 100	n = 150	n = 200	n = 250	n = 300	n = 350	n = 400	n = 450	n = 500
$Re_{śr(s)}$	8201,3	5512,1	4153,3	3321,1	2746,5	2341,9	2063,4	1845,9	1672,0
$Pr_{śr(s)}$	0,78	0,78	0,78	0,78	0,78	0,78	0,78	0,78	0,78
$v_{śr(s)}$ [m/s]	13,98	10,05	6,9	6,0	4,64	4,34	3,47	3,33	3,01
k [W/m^{2}K]	33,8	24,5	19,3	15,3	11,5	9,1	8,7	8,3	8,1
$\Delta p_{(s)}$ [kPa]	0,191	0,100	0,056	0,041	0,028	0,023	0,016	0,014	0,012
$T_{śr(s)}$ [°C]	214,57	212,11	210,74	209,84	209,1	208,37	207,98	207,72	207,5
Q [kW]	175,8	184,2	189	188,5	180,7	175,3	182,4	188,9	195,3
A [m^2]	31,9	47,9	63,9	79,9	95,8	111,8	127,8	143,7	159,7
$T_{wl(s)}$ [°C]	510	510	510	510	510	510	510	510	510
$T_{wyl(s)}$ [°C]	288,58	277,73	271,62	272,21	282,82	289,34	281,25	271,61	263,38
$T_{wl(t)}$ [°C]	87,44	87,44	87,44	87,44	87,44	87,44	87,44	87,44	87,44
$T_{wyl(t)}$ [°C]	205,34	205,34	205,34	205,34	205,34	205,34	205,34	205,34	205,34
ΔT_{log} [°C]	249,33	243	239,39	239,74	245,99	249,76	245,07	239,38	234,42

Poszukiwane parametry, które posłużyły wytypowaniu najlepszego rozwiązania obejmowały: uśrednione wartości liczby Re, liczby Pr, prędkości przepływu v, współczynnika przenikania ciepła k, opory przepływu po stronie rur Δp , temperaturę ścianki rurki $T_{śr}$, moc cieplną Q oraz powierzchnię wymiany ciepła A, wynikającą z liczby zastosowanych rur. W tabelach uwzględniono dodatkowo temperatury na wlocie i wylocie spalin ($T_{wl(s)}$, $T_{wyl(s)}$) oraz czynnika roboczego ($T_{wl(t)}$, $T_{wyl(t)}$), a także średnia logarytmiczna różnica temperatur ΔT_{log}, bazująca na punktach końcowych. Do rozważań przyjęto rury procesowe w następujących rozmiarach (średnica zewnętrzna [mm] x grubość ścianki [mm]):ϕ35x1, ϕ30x1, ϕ25x1, ϕ20x1. Jako materiał rur procesowych przyjęto dla wszystkich opcji stal austenityczną kwasoodporną w gatunku 1.4404 (316L). Do analizy przyjęto następujące wartości ilościowe rur procesowych n=100, 150, 200, 250, 300, 350, 400, 450, 500. Dla konstrukcji wymien-

nika z rurami procesowymi ϕ35x1, wyniki obliczeń przedstawiono w Tab. 13.2.



Rys. 13.7: Zmiana mocy cieplnej wymiennika w zależności od liczby rur (ϕ35x1).



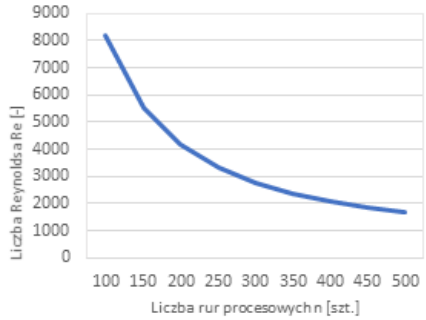
Rys. 13.8: Zmiana oporów po stronie spalin w zależności od liczby rur (ϕ35x1).

Właściwe rozwiązanie konstrukcyjne powinno spełniać dwa warunki. Pierwszy to osiągnięcie mocy cieplnej na poziomie 217 kW, wartość tę zaznaczono linią koloru pomarańczowego na wykresie Rys. 13.7, zaś drugi to maksymalny dopuszczalny spadek ciśnienia 0,4 kPa, co również zaznaczono linią koloru pomarańczowego na wykresie Rys. 13.8. Na wykresach dla kolejnych serii obliczeń, warunki te zostaną oznaczone w ten sam sposób. Opory przepływu po stronie

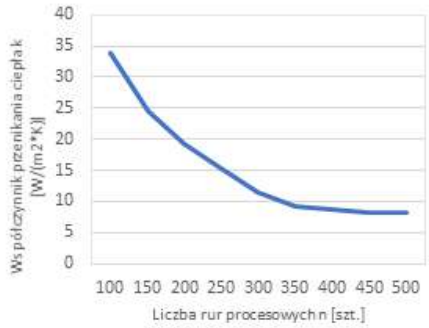
toluenu mają niskie wartości, dlatego też zostały pominięte.

Na wykresie Rys. 13.8, na którym przedstawiono wartości spadków ciśnienia w funkcji liczby rur, wyraźnie widać, iż warunek dotyczący oporów przepływu został spełniony dla wszystkich wariantów analizowanych ilości rur procesowych w zakresie od 100 do 500, co wynika z dość dużego całkowitego pola przekroju i niskich wartości prędkości spalin. Jak można zaobserwować na wykresie Rys. 13.7 warunek osiągnięcia mocy cieplnej na poziomie 217 kW nie został spełniony dla żadnego wariantu, a najwyższa wartość mocy cieplnej to 195,3 kW dla ilości rur n=500. Brak osiągnięcia wymaganej mocy cieplnej oznacza, że w żadnym z przypadków zadany strumień czynnika roboczego wynoszący 1422 kg/h nie został w całości odparowany. Dalsze zwiększanie powierzchni wymiany ciepła poprzez dodawanie rur mija się z celem, ponieważ już przy ilości rur n=500 powierzchnia wymiany ciepła wynosząca 159,7 m^2 nie jest dobrym rozwiązaniem pod względem konstrukcyjnym i ekonomicznym.

Na wykresie Rys. 13.10 przedstawiono współczynnik przenikania ciepła, będący jednym z najistotniejszych parametrów procesów wymiany ciepła. Służy on do określenia intensywności wymiany i uwzględnia transport ciepła, w tym przypadku ze spalin poprzez ściankę rurki do czynnika roboczego.



Rys. 13.9: Zmiana liczby Reynoldsa po stronie spalin w funkcji liczby rur (ϕ35x1).



Rys. 13.10: Wartości współczynnika przenikania ciepła w funkcji liczby rur (ϕ35x1).

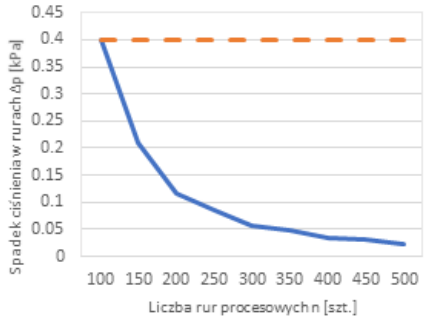
Wykres Rys. 13.10 pokazuje spadkowy trend współczynnika przenikania wraz ze wzrostem ilości rur. Ma to związek z coraz to niższymi wartościami prędkości spalin dla rosnącej ilości rur, co w rezultacie pogarsza warunki wymiany ciepła. Niższe prędkości spalin mają swoje odzwierciedlenie na wykresie Rys. 13.9 w postaci spadkowego trendu wartości liczb Re. Na potrzeby analiz, jako zakres przepływu przejściowego przyjęto $2300 \leq Re \leq 10000$. Dla liczby rur procesowych w zakresie od 100 do 350 odnotowuje się przepływ przejściowy, zaś już dla większej ilości rur jest to przepływ laminarny. W żadnym z analizowanych przypadków nie osiągnięto zadanej temperatury wylotowej czynnika roboczego na poziomie 211,8°C. Zgodnie z Tab. 13.2 dla wszystkich przypadków temperatura toluenu na wylocie wyniosła 205,34°C i jest to temperatura nasycenia dla ciśnienia 8,45bar(a). Zatem nie tylko nie udało się odparować wymaganego strumienia toluenu, ale także nie uzyskano nawet małego stopnia przegrzania par czynnika roboczego.

Kolejną serię obliczeń, tym razem dla rury o wymiarze ϕ30x1, przedstawiono w Tab. 13.3.

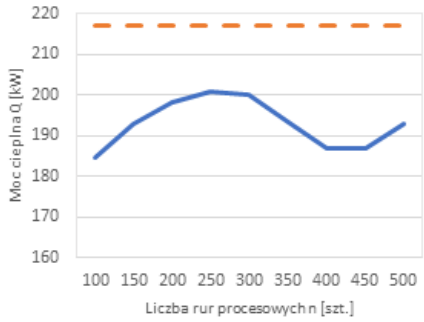
Tab. 13.3: Tabela porównawcza parametrów cieplno-przepływowych dla rury procesowej o wymiarze ϕ30x1 i ilości rur n=100-500.

	n = 100	n = 150	n = 200	n = 250	n = 300	n = 350	n = 400	n = 450	n = 500
$Re_{śr(s)}$	9747,8	6553,2	4941,5	3964,4	3299,9	2808,8	2442,1	2170,9	1965,9
$Pr_{śr(s)}$	0,78	0,78	0,78	0,78	0,78	0,78	0,78	0,78	0,78
$v_{śr(s)}$ [m/s]	19,27	13,84	9,5	8,23	6,32	5,93	4,8	4,64	4,15
k [W/m^{2}K]	44,5	32,3	25,6	21,1	17,3	13,5	10,8	9,5	9,2
$\Delta p_{(s)}$ [kPa]	0,401	0,209	0,115	0,085	0,057	0,048	0,034	0,030	0,024
$T_{śr(s)}$ [°C]	215,9	213,16	211,58	210,54	209,79	209,21	208,6	208,1	207,84
Q [kW]	184,6	193,0	198,3	201,0	199,9	193,1	186,8	186,8	193
A [m^2]	31,9	47,9	63,9	79,9	95,8	111,8	127,8	143,7	159,7
$T_{wl(s)}$ [°C]	510	510	510	510	510	510	510	510	510
$T_{wyl(s)}$ [°C]	277,25	266,35	259,5	255,94	257,49	266,22	275,19	274,43	266,38
$T_{wl(t)}$ [°C]	87,44	87,44	87,44	87,44	87,44	87,44	87,44	87,44	87,44
$T_{wyl(t)}$ [°C]	205,34	205,34	205,34	205,34	205,34	205,34	205,34	205,34	205,34
ΔT_{log} [°C]	242,72	236,23	232,08	229,9	230,85	236,15	241,51	241,05	236,25

Po zilustrowaniu wyników z Tab. 13.3 na wykresie Rys. 13.11 można zauważyć nieznaczne przekroczenie dopuszczalnego spadku ciśnienia dla liczby rur n=100. W pozostałych przypadkach opory są poniżej maksymalnej dopuszczalnej wartości oznaczonej linia koloru pomarańczowego i nie przekraczają 0,209 kPa. Warunek spadku ciśnienia dla ilości rur n≥150 został spełniony.

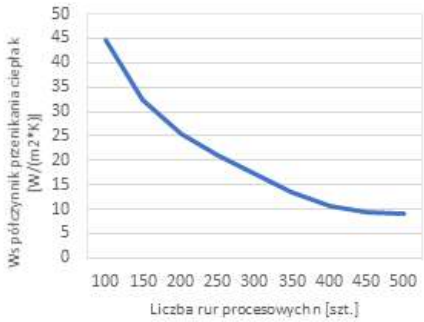


Rys. 13.11: Zmiana oporów przepływu spalin w funkcji liczby rur (ϕ30x1).



Rys. 13.12: Zmiana mocy cieplnej w funkcji liczby rur (ϕ30x1).

Drugi warunek, dotyczący osiągnięcia mocy cieplnej na poziomie 217 kW podobnie jak dla rury ϕ35x1 nie został spełniony dla całego badanego zakresu. Maksymalna wartość mocy cieplnej (Rys. 13.12) dla liczby rur n=250 wynosi 201 kW.



Rys. 13.13: Zmiana współczynnika przenikania ciepła w funkcji liczby rur (ϕ30x1).

Analizując Rys. 13.12 można zaobserwować pewną nieliniowość, która pojawiła się dość niewyraźnie również na Rys. 13.7. Nieliniowość ta polega na spadku wydajności wymiennika ciepła przy liczbie rur n=300 i ponownym wzroście przy n$\geq$450. W efekcie zmniejszenia średnicy rury procesowej z 35 mm do 30 mm uzyskano nieco wyższe wartości współczynników przenikania ciepła (Rys. 13.13),

Rys. 13.14: Zmiana liczby Reynoldsa po stronie spalin w funkcji liczby rur (ϕ30x1).

które kształtują się w zakresie od 44,5 [W/m^{2}K] do 9,2 [W/m^{2}K].

Wykres liczby Reynolds'a (Rys. 13.14) podobnie jak dla większej średnicy rur ϕ35x1, w większości analizowanych przypadków zawiera się w zakresie przepływu przejściowego, zaś dla ilości rur n$\geq$450 jest to już przepływ laminarny.

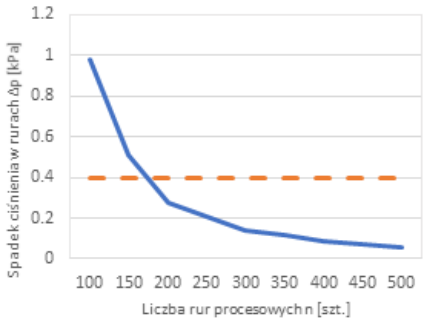
Wynik dla wymiennika z rurą ϕ25x1 przedstawiono w Tab. 13.4.

Tab. 13.4: Tabela porównawcza parametrów cieplno-przepływowych dla rury procesowej o wymiarze ϕ25x1 i ilości rur n=100-500.

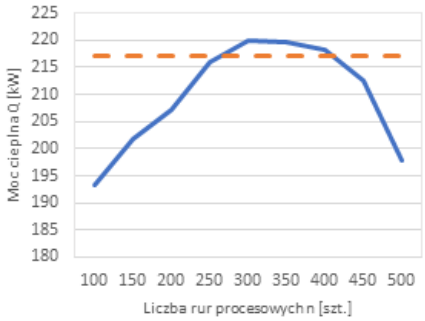
	n = 100	n = 150	n = 200	n = 250	n = 300	n = 350	n = 400	n = 450	n = 500
$Re_{śr(s)}$	11972,1	8051,0	6071,8	4904,9	4104,4	3517,9	3073,3	2745,0	2405,3
$Pr_{śr(s)}$	0,78	0,78	0,78	0,78	0,78	0,78	0,78	0,78	0,78
$v_{śr(s)}$ [m/s]	28,38	20,36	13,96	12,03	9,19	8,56	6,9	6,7	6,12
k [W/m^{2}K]	61,6	44,6	35,5	24,4	21,4	18,4	15,7	13,0	12,0
$\Delta p_{(s)}$ [kPa]	0,982	0,51	0,279	0,208	0,139	0,115	0,084	0,074	0,060
$T_{śr(s)}$ [°C]	218,02	214,68	212,77	223,3	219,35	217,54	211,03	205,95	208,54
Q [kW]	193,4	201,9	207,1	216,1	219,8	219,7	218,3	212,7	197,9
A [m^2]	22,9	34,3	45,8	57,2	68,7	80,1	91,6	103,0	114,5
$T_{wl(s)}$ [°C]	510	510	510	510	510	510	510	510	510
$T_{wyl(s)}$ [°C]	265,76	254,81	239,13	236,27	232,16	231,45	234,56	240,64	260,02
$T_{wl(t)}$ [°C]	87,44	87,44	87,44	87,44	87,44	87,44	87,44	87,44	87,44
$T_{wyl(t)}$ [°C]	205,34	205,34	206,93	209,86	214,06	214,79	211,61	205,38	205,34
ΔT_{log} [°C]	235,88	229,2	218,72	215,71	211,39	210,65	213,91	220,30	232,4

Jak można zauważyć na Rys. 13.15, zmniejszenie średnicy rury procesowej o kolejne 5 mm znacząco wpływa na wzrost oporów

przepływu, wartość dopuszczalna 0,4 kPa została przekroczona dla dwóch badanych przypadków, n=100 i n=150.



Rys. 13.15: Zmiana oporów przepływu po stronie spalin w funkcji liczby rur (ϕ25x1).



Rys. 13.16: Zmiana mocy cieplnej w funkcji liczby rur (ϕ25x1).

Wzrost oporów przepływu jest skutkiem zwiększenia prędkości spalin o około 50% dla każdego z analizowanych przypadków, w porównaniu do obliczeń dla rury ϕ30x1. Warunek zachowania maksymalnych oporów przepływu został spełniony dla ilości rur procesowych n$\geq$200. Z wykresu na Rys. 13.16 wynika, że zmniejszenie średnicy rur pozwoliło także na uzyskanie w większości analizowanych przypadków wymaganej wydajności na poziomie minimum 217 kW, z podobną nieliniowością spadku mocy przy n$\geq$400, jednak

tym razem bez wyraźnego odbicia. Dla badanej liczby rur w zakresie 200-450 uzyskano przegrzanie czynnika roboczego nawet do 214,97°C dla ilości rur n=350. Warto dodać, iż dla poprzednich analizowanych średnic rur nie odnotowano przegrzania par czynnika roboczego w żadnym przypadku.

Rys. 13.17: Zmiana liczby Reynoldsa po stronie spalin w funkcji liczby rur (ϕ25x1).

Rys. 13.18: Zmiana współczynnika przenikania ciepła w funkcji liczby rur (ϕ25x1).

Na wykresie Rys. 13.18 można zauważyć wzrost współczynnika przenikania ciepła dla rury ϕ25x1 w porównaniu z rurą ϕ30x1 (Rys. 13.13), najbardziej widoczne jest to dla najmniejszych analizowanych wartości n w zakresie od 100 do 200. Dla pozostałych ilości rur wzrost

ten jest niewielki. Zarówno współczynnik k jak i liczba Re zachowuje stały spadkowy trend, co miało miejsce także w poprzednich analizowanych przypadkach. Analizując wykres Rys. 13.17 można zauważyć iż, po raz pierwszy przekroczono granicę przepływu przejściowego dla najmniejszej analizowanej liczby rur n=100.

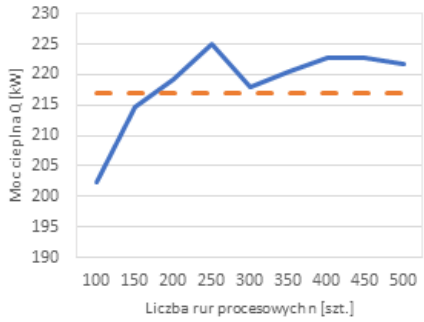
Wyniki obliczeń dla rury ϕ20x1 przedstawiono w Tab. 13.5.

Tab. 13.5: Tabela porównawcza parametrów cieplno-przepływowych dla rury procesowej o wymiarze ϕ20x1 i ilości rur n=100-500.

	n = 100	n = 150	n = 200	n = 250	n = 300	n = 350	n = 400	n = 450	n = 500
$Re_{śr(s)}$	15438,2	10430,3	7862,2	6331,2	5233,1	4498,5	3946,5	3508,0	3153,33
$Pr_{śr(s)}$	0,78	0,78	0,78	0,78	0,78	0,78	0,78	0,78	0,78
$v_{śr(s)}$ [m/s]	46,28	32,95	22,57	19,49	15,04	8,97	6,22	10,84	9,76
k [W/m^{2}K]	91,7	54,6	40,3	34,8	30,4	27,5	24,6	22,1	19,4
$\Delta p_{(s)}$ [kPa]	3,041	1,595	0,884	0,641	0,426	0,352	0,255	0,226	0,188
$T_{śr(s)}$ [°C]	221,24	225,77	234,33	231,14	229,12	227,84	224,55	222,54	219,33
Q [kW]	202,3	214,7	219,3	225,1	217,9	220,4	222,8	222,7	221,7
A [m^2]	18,3	27,5	36,6	45,8	54,9	64,1	73,3	82,4	91,6
$T_{wl(s)}$ [°C]	510	510	510	510	510	510	510	510	510
$T_{wyl(s)}$ [°C]	254,20	238,02	233,60	224,48	234,34	230,56	227,61	227,52	228,92
$T_{wl(t)}$ [°C]	87,44	87,44	87,44	87,44	87,44	87,44	87,44	87,44	87,44
$T_{wyl(t)}$ [°C]	205,34	208,08	212,6	221,9	211,84	215,7	218,71	218,8	217,37
ΔT_{log} [°C]	228,83	217,55	212,90	203,3	213,68	209,71	206,6	206,51	207,98

Wyniki obliczeń zilustrowano na wykresach.

Rys. 13.19: Zmiana oporów przepływu po stronie spalin w funkcji liczby rur (ϕ20x1).



Rys. 13.20: Zmiana wartości mocy cieplnej w funkcji liczby rur (ϕ20x1).

Po ostatniej analizie aparatu wyparnego z rurą procesową o wymiarze ϕ20x1, na podstawie wykresów Rys. 13.19 i Rys. 13.20 można stwierdzić, iż ten rozmiar rur przy stosunkowo niedużej powierzchni wymiany ciepła spełnia warunek dopuszczalnego spadku ciśnienia oraz osiągnięcia zadanej mocy cieplnej. Optymalne rozwiązanie znajduje się w przedziale liczby rur pomiędzy $300<n<350$, co przekłada się na przedział powierzchni wymiany ciepła pomiędzy 54,9 m^2 a 64,1 m^2. Oba warunki są także spełnione dla ilości rur $n\geq400$, jednak z uwagi na większą powierzchnię wymiany ciepła i związane z tym koszty produkcji, rozwiązania te nie są brane pod uwagę.

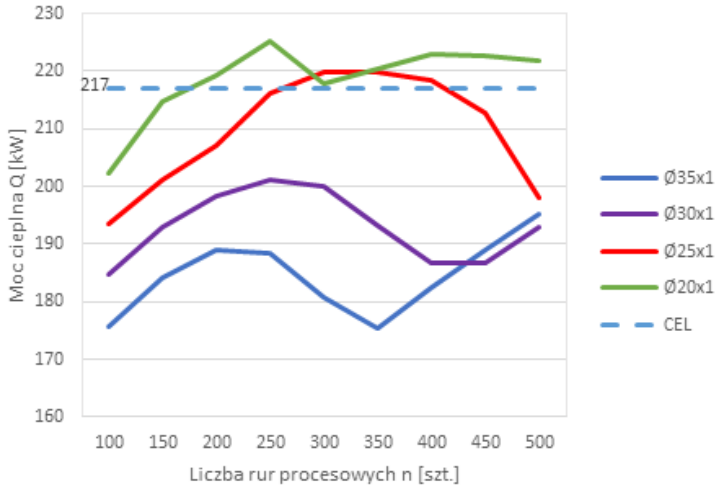
Współczynnik przenikania ciepła (Rys. 13.21) dla optymalnego rozwiązania $300<n<350$ zawiera się w zakresie pomiędzy 30,4 a 27,5 [W/m^2K]. Wyższe wartości prędkości spalin i w efekcie wyższe wartości liczby Reynoldsa, (Rys. 13.22) wynikają ze zmniejszenia przekroju przepływu po stronie spalin, w stosunku do wcześniejszych serii obliczeń z rurami o większych średnicach. Na wykresie Rys. 13.22 można zauważyć, iż optymalne rozwiązanie znajduje się w zakresie przepływu przejściowego, niemniej jednak jest to najlepsze i najbardziej optymalne rozwiązanie do konstrukcji analizowanego wymiennika ciepła typu NKN. Na Rys. 13.23 zestawiono wartości mocy cieplnych dla różnych wariantów wymiennika z rurą ϕ35x1, ϕ30x1, ϕ25x1 oraz ϕ20x1.

Rys. 13.21: Zmiana współczynnika przenikania ciepła w funkcji liczby rur (ϕ20x1).

Rys. 13.22: Zmiana liczby Reynoldsa po stronie spalin w funkcji liczby rur (ϕ20x1).

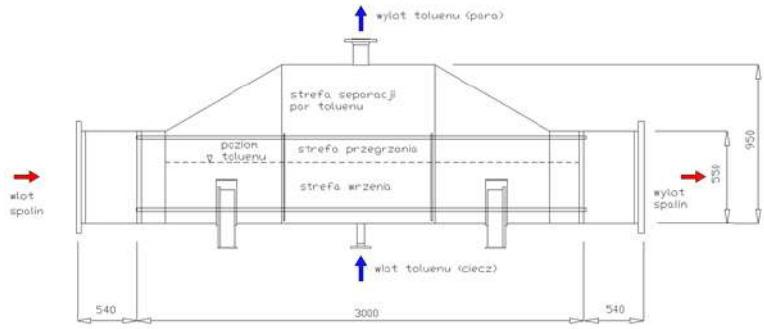
Analizując Rys. 13.23 zaobserwowano, iż warunek spełnienia mocy cieplnej został osiągnięty jedynie dla kilku badanych rozwiązań konstrukcyjnych z rurą procesową ϕ25x1 oraz ϕ20x1. Rury o większych rozmiarach są dla rozważanego przypadku nieefektywne, o czym świadczą niskie wartości liczb Reynoldsa i w rezultacie współczynnika przenikania ciepła. Interesującym zjawiskiem jest efekt nieliniowości zmiany mocy cieplnej obserwowany wraz ze wzrostem liczby rur procesowych. Dla wszystkich badanych rozmiarów rur (Rys. 13.9, Rys. 13.14, Rys. 13.17 i Rys. 13.22) obserwuje się spadek mocy przy wartościach Re około 6000 i ponowny wzrost przy

wartościach od 3000 do 4000. Nieliniowość ta ma związek z charakterem przejściowym przepływu i być może wrzeniem toluenu w przestrzeni międzyrurowej.



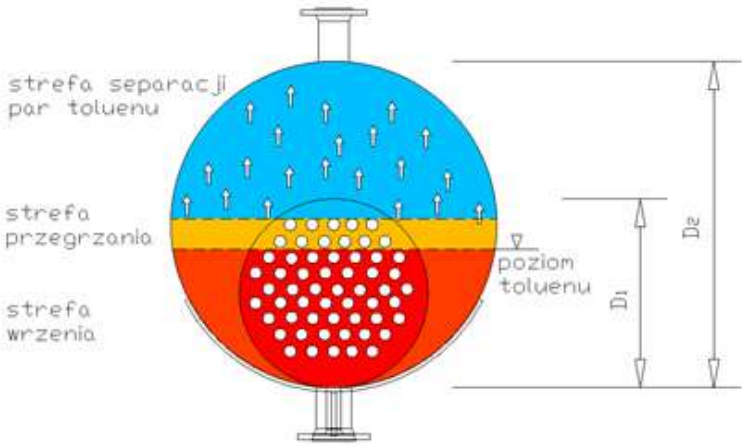
Rys. 13.23: Zestawienie charakterystyk mocy cieplnej dla różnych rozmiarów rur wymiennika NKN.

W programie Aspen Exchanger Design&Rating wykonano kilka dodatkowych serii obliczeń w celu określenia dokładnej ilości rur ϕ20x1 z przedziału 300<n<350. Obliczenia wykonano dla ilości rur procesowych od 305 do 345 z częstotliwością co 5, a następnie od 341 do 344 z częstotliwością co 1. Jako kryterium przyjęto minimalną ilość rur procesowych, niezbędną do uzyskania właściwego stopnia przegrzania par toluenu na wylocie z wymiennika 211,8°C. Wyniki obliczeń nie zostały zamieszczone, a jako rezultat wskazano jedynie dokładną ilość rur procesowych w ilości 344. Każda z rur procesowych ma długość 3000 mm, zaś całkowita powierzchnia wymiany ciepła dla optymalnej konstrukcji wynosi 63m^2. Szkic konstrukcji parownika przedstawiono na Rys. 13.24.



Rys. 13.24: Szkic parownika w postaci aparatu wyparnego typu NKN z prostymi rurkami.

Pęk rurowy zalany jest częściowo czynnikiem roboczym gdzie dochodzi do wrzenia, pozostała jego cześć przeznaczona jest na strefę przegrzania par toluenu (Rys. 13.24, Rys. 13.25).

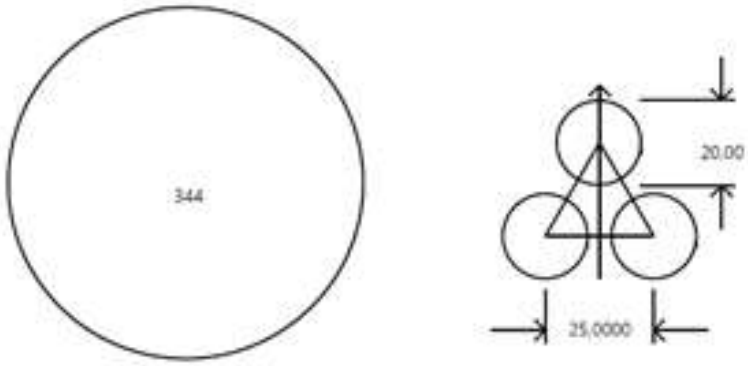


Rys. 13.25: Przekrój przez płaszcz aparatu wyparnego NKN.

Zarówno wrzenie jak i odparowanie ma miejsce jedynie w obrębie obrysu pęku rurowego. Objętość płaszcza powyżej poziomu skrajnych rur procesowych pełni rolę strefy odseparowania par toluenu (Rys. 13.24, Rys. 13.25), która wykorzystywana jest do swobodnego rozdziału faz. Bardzo istotnym elementem jest zachowanie stosunku

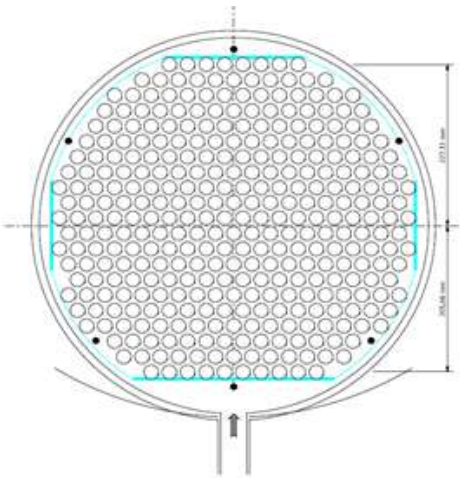
średnicy płaszcza $D_2/D_1 \geq 1{,}4$ (Rys. 13.25). Konieczność zapewnienia odpowiednio dużej objętości nie zalanej czynnikiem roboczym zapewnia właściwe oddzielenie się pęcherzy parowych od kropli cieczy, co wyklucza potencjalne uszkodzenie łopatek turbiny.

W tym przypadku zastosowano trójkątny układ rurek z podziałką 25 mm (Rys. 13.26). Dla nowej konstrukcji zastosowano nieco większą różnicę stosunku średnic D_2/D_1 (Rys. 13.25) aniżeli mówi o tym literatura. W rozważanym przypadku D1=950 mm, zaś D2=550 mm co w rezultacie daje D_2/D_1 =(950 mm)/(550 mm)=1,73. Rezultaty obliczeń cieplnych przedstawiono w karcie doboru urządzenia (Rys. 13.28). Spadek ciśnienia dla spalin wyniósł nieco ponad 0,3 kPa, ale nie przekroczył zadanej wartości 0,4 kPa przy prędkości przepływu przez rurki w zakresie 12,53 – 15,9 m/s. Po stronie toluenu projektowa maksymalna wartość spadku ciśnienia została przekroczona i wyniosła 3,41 kPa, przy ograniczeniu 3 kPa. Wynika to z konstrukcji urządzenia, zwartej zabudowy pęku rurowego i jego długości. Jest to jednak najlepszy możliwy kompromis pomiędzy wydajnością, wielkością wymiennika oraz ograniczeniem wynikającym ze spadku ciśnienia głównie po stronie spalin.



Rys. 13.26: Układ rur procesowych.

Przekrój przez pęk rurowy parownika pokazuje Rys. 13.27.



Rys. 13.27: Przekrój przez płaszcz.

Heat Exchanger Specification Sheet

1	Company: SECESPOL Sp. z o.o.					
2	Location: Nowy Dwór Gdański, Poland					
3	Service of Unit:	Our Reference:				
4	Item No.: Parownik	Your Reference:				
5	Date: 01.11.2021	Rev No.: 0	Job No.:			
6	Size: 550 /950 -3000 mm	Type: NKN Horizontal		Connected in: 1 parallel	1 series	
7	Surf/unit(eff.) 63 m²	Shells/unit 1		Surf/shell(eff.)	63 m²	
8	**PERFORMANCE OF ONE UNIT**					
9	Fluid allocation	Shell Side		Tube Side		
10	Fluid name	Toluene		Flue gas		
11	Fluid quantity, Total kg/h	1422		2340		
12	Vapor (In/Out) kg/h	0	1422	2340	2340	
13	Liquid kg/h	1422	0	0	0	
14	Noncondensable kg/s	0	0	0	0	
15						
16	Temperature (In/Out) °C	87,44	211,8	510	234,4	
17	Bubble / Dew point °C	205,34 /	/	/	/	
18	Density Vapor/Liquid kg/m³	/ 805,05	22,7 / 674,61	0,47 /	0,72 /	
19	Viscosity mPa-s	/ 0,2983	0,0116 / 0,1364	0,0348 /	0,0253 /	
20	Molecular wt, Vap		92,13	26,91	26,91	
21	Molecular wt, NC					
22	Specific heat kJ/(kg-K)	/ 1,895	1,879 / 2,258	1,256 /	1,176 /	
23	Thermal conductivity W/(m-K)	/ 0,1187	0,0267 / 0,0842	0,0577 /	0,0376 /	
24	Latent heat kJ/kg		293,2			
25	Pressure (abs) bar	8,45	8,41589	1,13	1,12675	
26	Velocity (Mean/Max) m/s	0,13 / 0,16		12,53 / 15,9		
27	Pressure drop, allow./calc. kPa	3	3,411	0,4	0,325	
28	Fouling resistance (min) m²-K/W	9E-05		0,00176	0,00196 Ao based	
29	Heat exchanged 217,5 kW			MTD (corrected) 115,93 °C		
30	Transfer rate, Service 30,1	Dirty 32		Clean 34,2	W/(m²-K)	

Rys. 13.28: Karta obliczeń cieplnych z programu Aspen Exchanger Design&Rating.

W wierszu numer 30 karty obliczeń (Rys. 13.28) znajdują się wartości współczynników przenikania ciepła. Wartość oznaczona jako „Transfer rate" jest wartością obliczeniową bazową, czyli wartością w sytuacji gdy wymiennik jest dobrany „w punkt", bez jakiegokolwiek naddatku powierzchni wymiany ciepła. Wartość „ Service dirty" charakteryzuje współczynnik przenikania ciepła dla wymiennika w sytuacji, w której zakłada się naddatek w postaci zwiększenia powierzchni wymiany ciepła poprzez wydłużenie rur, zwiększenie ich ilości itp. Dzieląc wartość „Service dirty" przez wartość obliczeniową bazową otrzymuje się procentową wartość przewymiarowania powierzchni wymiany ciepła, czyli $\frac{32\ [W/m^2K]}{30{,}1\ [W/m^2K]}$=6,3%. Wartość „Service clean" z kolei charakteryzuje współczynnik, który uwzględnia nie tylko naddatek w postaci bezpośrednio powierzchni wymiany ciepła, ale także współczynnik zanieczyszczenia płynów po obu stronach. Czyli dzieląc wartość „Service clean" przez wartość obliczeniową bazową uzyskamy całkowite przewymiarowanie powierzchni wymiany ciepła, które wyniosło $\frac{34{,}2\ [W/m^2K]}{30{,}1\ [W/m^2K]}$=13,6%. Program doboru uwzględnił degradację powierzchni wymiany ciepła na skutek zastosowania współczynników zanieczyszczenia o wartości równej 13,6%-6,3%=7,6% całkowitej powierzchni wymiany ciepła.

Rating / Checking		Shell Side		Tube Side	
Total mass flow rate	kg/h	1422		2340	
Vapor mass flow rate (In/Out)	kg/s	0	0.395	0.65	0.65
Liquid mass flow rate	kg/h	1422	0	0	0
Vapor mass fraction		0	1	1	1
Temperatures	°C	87.44	211.8	510	234.4
Bubble / Dew point	°C	205.34 /	/	/	/
Operating Pressures	bar	8.45	8.41589	1.13	1.12675
Film coefficient	W/(m²-K)	832.1		35.8	
Fouling resistance	m²-K/W	9E-05		0.00176	
Velocity (highest)	m/s	0.18		15.9	
Pressure drop (allow./calc.)	kPa	3 /	3.411	0.4 /	0.325
Total heat exchanged	kW	217.5			
Overall clean coeff. (plain/finned)	W/(m²-K)	34.2 /			
Overall dirty coeff. (plain/finned)	W/(m²-K)	32 /			
Effective area (plain/finned)	m²	63 /			
Effective MTD	°C	115.93			
Actual/Required area ratio (dirty/clean)		1.06 / 1.14			
Vibration problem (HTFS)		No			
RhoV2 problem		No			

Unit	NKN	1 pass	1 ser	1 par		
Shell size	530	- 3000	mm	Hor		
Tubes	Plain					
Insert	None					
No.	344	OD	20	Tks	1	mm
Pattern	30	Pitch	25	mm		
Baffles	Unbaffled	Cut(%d)				
Total cost	50205	Dollar(US)				

Rys. 13.29: Zestawienie najważniejszych parametrów parownika.

Rys. 13.29 przedstawia zestawienie głównych parametrów roboczych wymiennika, zarówno po stronie płaszcza, jak i po stronie rur. Pierwszy wiersz patrząc od góry zawiera parametry robocze, ta-

kie jak przepływy masowe, rodzaj frakcji, temperatury oraz ciśnienia wejścia i wyjścia. Kolejny wiersz zawiera dane dotyczące współczynników przenikania ciepła po stronie płaszcza i rur (z ang. *film coefficient*), przyjęte wartości zanieczyszczenia płynów, średnie prędkości płynów oraz wartości dopuszczalnego oraz obliczeniowego spadku ciśnienia. Współczynnik przejmowania po stronie rur wynosi α_1= 35,8 W/m^{2}K, zaś po stronie płaszcza α_2= 832,1 W/m^{2}K, gdzie wartość ta jest wynikiem przemiany fazowej na skutek odparowania czynnika roboczego. Oba współczynniki są niezbędne do wyliczenia całkowitego współczynnika przenikania ciepła dla rozpatrywanego wymiennika płaszczowo-rurowego, w myśl wzoru:

$$k = \frac{1}{\frac{1}{\alpha_1} + \frac{\delta}{\lambda} + \frac{1}{\alpha_2}} \tag{13.1}$$

Na Rys. 13.30 pokazano wartości liczb kryterialnych na wlocie i wylocie wymiennika, z lewej strony dla toluenu, z prawej dla spalin.

Heat Transfer Parameters		In	Out	In	Out
Prandtl numbers	Vapor		0.82	0.75	0.79
	Liquid	4.76			
Reynolds numbers	Vapor Nominal		81156.78	3844.41	5279.45
	Liquid Nominal	3160.54			

Rys. 13.30: Liczby kryterialne dla przepływu toluenu i spalin.

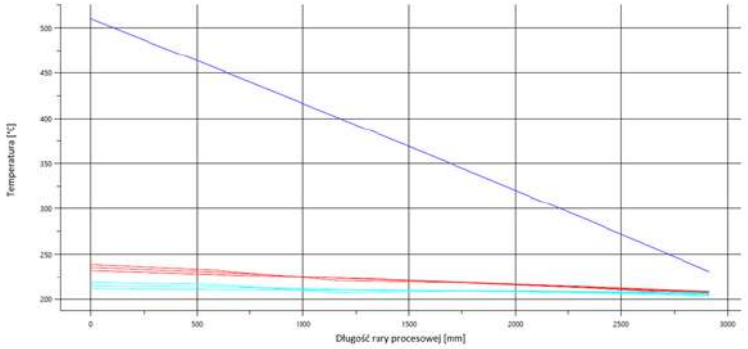
Liczba Reynoldsa po stronie toluenu na wlocie do płaszcza wynosi 3160,54 i plasuje się w zakresie przepływu przejściowego, zaś po zmianie fazy na gazową na wylocie z płaszcza wynosi już 81156,78 i jest to już zdecydowanie przepływ burzliwy. Po stronie spalin wartość liczby Re na wlocie wynosi 3844,41, zaś na wylocie 5279,45, obie wartości znajdują się w zakresie przepływu przejściowego, a różnica pomiędzy nimi wynika z różnic temperatury pomiędzy wlotem a wylotem, co w rezultacie wpływa na gęstość i prędkość przepływu spalin.

Dane dotyczące średnich temperatur ścianki płaszcza i ścianek rur (Rys. 13.31) w wymiennikach płaszczowo-rurowych wykorzystuje się do obliczeń wydłużeń cieplnych i ich ewentualnej kompensacji. Kompensacja wydłużeń odbywa się poprzez zastosowanie na płaszczu wymiennika odpowiedniego kompensatora typu soczewkowego

Temperature Difference	°C	Heat Flux (based on tube O.D)	kW/m²
Overall effective MTD	115,93	Overall flux	3,71
One pass counterflow MTD	118,74	Critical heat flux (at highest ratio)	418,43
LMTD based on end points	213,74	Highest local flux	10,75
Effective MTD correction factor	0,54	Highest local/critical flux	0,03
Wall Temperatures	°C		
Shell mean metal temperature		204,42	
Tube mean metal temperature		211,88	
Tube wall temperatures (highest/lowest)		216,82 / 203,14	

Rys. 13.31: Średnie temperatury ścianek oraz logarytmiczna różnica temperatur.

lub mieszkowego. Na Rys. 13.32 pokazano kolorem błękitnym rozkład temperatury ścianki rur na całej długości rur procesowych czyli 3 000 mm, gdzie 0 na osi y oznacza początek wymiennika, gdzie wpływają gorące spaliny.



Rys. 13.32: Rozkład temperatur w rurkach wymiennika po stronie spalin (kolor niebieski – temperatura splin, kolor czerwony – średnia temperatura ścianki rury z uwzględnieniem współczynników zanieczyszczenia, kolor błękitny – średnia temperatura ścianki rury bez uwzględnienia współczynników zanieczyszczenia).

Temperatura na całej długości rur kształtuje się w zakresie 225-200°C i maleje w kierunku króćca wylotowego spalin. Kolorem czerwonym zaznaczono temperaturę ścianki uwzględniającą dodatkowy opór cieplny wynikający z zanieczyszczenia powierzchni wymiany ciepła na skutek działania uwzględnionych współczynników zanieczyszczenia. Dla tego przypadku intensywność wymiany ciepła jest pogorszona, a średnie temperatury ścianek są nieco wyższe

i kształtują się w zakresie 240-200°C. Kolorem niebieskim oznaczono krzywą wartości średniej temperatury spalin.

13.5 Podsumowanie i wnioski

W wyniku przeprowadzonej analizy i obliczeń cieplnych za pomocą programu Aspen Exchanger Design&Rating wskazano optymalne rozwiązanie dla przyjętej konstrukcji wymiennika typu NKN. Rozwiązaniem spełniającym zadane parametry pracy jest urządzenie zawierające 344 rury procesowe w rozmiarze ϕ20x1 i długości jednostkowej odcinka l=3000 mm. Objętość przestrzeni płaszcza to 0,99 m^3, objętość przestrzeni rur 0,41 m^3. Całkowita powierzchnia wymiany ciepła parownika to 63 m^2, a waga pustego urządzenia pozbawionego płynów roboczych to ∼1650 kg. Najlepsze wyniki zostały osiągnięte dla najmniejszej z analizowanych średnic rur procesowych ϕ20x1, co jest bezpośrednio związane z wyższymi prędkościami płynów roboczych, liczbami Reynoldsa i w efekcie wyższym współczynnikiem przenikania ciepła. Im mniejsza średnica rury procesowej, tym lepsze współczynniki wymiany ciepła i w rezultacie mniejsza powierzchnia wymiany ciepła. Wyniki obliczeń cieplnych ukazujące nieliniowe charakterystyki mocy wskazują, iż procesy wymiany ciepła polegające na przemianach fazowych należy traktować ze szczególną uwagą i poddawać wielokrotnej analizie.

W obecnie eksploatowanej mikrosiłowni o mocy 30 kWe bazuje się w dalszym ciągu na odparowaniu pośrednim. Zamiast jednego parownika w układzie zainstalowane są dwa wymienniki. Pierwszym z nich jest ekonomizer, którego zadaniem jest odbiór ciepła ze spalin i przekazanie go do dodatkowego obiegu olejowego. Podgrzany przez spaliny olej termalny zasila kolejny wymiennik, w którym zachodzi odparowanie czynnika roboczego. Rozwiązanie to jest znacznie droższe, ponieważ wymaga rozbudowy całego układu o obieg olejowy, na który składa się ekonomizer, parownik, pompa obiegowa, rurociągi, olej termalny, armatura zabezpieczająca i kontrolno-pomiarowa.

W niniejszym opracowaniu wytypowano optymalne rozwiązanie konstrukcyjne płaszczowo-rurowego wymiennika ciepła, służącego do

bezpośredniego odparowania czynnika roboczego. Rozwiązanie to umożliwia usunięcie z układu całego obiegu olejowego, co stanowi realne oszczędności nie tylko w budowie tego typu instalacji, ale także w trakcie długoletniej eksploatacji.

Podziękowanie

Praca została zrealizowana w ramach II edycji programu Ministerstwa Edukacji i Nauki „Doktorat wdrożeniowy”.

Bibliografia

[1] Guoquan Qiu, Yingjuan Shao, Jinxing Li, Hao Liu, and Saffa B Riffat. Experimental investigation of a biomass-fired orc-based micro-chp for domestic applications. *Fuel*, 96:374–382, 2012.

[2] Hao Liu, Yingjuan Shao, and Jinxing Li. A biomass-fired micro-scale chp system with organic rankine cycle (orc)–thermodynamic modelling studies. *Biomass and bioenergy*, 35(9):3985–3994, 2011.

[3] Serafim Georgousopoulos, Konstantinos Braimakis, Dimitrios Grimekis, and Sotirios Karellas. Thermodynamic and techno-economic assessment of pure and zeotropic fluid orcs for waste heat recovery in a biomass igcc plant. *Applied Thermal Engineering*, 183:116202, 2021.

[4] Zhang Shengjun, Wang Huaixin, and Guo Tao. Performance comparison and parametric optimization of subcritical organic rankine cycle (orc) and transcritical power cycle system for low-temperature geothermal power generation. *Applied energy*, 88(8):2740–2754, 2011.

[5] Zvonimir Guzović, Boris Majcen, and Svetislav Cvetković. Possibilities of electricity generation in the republic of croatia from medium-temperature geothermal sources. *Applied energy*, 98:404–414, 2012.

[6] Tao Guo, HX Wang, and SJ Zhang. Fluids and parameters optimization for a novel cogeneration system driven by

low-temperature geothermal sources. *Energy*, 36(5):2639–2649, 2011.

[7] Paola Bombarda, Costante M Invernizzi, and Claudio Pietra. Heat recovery from diesel engines: A thermodynamic comparison between kalina and orc cycles. *Applied thermal engineering*, 30(2-3):212–219, 2010.

[8] EH Wang, HG Zhang, BY Fan, MG Ouyang, Y Zhao, and QH Mu. Study of working fluid selection of organic rankine cycle (orc) for engine waste heat recovery. *Energy*, 36(5):3406–3418, 2011.

[9] Manuel Jiménez-Arreola, Roberto Pili, Christoph Wieland, and Alessandro Romagnoli. Analysis and comparison of dynamic behavior of heat exchangers for direct evaporation in orc waste heat recovery applications from fluctuating sources. *Applied energy*, 216:724–740, 2018.

[10] Manuel Jiménez-Arreola, Christoph Wieland, and Alessandro Romagnoli. Direct vs indirect evaporation in organic rankine cycle (orc) systems: A comparison of the dynamic behavior for waste heat recovery of engine exhaust. *Applied Energy*, 242:439–452, 2019.

[11] R Chacartegui, D Sánchez, JM Muñoz, and T Sánchez. Alternative orc bottoming cycles for combined cycle power plants. *Applied Energy*, 86(10):2162–2170, 2009.

[12] Fahad A Al-Sulaiman, Ibrahim Dincer, and Feridun Hamdullahpur. Energy and exergy analyses of a biomass trigeneration system using an organic rankine cycle. *Energy*, 45(1):975–985, 2012.

[13] T Tartière and M Astolfi. Orc world map. *https://orc-world-map.org/*, 2021.

[14] Thomas Tartière and Marco Astolfi. A world overview of the organic rankine cycle market. *Energy procedia*, 129:2–9, 2017.

[15] Ivanka Milcheva, Florian Heberle, and Dieter Brüggemann. Modeling and simulation of a shell-and-tube heat exchanger for organic rankine cycle systems with double-segmental baffles by adapting the bell-delaware method. *Applied Thermal Engineering*, 126:507–517, 2017.

[16] Rita Mastrullo, Alfonso William Mauro, Rémi Revellin, and Luca Viscito. Modeling and optimization of a shell and louvered fin mini-tubes heat exchanger in an orc powered by an internal combustion engine. *Energy conversion and management*, 101:697–712, 2015.

[17] Fan Wei, Guo Senchuang, and Han Zhonghe. Economic analysis of organic rankine cycle (orc) and organic rankine cycle with internal heat exchanger (iorc) based on industrial waste heat source constraint. *Energy Procedia*, 158:2403–2408, 2019.

[18] Maria Anna Chatzopoulou, Steven Lecompte, Michel De Paepe, and Christos N Markides. Off-design optimisation of organic rankine cycle (orc) engines with different heat exchangers and volumetric expanders in waste heat recovery applications. *Applied energy*, 253:113442, 2019.

[19] Cheng Zhang, Chao Liu, Shukun Wang, Xiaoxiao Xu, and Qibin Li. Thermo-economic comparison of subcritical organic rankine cycle based on different heat exchanger configurations. *Energy*, 123:728–741, 2017.

[20] Tallapureddy Subba Reddy, Thimmasandra Venkataswamy Sreerama Reddy, Erappa Rajj Babu, and Hosur Nanji Reddy Reddappa. Modified kinetic gas molecule optimization for four fluids of organic rankine cycle in shell and tube heat exchanger. *Materials Today: Proceedings*, 47:2258–2262, 2021.

[21] Aspen technology inc., aspentech exchanger design&rating (edr). *https://www.aspentech.com/en/products/engineering/aspen-exchanger-design-and-rating*, 2021.

[22] Heat transfer research, inc. *https://www.htri.net/software*, 2021.

[23] TEMA. 10th edition tema standards. 2020.

[24] Ennio Macchi and Marco Astolfi. *Organic rankine cycle (ORC) power systems: technologies and applications*. Woodhead Publishing, 2016.

Rozdział 14

Uproszczona analiza numeryczna flatteru łopatek wirnikowych ostatniego stopnia turbiny parowej niskiego ciśnienia

Arkadiusz Koprowski, Romuald Rządkowski

Institute of Fluid Flow Machinery, Polish Academy of Sciences,
Fiszera 14, 80-231 Gdansk, Poland

14.1 Flatter

Flatter to zjawisko drgań samowzbudnych powstających pod wpływem sił aerodynamicznych, sprężystej reakcji oraz sił bezwładności [1]. Zwykle pojawia się przy niskich częstotliwościach drgań własnych łopatek lub układu łopatek połączonych z tarczą. Jest ono niebezpieczne, ponieważ może powodować pęknięcie łopatek, skrócić okres między remontami oraz prowadzić do awarii.

Drgania łopatek w turbinach wynikają głównie z śladów krawędziowych wywoływanych przez łopatki poprzedzającej palisady oraz niedokładności wykonania palisad łopatkowych. Ślady krawędziowe powodują wymuszenia wysokoczęstotliwościowe [2]. Zjawisko flatteru występuje natomiast zwykle przy niskich częstotliwościach. Wpływ łopatek kierowniczych wydaje się być jednak istotny dla tego zjawiska. Artykuł ten stanowi wstęp do rozważań mających na celu ustalenie wpływu niestacjonarnych zjawisk powodowanych przez łopatki kierownicze na wartości oraz rozkład współczynników tłumienia aerodynamicznego.

W tym artykule analizie poddano ostatni stopień niskociśnieniowej części turbiny 18K370 wyprodukowanej w zakładach ZAMECH w Elblągu na licencji firmy BBC Baden [3]. Przeprowadzono analizę modalną, obliczenia przepływu oraz analizę flatteru dla zadanego kąta przesunięcia fazowego między łopatkami wynoszącego -90° metodą energetyczną dla modelu uwzględniającego cały stopień oraz dla modelu zakładającego jednorodny rozkład parametrów na wylocie z kierownic.

14.2 Wpływ łopatek kierowniczych na drgania łopatek wirnikowych w turbinach i sprężarkach osiowych

Istnienie wpływu łopatek kierowniczych na flatter łopatek wirnikowych w turbinach i sprężarkach osiowych opisano dotychczas w różnych publikacjach.

Badania przeprowadzone przez Manwaringa i Wislera w [4] poka-

zują, że sąsiadujące palisady mają istotny wpływ na siły aerodynamiczne działające na łopatki wirnikowe, które udało im się zmierzyć.

Xiaojie w [5] opisał wpływ wlotowych łopatek kierowniczych do sprężarki oraz łopatek kierowniczych znajdujących się za łopatką wirnikową na charakterystykę flatterową łopatek wirnikowych. W artykule wykonano obliczenia numeryczne metodą objętości skończonych dla trójwymiarowych modeli. Na ich podstawie stwierdzono, że w każdym z rozważanych przypadków, czyli dla uwzględnienia wpływu jedynie łopatek wlotowych, dla uwzględnienia wpływu łopatek kierowniczych za wirnikiem oraz dla przypadku kontrolnego, w którym nie uwzględniono ani poprzedzających ani następujących po wirniku kierownic współczynnik tłumienia aerodynamicznego jest inny. Autorzy wskazali, na to, że współczynnik tłumienia aerodynamicznego w przypadku uwzględniającym kierownice wlotowe ma wartość o 150% wyższą niż w przypadku izolowanego wirnika. Autorzy poszukiwali również optymalnej metody modelowania wpływu poprzedzających i następujących palisad i analizują wpływ odstępu pomiędzy poszczególnymi palisadami.

Podobnym zagadnieniem zajął się Zhang w [6]. Przeprowadził on obliczenia metodami numerycznej mechaniki płynów dla izolowanego wirnika wentylatora NASA Rotor 67 oraz dla tego samego wirnika z uwzględnieniem kierownic wlotowych. Pokazano, że jeżeli częstotliwość zakłóceń wywoływanych przez kierownice wlotowe w wyniku ruchu obrotowego wirnika jest bliska częstotliwości drgań własnych, to może ona spowodować wystąpienie flatteru, nawet gdy w wyniku obliczeń wirnika izolowanego uzyskano wysokie wartości współczynnika tłumienia aerodynamicznego. Gdy częstotliwość wynikająca z łopatek kierowniczych była znacząco różna od częstotliwości drgań własnych, wartość współczynnik tłumienia aerodynamicznego istotnie się różniła między w zależności od uwzględnienia kierownic i ich nie uwzględniania. W przedstawionych wynikach, przy uwzględnieniu kierownic, wartość współczynnika tłumienia aerodynamicznego była niższa.

Zagadnieniem flatteru z uwzględnieniem łopatek kierowniczych odniesieniu do turbin parowych zajął się Huang w [7]. Analizował

on ostatni stopnień turbiny parowej niskiego ciśnienia. Przeprowadził obliczenia dla modeli dwuwymiarowych oraz trójwymiarowych. Z jego obliczeń wynika, że obecność łopatek kierowniczych w modelu znacznie zmienia charakterystykę stabilności łopatek wirnika pod względem drgań samowzbudnych. Wskazana jest również możliwość optymalizacji wielkości odstępu między łopatkami wirnikowymi i kierowniczymi, tam aby kontrolować występowanie flatteru.

Kubitz w [8] przestawił wyniki obliczeń dla łopatki sprężarki silnika SO-3. W symulacjach została wykorzystana metoda energetyczna. Uzyskane wartości współczynnika tłumienia aerodynamicznego różnią się między symulacjami uwzględniającymi łopatki kierownicze i bez łopatek kierowniczych. Ponadto w obu przypadkach ujemne wartości współczynnika tłumienia aerodynamicznego pojawiają się dla innych kątów przesunięcia fazowego między łopatkami (IPBA).

Dynamika analizowanej w tym artykule łopatki wirnikowej była analizowana w [3]. W przypadku metody energetycznej ujemne współczynniki tłumienia aerodynamicznego wystąpiły dla pierwszej częstotliwości drgań własnych dla kątów IPBA zbliżonych do -90°.

W artykule przeanalizowano ostatni stopień turbiny niskiego ciśnienia. Przeprowadzona została analiza modalna, obliczenia przepływu oraz analiza flatteru dla zadanego kąta przesunięcia fazowego między łopatkami wynoszącego -90°.

14.3 Metody przewidywania wystąpienia flatteru łopatek turbinowych

Analiza flatteru w turbinach opiera się o rozwiązanie równania ruchu:

$$M\ddot{x} + D\dot{x} + Kx = F(t) \tag{14.1}$$

gdzie M, D i K to macierze mas, tłumienia i sztywności, x to wektor przemieszczenia, a $F(t)$ to wektor sił aerodynamicznych, który zmienia się w czasie.

Metody modelowania flatteru można podzielić pod względem podejścia do analizy oddziaływania przepływu z częścią mechaniczną na dwie kategorie. Wyróżnia się metody klasyczne oraz metody cał-

kowe. Metody klasyczne charakteryzują traktowaniem nieliniowego zjawiska wzajemnego oddziaływania przepływu i struktury jako oddzielne zagadnienia o charakterze liniowym. Metody całkowe rozpatrują obszar przepływowy oraz strukturalny jako jeden ciągły ośrodek [8].

14.3.1 Metody klasyczne

Metody klasyczne traktują zjawiska wzajemnego oddziaływania przepływu i struktury jako oddzielne zagadnienia.

Jednym z ważnych pojęć w aerosprężystości w turbinach jest kąt przesunięcia fazowego między łopatkami (interblade phase angle IBPA). Zakłada się, że gdy nie występuje rozstrojenie łopatek, to łopatki w pojedynczej palisadzie drgają z tą samą amplitudą, częstotliwością oraz formą. Lecz między łopatkami występuje pewne przesunięcie w fazie, którym jest IBPA. Kąt ten jednocześnie odpowiada liczbie średnic węzłowych ułopatkowanej tarczy [8]. Metody klasyczne często opierają się o wyznaczenie drgań swobodnych. Forma oraz częstotliwość drgań swobodnych następnie służy do wyznaczania sił niestacjonarnych. Na podstawie wyznaczonych sił niestacjonarnych określa się stabilność układu pod względem wystąpienia flatteru.

Zalety metod klasycznych to między innymi możliwość uproszczenia modeli, a wraz z tym mniejszy koszt obliczeniowy. W wielu przypadkach uzyskiwana z obliczeń informacja o stabilności lub jej braku jest wystarczająca, szczególnie przy wstępnych obliczeniach projektowych.

Wadami tych metod jest upraszczanie przez nie zjawiska, szczególnie w przypadkach, w których w przepływie występują fale uderzeniowe. Zapobiega się temu zadając drgania o małej amplitudzie, pozwalającej osiągnąć do pewnego stopnia liniową odpowiedź przepływu na drgania łopatek.

Jedną z metod klasycznych jest metoda energetyczna. Polega na wyznaczeniu sumy pracy wykonywanej przez siły aerodynamiczne działające na łopatki, które poruszają się zgodnie z przypisanym ruchem zgodnym z formą i częstotliwością drgań swobodnych. Metoda

ta zakłada, że zjawisko flatteru zachowuje się liniowo. Jeżeli energia jest doprowadzana do łopatki, czyli gdy praca aerodynamiczna jest dodatnia, oznacza to, że przypadek nie jest stabilny. W przeciwnej sytuacji przypadek jest stabilny. Ustalenie, czy ułopatkowana palisada jest stabilna wymaga przeprowadzenia obliczeń dla każdego kąta fazowego oraz dla każdej częstotliwości i formy drgań własnych.

14.3.2 Metody całkowe

Metody całkowe nie rozdzielają analiz przepływu od analiz ruchu części mechanicznych. Używane są, aby uwzględnić nieliniowe zjawiska pojawiające się w niestacjonarnych analizach przepływów płynów lepkich o wysokiej prędkości. W szczególności dąży się do uchwycenia wpływu zjawisk takich jak fale uderzeniowe i oderwania warstwy przyściennej. Kolejnym uzasadnieniem dla stosowania tych metod jest uwzględnienie zjawisk nieliniowych takich jak tłumienie mechaniczne w obliczeniach strukturalnych [8].

- Całkowanie częściowe

  Przy całkowaniu częściowym rozwiązania równań dla przepływu oraz dla części mechanicznych są wyznaczane osobno, ale przy każdym kroku czasowym informacje dotyczące obciążeń i przemieszczeń są przekazywane między domenami. Powierzchnie, którymi się stykają dwie domeny stanowią warunki brzegowe dla siebie nawzajem.

- Periodic Mode Updating Method

  Jest to metoda będąca połączeniem metody klasycznej i metody całkowania częściowego. Drgania swobodne są wyznaczane w domenie częstotliwościowej. Wyniki służą do przeprowadzenia obliczeń metodami numerycznej mechaniki płynów dla odpowiedniej częstotliwości i formy drgań uzyskanej z obliczeń drgań swobodnych. Po zasymulowaniu pierwszego okresu, forma drgań jest wyznaczana ponownie, tym razem z uwzględnieniem współczynnika tłumienia aerodynamicznego wyznaczonego z obliczeń przepływu. Obliczenia są powtarzane w takim cyklu, aż do osiągnięcia zbieżności form

drgań własnych.

- Całkowanie pełne

  Metody wykorzystujące całkowanie pełne polegają opisaniu całego modelu jednym układem równań. W przeciwieństwie całkowania częściowego równania dotyczące przepływu i struktury są rozwiązywane jednocześnie. Metody te ze względu na brak różnicy w czasie między przemieszczeniem struktury a przekazaniem obciążeń na strukturę dokładniej modelują transfer energii.

14.3.3 Modele o zredukowanym rzędzie dokładności

Modele o zredukowanym rzędzie dokładności cieszyły się zainteresowaniem w przeszłości, ze względu na niski koszt obliczeń. Niestety często prowadziły do niedokładnych wyników. Współcześnie większe znaczenie mają metody wykorzystujące metody numerycznej mechaniki płynów. Jednak przy modelowaniu na przykład przypadków, które nie zachowują osiowej symetrii lub innych wymagających zamodelowania całej palisady metody te nadal pozostają przydatne. Są to między innymi:

- Projection methods [9]

  Metody te polegają na wykorzystaniu niewielkiej liczby rozwiązań dla pełnego modelu. Wykonuje się obliczenia dla różnych warunków w zakresie roboczym, w których istotna jest dokładność wyników dla przepływu. Następnie tworzone są wektory wychwytujące zachowanie przepływu. W zależności od zagadnienia mogą to być na przykład rozkłady przestrzenne parametrów przepływu. Uzyskane wcześniej wyniki są rzutowane na przestrzeń reprezentowaną przez wektory. Na tej podstawie rekonstruowane jest rozwiązanie dla całego przepływu. W tej metodzie często jest wykorzystywana metoda pochodząca z analizy sygnałów proper orthogonal decomposition (POD).

- Aerodynamic Influence Coefficients [9]

  W metodzie tej przyjmuje się założenia dotyczące małych

przemieszczeń, zakłada się liniowość, nakładanie się zjawisk oraz że zakłócenia przepływu wynikające z drgań jednej łopatki wpływają jedynie na kilka sąsiadujących kanałów i szybko wygasają. Symulacji podlega nieparzysta liczba łopatek. Środkowa łopatka drga zgodnie z przypisanym wcześniej ruchem. Wpływ jej drgań jest wyrażony niestacjonarnym rozkładem ciśnienia jest zauważalny na pozostałych nieporuszających się łopatkach. Niestacjonarne obciążenia można wyrazić liczbami zespolonymi przy pomocy transformacji Fouriera. Taki zabieg można przeprowadzić dla każdej z łopatek i następnie połączyć wpływ wszystkich. Pozwala to na analizę rozstrojonych łopatek. Mimo przyjmowania założenia o liniowości i nakładaniu się zjawisk metoda ta pozwala na uzyskanie dość dokładnych wyników również dla przepływów naddźwiękowych.

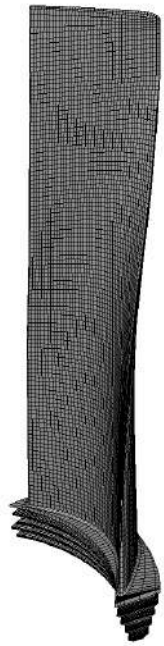
W ostatnim czasie pojawiają się nowe metody wykorzystujące sztuczne sieci neuronowe pozwalające symulować odpowiedź przepływu na różne zakłócenia. Ich zaletą jest możliwość uwzględnienia zjawisk nieliniowych. Wadą jest natomiast wysoki koszt obliczeniowy związany z uczeniem sieci [9].

14.4 Obliczenia modalne

Obliczenia modalne wykonano przy pomocy programu Ansys Mechanical APDL. Obliczenia modalne wykonano dla modelu łopatki, który uwzględniał jej stopkę oraz pióro. Na Rys. 14.1 przedstawiono siatkę obliczeniową modelu. Liczba elementów wynosi 21724. Własności materiału, dla którego wykonywano obliczenia przedstawiono w Tab. 14.1.

Tab. 14.1: Własności materiału łopatki w obliczeniach modalnych.

Własność	**Wartość**
Moduł Younga	211.125 GPa
Gęstość	7800 $\frac{kg}{m^3}$
Współczynnik Poissona	0.2865



Rys. 14.1: Siatka obliczeniowa do obliczeń modalnych łopatki.

W odpowiednich miejscach stopki łopatki odebrano wszystkie stopnie swobody. W pierwszym kroku zadano prędkość obrotową wirnika wynoszącą 3000 $\frac{obr}{min}$. Wyniki z tego kroku takie jak odkształcenie łopatki przekazywane są jako wstępne naprężenie do drugiego kroku. W drugim kroku wykonywane są obliczenia modalne. W ich wyniku uzyskiwane są częstości oraz formy drgań własnych łopatki. Częstości drgań łopatki przedstawiono w Tab. 14.2. Formę drgań dla pierwszej częstości przedstawiono na Rys. 14.2. Maksymalne przemieszczenia występują w lewym górnym rogu łopatki.

Wyniki dotyczące częstości i formy drgań własnych łopatki zostaną w dalszej części wykorzystane przy wyznaczaniu współczynnika tłumienia aerodynamicznego.

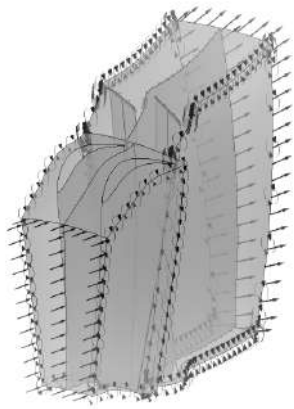
Tab. 14.2: Częstości drgań własnych łopatki

lp.	Częstotliwość [Hz]
1	119.13
2	179.44
3	257.03
4	344.46
5	496.94
6	530.68
7	605.51
8	765.93
9	920.92
10	1014.4

Rys. 14.2: Forma drgań dla pierwszej częstości drgań własnych łopatki.

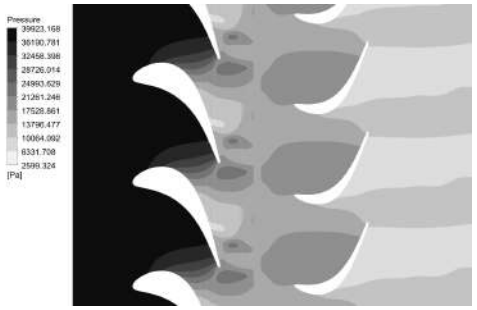
14.5 Obliczenia przepływowe

Obliczenia przepływowe wykonano przy pomocy programu Ansys CFX [10]. W pierwszym kroku wykonano stacjonarne obliczenia inicjalizacyjne dla przypadku uwzględniającego łopatki kierownicze. Warunki brzegowe, jakie przyjęto na wlocie do kierownic to ciśnienie całkowite na poziome 39900 Pa i temperaturą całkowitą wynoszącą 75,8°C. Na wylocie z wirnika przyjęto natomiast średnie statyczne ciśnienie wynoszące 10450 Pa. Prędkość obrotowa wirnika wynosiła 3000 $\frac{obr}{min}$. Interfejsem pomiędzy siatkami zawierającymi łopatki kierownicze a siatkami zawierającymi łopatki wirnikowe w obliczeniach stacjonarnych był Mixing-Plane z ustawieniem constant total pressure a w obliczeniach niestacjonarnych Transient Rotor Stator. Model przedstawiono na Rys. 14.3. Natomiast rozkład ciśnienia uzyskany w wyniku obliczeń przedstawiono na Rys. 14.4. Liczba objętości skończonych w obliczeniach wynosiła 3 270 062.



Rys. 14.3: Model obliczeniowy przepływu.

Jako model gazu przyjęto zawarty w programie Ansys CFX model Water Ideal Gas [10]. Jest to model oparty o równanie stanu gazu doskonałego. Ze względu na pracę stopnia w warunkach pary mokrej, uzyskane przy pomocy tego modelu wyniki dotyczące temperatury są nierealistyczne. Pociąga to za sobą błędy dotyczące parametrów takich jak gęstość. Może także spowodować, że w wynikach



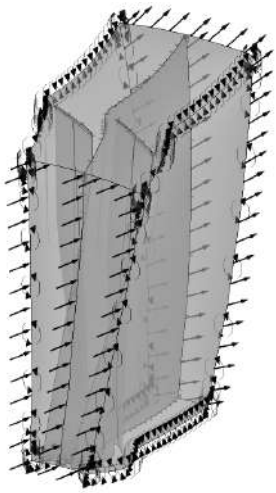
Rys. 14.4: Rozkład ciśnienia w przepływie w przekroju w połowie wysokości kanału przepływowego.

zawirowania będą pojawiać się w innych miejscach, niż w rzeczywistości lub przy wykorzystaniu modelu dwufazowego gazu rzeczywistego. Wpływ modelu gazu na współczynnik tłumienia aerodynamicznego rozpatrzono na przykład w [11].

W symulacjach przyjęto model turbulencji SST [10], który jest odpowiedni do tego typu analiz [8].

W kolejnym kroku przygotowano obliczenia nie uwzględniające łopatek kierowniczych. Warunek brzegowy dotyczący wlotu, to w tym przypadku rozkład ciśnienia całkowitego, temperatury całkowitej oraz rzutów wektorów jednostkowych prędkości na osie układu współrzędnych na wlocie do łopatek wirnikowych. Rozkłady te uzyskano na podstawie wyników z modelu uwzględniającego łopatki kierownicze. Pozostałe warunki brzegowe pozostały takie same, jak w modelu uwzględniającym łopatki kierownicze. Model przedstawiono na Rys. 14.5 Natomiast rozkład ciśnienia uzyskany w wyniku obliczeń przedstawiono na Rys. 14.6. Liczba objętości skończonych w obliczeniach wynosiła 2992268.

W obliczeniach przepływowych nie uwzględniono odkształcenia łopatki wynikającego z działania sił odśrodkowych.



Rys. 14.5: Model obliczeniowy przepływu.

Rys. 14.6: Rozkład ciśnienia w przepływie w przekroju w połowie wysokości kanału przepływowego.

14.6 Obliczenia flatteru

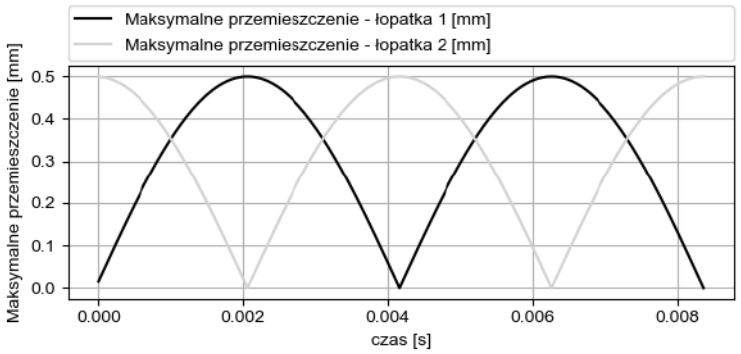
Obliczenia flatteru wykonano przy pomocy programu Ansys CFX, wcześniej opisaną metodą energetyczną. Obliczenia wykonano przy pomocy metody transformacji Fouriera. W tym przypadku pozwala ona zmniejszenie liczby kanałów przepływowych do dwóch. W metodzie tej fazie warunek brzegowy związany z periodycznością przepływu jest przesunięty w fazie. Przepływ w przekrojach sąsiednich kanałów przepływowych jest periodyczny. Pozwala to na zastosowanie szeregu Fouriera do przechowywania wyników symulacji na brzegach obszaru.

Wymuszenia związanie z oddziaływaniami między łopatkami wirnikowymi a kierownicami są przekazywane przez odpowiednie równanie, które przekazuje dane o przepływie między wylotem z kanałów przepływowych kierownic a wlotem do kanałów przepływowych łopatek wirnikowych, tak aby zachować odpowiedni rozkład mimo różnej liczby łopatek w obu palisadach.

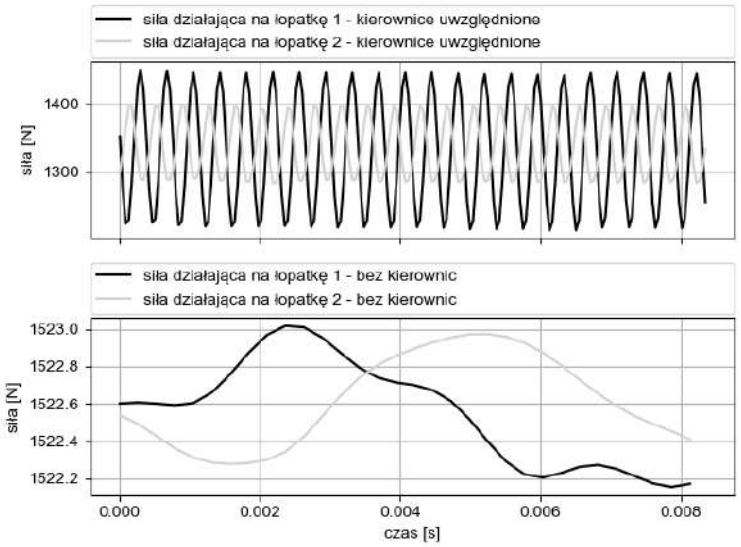
Celem takich obliczeń jest ustalenie, czy stopień jest stabilny pod względem flatteru, czy nie. Aby to ustalić wyznaczany jest współczynnik tłumienia aerodynamicznego [8].

Obliczenia przeprowadzono dla kąta przesunięcia fazowego IBPA równego -90° i dla formy oraz częstotliwości odpowiadającym pierwszej częstotliwości drgań własnych. Forma została znormalizowana do amplitudy wynoszącej 1 mm. Wykonano symulację dla przypadku pomijającego wpływ zakłóceń przepływu wynikających z łopatek kierowniczych oraz symulację je uwzględniającą.

Rysunek 14.7 przedstawia największe przemieszczenie węzłów w danym czasie. Przemieszczenia te są informacją wejściową do obliczeń flatteru. Rysunek 14.8 przedstawia zmienność sił wypadkowych działających na łopatki występujące w modelu. W przypadku braku uwzględnienia kierownic siły działające na łopatki były wyższe, co prawdopodobnie wynika z umieszczenia warunku brzegowego blisko rozpatrywanego obszaru. Przypadek ten charakteryzuje się znacznie mniejszą zmiennością sił, prawie pomijalną w porównaniu do zmienności zauważalnej w modelu uwzględniającym kierownice.



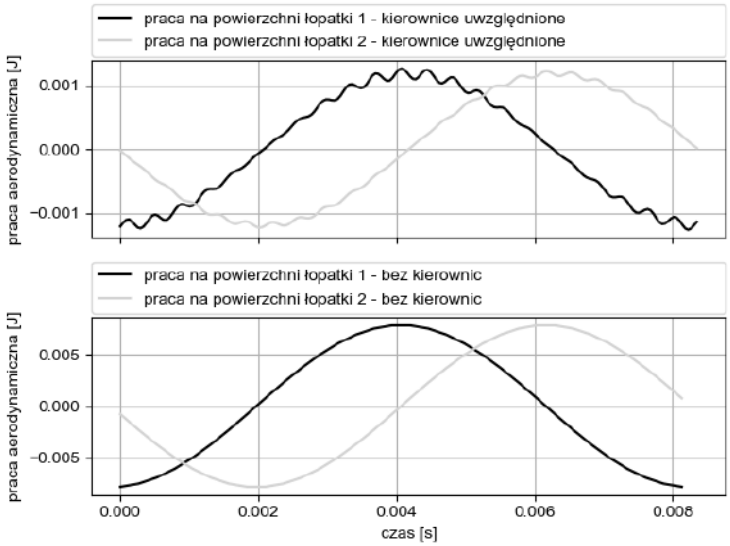
Rys. 14.7: Maksymalne przemieszczenie węzłów łopatki w funkcji czasu.



Rys. 14.8: Siły działające na łopatki.

Na Rys. 14.9 przedstawiono zmienność pracy aerodynamicznej na powierzchni łopatki. Amplituda tej pracy jest wyższa, gdy kierownice nie są uwzględniane, co może być związane z wyższą wartością siły w tym przypadku. Kierownice oraz występujący wraz z nimi

ślad krawędziowy powodują dodatkowe zakłócenia pracy aerodynamicznej.



Rys. 14.9: Praca aerodynamiczna na powierzchni łopatek

Współczynnik tłumienia aerodynamicznego znormalizowano przy pomocy zależności [8]

$$D = \frac{W}{m \cdot f^2 \cdot A^2} \tag{14.2}$$

gdzie:

- W – praca aerodynamiczna uzyskana w wyniku obliczeń [J]
- m – masa łopatki [kg]
- f – częstotliwość [Hz]
- A – amplituda [m]

Dla modelu pomijającego wpływ kierownic wyniósł 0,000425 a w przypadku uwzględniającym 0,000922. Uwzględnienie kierownic wpłynęło na wartość współczynnika tłumienia aerodynamicznego. Jednak obie te wartości są bardzo bliskie zera. Ich znak i różnica wartości może się zmienić w wyniku błędu numerycznego.

14.7 Podsumowanie

Przeprowadzone obliczenia potwierdzają, że uwzględnienie w obliczeniach niestacjonarnych zjawisk wywoływanych przez łopatki kierownicze może wpływać na wartość współczynnika tłumienia aerodynamicznego. Jednak uzyskane wyniki nie są wystarczające. Należy kontynuować obliczenia, aby otrzymać wartości współczynnika tłumienia dla większej liczby kątów IBPA.

Uzyskane wyniki nie są zgodne z występującymi w literaturze [3]. Może to być spowodowane brakiem uwzględnienia w obliczeniach przepływowych odkształcenia łopatki wynikającego z prędkości obrotowej lub przyjęciem innych warunków brzegowych.

Z przeprowadzonych obliczeń wynika, że przy kącie przesunięcia fazowego wynoszącym -90°, dla warunków w podanych w rozdziale 14.5 zjawisko flatteru nie powinno wystąpić.

Bibliografia

[1] Earl H. Dowell. *A Modern Course in Aeroelasticity*. Springer, 2015.

[2] Stefan Perycz. *Turbiny parowe i gazowe*. Zakład Narodowy im. Ossolińskich, 1992.

[3] Romuald Rządkowski and Ryszard Szczepanik. *Dynamics of Last Stage Low Pressure Steam Turbine Rotor Blades*. Wydawnictwo Instytutu Technicznego Wojsk Lotniczych, 2017.

[4] S. R. Manwaring and D. C. Wisler. Unsteady aerodynamics and gust response incompressors and turbines. *Journal of turbomachinery*, 115(4):724–740, 1993.

[5] Zhang Xiaojie, Wang Yanrong, Han Le, Zhao Jiazhe, and Luo Yanbin. Influencełł of upstream and downstream stator blades on the rotor blade flutter characteristics. *ASME 2018 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 8, 2018.

[6] Chen an Zhang, Zhengyin Ye, and Feng Liu. Numerical researches on aeroelastic problem of a rotor due to igv/fan interaction. *47th AIAA Aerospace Sciences Meeting including The New Horizons Forum and Aerospace Exposition*, 2009.

[7] X. Q. Huang, L He, and D. L. Bell. Influence of upstream stator on rotor flutter stability in a low pressure steam turbine stage. *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, 220(1):25–35, 2006.

[8] Leszek Kubitz, Romuald Rządkowski, and Ryszard Szczepanik. *Flatter łopatek wirnikowych turbin gazowych i silników lotniczych.* Wydawnictwo Instytutu Technicznego Wojsk Lotniczych, 2017.

[9] Marco Casoni and Ernesto Benini. A review of computational methods and reduced order models for flutter prediction in turbomachinery. *Aerospace*, 8, 2021.

[10] Inc. ANSYS. *ANSYS CFX-Solver Theory Guide Release 2020-R1.* 2020.

[11] Christopher Fuhrer, Marius Grübel, Damian M. Vogt, and Paul Petrie-Repar. The influence of non-equilibrium wet steam effects on the aeroelastic properties of a turbine blade row. *ASME Turbo Expo 2016: Turbomachinery Technical Conference and Exposition*, 8, 2016.

Rozdział 15

Wprowadzenie do metod optymalizacyjnych i przedstawienie koncepcji ich wykorzystania przy rozruchu turbiny parowej

Jan Przytulski

Institute of Fluid Flow Machinery, Polish Academy of Sciences,
Fiszera 14, 80-231 Gdansk, Poland

15.1 Wstęp

Polski system elektroenergetyczny stoi w najbliższych latach przed dużym wyzwaniem przystosowania istniejącego systemu do nowej specyfiki pracy. Zwiększający się udział niestabilnych źródeł energii (OZE) w miksie energetycznym wymusza pracę JWCD (pol. Jednostek Wytwórczych Centralnie Dysponowanych) ze zwiększoną elastycznością w szerokim tego słowa znaczeniu, tj. krótsze czasy uruchomień/ odstawień, dynamiczniejsze nabory i zrzuty mocy, większa liczba cykli w czasie roku, czy zmniejszenie minimum technicznego.

Nowe jednostki wytwórcze projektowane są na takie uwarunkowania. Biorąc jednak pod uwagę uwarunkowania polityczne (polityka dekarbonizacyjna), ekonomiczne (koszt i czas nowej inwestycji) oraz bezpieczeństwa energetycznego RP (zapewnienie nieprzerwanych dostaw prądu przy rosnącym zapotrzebowaniu), nowe jednostki nie rozwiążą problemu w pełni. Konieczne jest modernizowanie istniejących JWCD pod kątem wydłużenia czasu eksploatacji i zwiększenia ich elastyczności.

Określenie stanu naprężeń w czasie rzeczywistym pozwala przeprowadzać rozruchy/ odstawienia/ zmiany obciążeń w sposób zarówno bezpieczny dla głównych węzłów konstrukcyjnych jak i optymalny pod względem czasu trwania danego stanu. Najczęściej w tym celu stosuje się metody pośrednie. Za pomocą sond rozruchowych mierzy się temperaturę w lokalizacjach krytycznych, określanych drogą analiz teoretycznych i doświadczeń eksploatacyjnych. Gradient temperatury w czasie rzeczywistym jest następnie wyko-rzystywany do określenia naprężeń. Wadą takiego rozwiązania jest skomplikowana konstrukcja i eksploatacja sond rozruchowych, trudność uzyskania rzeczywistej wartości pomiarowej i lokalny charakter pomiaru.

Celem rozprawy doktorskiej ”Optymalizacja rozruchu turbiny parowej z wykorzystaniem algorytmów genetycznych”realizowanej w ramach III edycji programu ”Doktorat Wdrożeniowy”jest skrócenie czasu rozruchów, czy też naborów/zrzutów mocy z zachowaniem optymalnego poziomu naprężeń i minimalizacji zużycia głównych komponentów turbiny. Nowe możliwości stwarza wykorzystanie al-

gorytmów genetycznych, w systemie ograniczeń termicznych turbiny. W literaturze [1–3] udowodniono, że sieci neuronowe i algorytmy genetyczne z dużą dokładnością mogą odwzorować rzeczywisty stan naprężeń w krytycznych lokalizacjach w czasie rzeczywistym (ang. on-line). Modelowanie stanu naprężeń algorytmami genetycznymi otwiera pole do optymalizacji rozruchu w czasie rzeczywistym, co nie było wcześniej rozpatrywane, a co może przynieść wielorakie korzyści.

W pracy pt. "Mechanizmy zużycia i żywotność turbin parowych" [4] w ramach corocznej monografii Trójmiejskiej Szkoły Doktorskiej PAN Doktorant przedstawił wiedzę pozyskaną z szerokiego zakresu literatury. Temat ten został wybrany również ściśle pod kątem tematu doktoratu. Istotne jest bowiem poznać fizykę zjawisk wpływających na zużycie i żywotność turbin parowych, jeśli chce się optymalizować rozruchy pod kątem elastyczności przy jednoczesnej minimalizacji spadku żywotności kluczowych komponentów. W pracy tej przedstawiono aspekty eksploatacyjne, żywotnościowe i związane z systemami ograniczeń termicznych turbin parowych.

Kolejnym krokiem było wykonanie modelu obliczeniowego i przeprowadzenie analizy MES rozruchu turbiny. Wyniki prac Doktorant przedstawił w pracy pt. "Obliczenia zmęczenia niskocyklowego wirnika turbiny parowej metodą elementów skończonych pod kątem optymalizacji rozruchów" [5] w ramach corocznej monografii Trójmiejskiej Szkoły Doktorskiej PAN. Na wstępie w pracy opisano po krótce wybraną maszynę i uzasadniono wybór konkretnego węzła konstrukcyjnego i typu rozruchu poddanego analizie. W następnym kroku przedstawiono tworzenie modelu i definicję warunków brzegowych. Obliczenia wykonano z wykorzystaniem komercyjnego programu ANSYS Mechanical oraz dodatkowych narzędzi będących własnością GE Power. Dla stworzonego modelu wykonano obliczenia bazując na krzywych rozruchowych dostarczonych przez producenta turbiny. W następnym kroku wykonano obliczenia dla tego samego typu rozruchu, bazując jednak na pomiarach z rzeczywistego obiektu. Dało to możliwość porównania założeń projektowych z rzeczywistością eksploatacyjną. Możliwe było również dzięki temu pokazanie, jak prowadzenie turbiny podczas rozruchu przekłada się

na jej poziom zużycia.

Na końcu zaproponowano we wnioskach kierunek dalszych prac, w którym opisano możliwe wykorzystanie algorytmu genetycznego, który po zaimplementowaniu do systemu ograniczeń termicznych turbiny, pomagałby w prowadzeniu rozruchów w sposób optymalny. Obliczenia stworzonym na potrzeby niniejszej pracy modelem MES mogą posłużyć jako dane wejściowe dla algorytmu genetycznego optymalizującego rozruch turbiny parowej.

Celem niniejszej pracy jest przedstawienie kolejnego etapu realizacji zadania, tj. obliczeń optymalizacyjnych rozruchu turbiny parowej. W tym celu należy opisać zjawisko, które podlega optymalizacji, zdefiniować funkcję celu i kryteria optymalizacji. Przedstawiono to w Sek. 15.2 Dodatkowo w Sek. 15.2.1 został opisany główny czynnik ograniczający czas rozruchu, tj. żywotność turbiny parowej. W Sek. 15.2.2 dokonano natomiast przeglądu dotychczasowych prób wykonania optymalizacji rozruchów turbin parowych.

Następnie opracowano koncepcję wykonania optymalizacji roz-ruchu turbiny parowej, którą przedstawiono w Sek. 15.3 opisano szczegółowo metodę, jaką Doktorant zamierza rozwiązać problem postawiony w ramach pracy doktorskiej.

Środowisko programistyczne, które zostało wybrane do wykonania optymalizacji z wykorzystaniem algorytmu genetycznego jest Python. Język ten został wybrany, ponieważ dobrze współpracuje ze środowiskiem Workbench, w którym to wykonany został model MES w ramach pracy [5] i prowadzone były obliczenia niestacjonarne. Język ten ponadto dobrze sprawdza się w przeprowadzaniu optymalizacji z użyciem algorytmów genetycznych. Środowisko Python posiada bogatą bibliotekę narzędzi do zagadnień tzw. „machine learningu". Jako, że zagadnienie optymalizacji i algorytmów genetycznych w środowisku Python, jak i samo to środowisko programistyczne jest nową dziedziną nauki dla Doktoranta, należało się z tą dziedziną zapoznać. W Sek. 15.3 przedstawiono to środowisko i opisano, jakie daje ono możliwości, które planuje się wykorzystać w celu realizacji postawionego zadania.

15.2 Opis zjawiska oraz definicja celu i kryteriów optymalizacyjnych

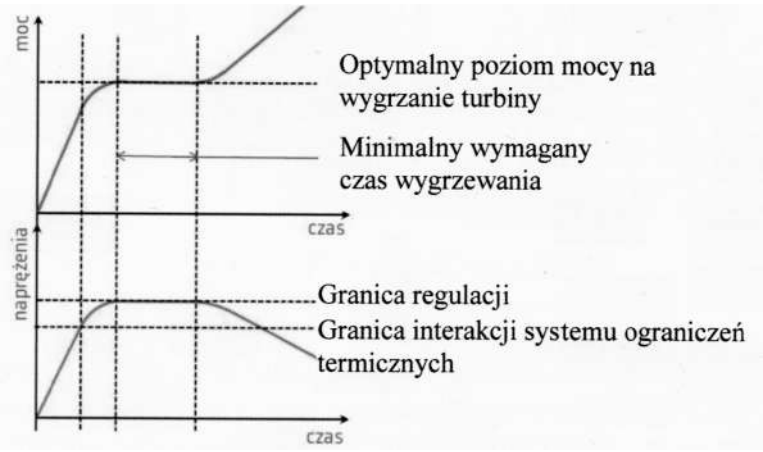
Celem optymalizacji jest maksymalne skrócenie czasu rozruchu przy jednoczesnej minimalizacji zużycia, czyli ograniczeniu wpływu mechanizmów zużycia na główne komponenty konstrukcyjne turbiny.

Należy pamiętać, że ograniczony pod względem elastyczności pracy jest proces nagrzewania podczas rozruchów i studzenia podczas odstawień elementów grubościennych, który musi być kontrolowany w taki sposób, aby naprężenia powstające w wyniku działania niestacjonarnego pola temperatury nie przekroczyły naprężeń dopuszczalnych. Przekroczenie podczas rozruchu naprężeń dopuszczalnych jest równoznaczne z przekroczeniem projektowej amplitudy naprężeń podczas rozruchu, co skutkuje intensyfikacją zmęczenia niskocyklowego, które to jest głównym mechanizmem zużycia turbiny parowej i wpływa w największym stopniu na żywotność turbiny parowej. Pojęcie żywotności turbiny parowej i zjawisko zmęczenia niskocyklowego opisano pokrótce w Sek. 15.2.1.

Główny wniosek płynący z fizyki zjawiska zmęczenia niskocyklowego jest następujący – im mniejsza będzie amplituda naprężeń podczas rozruchu turbiny, tym większą liczbę cykli obciążenia będzie można przeprowadzić. Sposobów na ograniczenie amplitudy naprężeń podczas rozruchów jest kilka – można zmieniać gradient naboru obrotów/mocy, można zmieniać gradient naboru parametrów pary podczas rozruchu, lub można stosować przystanki pomiędzy naborami. Ostatni sposób przedstawiono na Rys. 15.1. Stosując przystanek na wygrzanie turbiny, naprężenia przestają dalej rosnąć i po odpowiednio długim czasie nagrzewania można dalej prowadzić nabór mocy bez wzrostu naprężeń.

Bazując na powyższym opisie zjawiska, można określić zmienne, które będą brane pod uwagę podczas budowy programu do optymalizacji rozruchu turbiny parowej, polegającej na maksymalnym skróceniu czasu rozruchu, przy jednoczesnym zachowaniu projektowanej żywotności:

- gradient naboru obrotów - im większe tempo naboru obrotów, tym szybsze osiągnięcie wartości nominalnej, ale też i szyb-



Rys. 15.1: Przykład optymalnego uruchomienia turbiny [4].

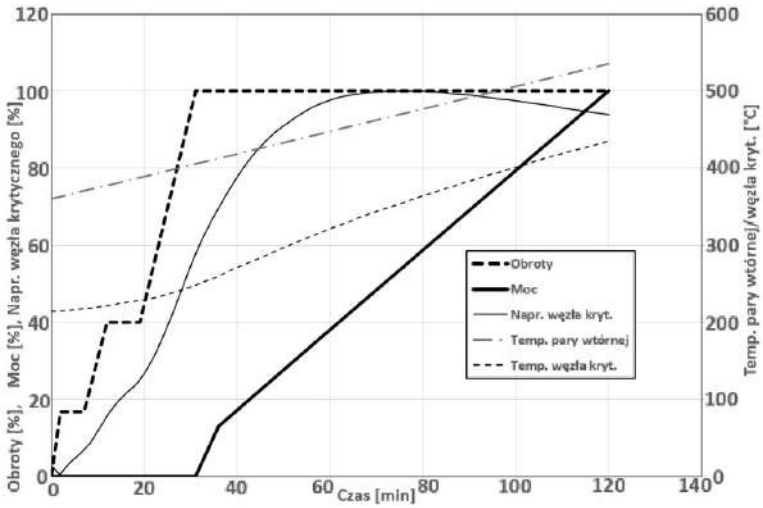
szy przyrost wartości naprężeń. Gradient naboru obrotów jest wartością wejściową zadaną przez operatora turbozespołu. W zależności od typu rozruchu (np. zimny, ciepły, gorący) producent turbiny rekomenduje różne wartości gradientu. Dodatkowo należy mieć na uwadze fakt, że dla obsza-rów prędkości krytycznych, gradient ten nie może być wynikiem optymalizacji, a musi przyjąć wartość zadaną przez producenta turbiny, w celu jak najszybszego przejścia przez obszar prędkości krytycznych. Prędkości krytyczne wyznaczane są w trakcie analizy dynamicznej linii wału i są to takie prędkości, przy której może mieć miejsce zjawisko rezonansu. W celu uniknięcia takiej sytuacji gradient obrotów przyjmuje wartość znacznie wyższy od wartości projektowych (np. 500 obr/min^2 zamiast 100-200 obr/min^2), gdyż bezpieczeństwo eksploatacji i uniknięcie poważnej awarii jest nadrzędne do aspektów żywotnościowych

- gradient naboru mocy - podobnie jak z gradientem naboru obrotów, im większe tempo naboru mocy, tym szybsze osiągnięcie wartości nominalnej, ale też i szybszy przyrost wartości naprężeń. Gradient naboru mocy jest wartością wejściową zadaną przez operatora turbozespołu. W zależności

od typu rozruchu (np. zimny, ciepły, gorący) producent turbiny rekomenduje różne wartości gradientu. Należy również pamiętać, że nabór mocy prowadzony jest po synchronizacji turbozespołu, tj. po osiągnięciu nominalnej prędkości obrotowej. Tak więc gradient naboru obrotów i mocy nie będzie optymalizowany jednocześnie, ale jeden po drugim. Dodatkowo warto wspomnieć, że istnieją różne tryby regulacji turbiny, czyli regulacja na zadaną moc, na zadany przepływ pary świeżej, czy na zadane ciśnienie przed turbiną. Natomiast wszystkie te tryby regulacji są ze sobą skorelowane, ponieważ sprowadzają się do odpowiedniego otwarcia zaworów regulacyjnych przez regulator turbiny. Tak więc w celu uproszczenia zadania rozp-atrywany będzie jedynie tryb regulacji mocy

- gradient naboru temperatury pary zasilającej- jest to zmienna mająca duży wpływ na wartość naprężeń, gdyż jak już wcześniej wspomniano nagrzewanie elementów grubościennych jest głównym czynnikiem limitującym elastyczność pracy turbiny. Gradient ten jest zadawany przez operatora kotła. Należy pamiętać, że nabór temperatury pary zasilającej może być prowadzony zarówno podczas naboru obrotów jak i mocy. Pokazano to na Rys. 15.2. Tak jak i przy poprzednich zmiennych, producent turbiny w zależności od typu rozruchu (np. rozruch zimny, ciepły, gorący) rekomenduje różne wartości tegoż gradientu
- początkowa wartość naprężenia i temperatury w lokalizacji krytycznej - wartości te są dostępne jako sygnał on-line z systemu ograniczeń termicznych turbiny

Koncepcja optymalizacji rozruchu zakłada, że gotowe narzędzie na podstawie stanu początkowego będzie w stanie zaproponować taką charakterystykę rozruchową, aby przeprowadzić rozruch w spo-sób optymalny, tj, najkrótszy, przy jednoczesnym zachowaniu projektowanej żywotności, co sprowadza się do zachowania dopuszczalnej wartości naprężeń.

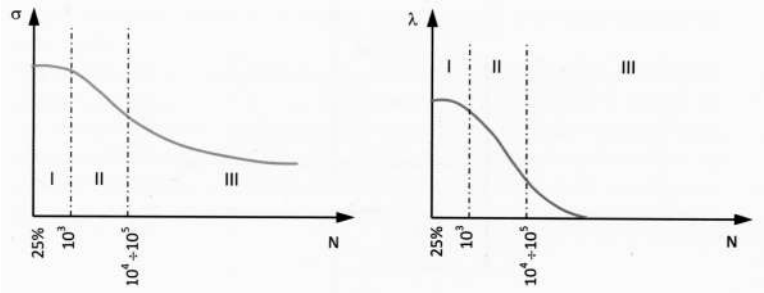


Rys. 15.2: Przykładowa krzywa rozruchowa – rozruch ciepły, turbina klasy 200MW [5].

15.2.1 Żywotność turbiny parowej

Najogólniejszą definicją żywotności jest zdolność elementu do wykonywania określonych funkcji w określonym czasie i określonych warunkach pracy [6]. W literaturze traktującej typowo o turbinach parowych [7] wprowadza się pojęcie zużycia, jako wyczerpanie żywotności, czyli utracenie przydatności danego elementu do dalszej eksploatacji.

W pracy [4] Doktorant pokazał, że głównym mechanizmem zużycia wpływającym na żywotność turbiny parowej jest zmęczenie niskocyklowe. Zjawisko zmęczenia niskocyklowego jest bardzo istotne dla konstrukcji narażonych na zmienne cykle obciążenia w dłuższym czasie. Taką konstrukcją jest turbina parowa, gdzie podczas jednego cyklu pracy, tj. rozruch, praca ciągła i odstawienie, pola temperatur, obciążeń mechanicznych i naprężeń zmieniają się w sposób znaczący. Wraz ze zwiększaniem liczby cykli, zniszczenie może nastąpić przy wartości naprężeń znacznie mniejszych od elementów obciążonych statycznie, co obrazuje wykres Wöhlera przedstawiony na Rys. 15.3.



Rys. 15.3: Wykres Wöhlera [4] σ – naprężenia; λ – odkształcenia, N – liczba cykli.

15.2.2 Przegląd podobnych prac

W pracy [1, 2] dzięki zastosowaniu systemu ograniczeń termicznych turbiny opartego na sieciach neuronowych typu NARX, oprócz wielu zalet (dokładne odwzorowanie stanu termicznego i stanu naprężeń w czasie rzeczywistym, możliwość rezygnacji z pomiarów sondami rozruchowymi i oparcie się na ogólnodostępnych pomiarach ruchowych (parametry pary oraz wydłużenia), możliwa jest również predykcja on-line naprężeń i temperatury w miejscu krytycznym na 5 minut do przodu z zadowalającą zbieżnością. W omawianej pracy jedynie sprawdzono i potwierdzono dokładność predykcji parametrów poprzez sieci NARX, jednak nie badano możliwych korzyści płynących z wykorzystania tejże predykcji w celu optymalnego prowadzenia rozruchu turbiny parowej.

Kolejną metodą możliwą do zastosowaniu w czasie rzeczywistym do optymalizacji rozruchów turbiny parowej są algorytmy genetyczne. W pracy [3] pokazano na prostym przykładzie (nagrzewanie wirnika), jak z pomocą algorytmu genetycznego można zapewnić optymalny nabór temperatury, tj. taki, przy którym naprężenia wynikające z niestacjonarnego pola temperatury będą w pobliżu naprężeń dopuszczalnych. Metoda ta nie była stosowana dla analizy pełnego rozruchu turbiny, niemniej jednak wyniki przedstawione w omawianej pracy pokazują, że algorytmy genetyczne mogą być zastosowane do tego celu i mogą z powodzeniem być zaimplementowane do regulatora turbiny.

Kolejna metoda optymalizacji rozruchów turbin parowych zosała przedstawiona w pracy [8], gdzie w tym celu użyto całki Duhamela, którą wykorzystuje się w złożonych algorytmach Bloków Ograniczeń Termicznych Turbiny. Wynikiem prac było uzyskanie optymalnego przebiegu rozruchu turbiny. W pracy tej również podkreślono możliwość stosowania tej metody w czasie rzeczywistym (ang. on-line).

Metody opisane powyżej zakładają zmiany w systemach ograniczeń termicznych turbiny. Jak zostanie pokazane w Sek. 15.3, koncepcja optymalizacji przedstawiona w niniejszej pracy nie będzie wymagała dokonywania zmian w istniejących systemach ograniczeń termicznych turbiny.

15.3 Środowisko programistyczne Python

Python to język programowania open source opracowany przez holenderskiego programistę Guido van Rossuma, nazwany na cześć brytyjskiej grupy Monty Python. Obecnie rozwijany perz organizację non-profit Python Software Foundation [9]. Python to język [10]:

- interpretowalny - nie jest kompilowany, lecz jest przechowywany w postaci kodu źródłowego i dopiero podczas uruchomienia wczytywany, interpretowany i wykonywany przez interpreter języka;
- wysokiego poziomu - czyli taki, którego kod jest w dużym stopniu zbliżony do języka używanego przez ludzi;
- dynamicznego typowania – to wartość a nie sama zmienna posiada typ, a wartości przekazywane są poprzez przypisanie;
- obiektowy;
- strukturalny;
- funkcyjny;
- modularny – swoiste funkcjonalności zostały wydzielone z języka i są dostępne w postaci bibliotek;
- używający wcięć akapitów – większość języków programowania używa nawiasów klamrowych;

Czytelność i relatywna łatwość użycia, szeroka społeczność programistyczna, bogaty wachlarz bibliotek rozszerzających możliwości środowiska (szczególnie w zakresie sztucznej inteligencji, przetwarzania danych i procesów w chmurze), wszechstronność, wydajność, niezawodność i szybkość spowodowały, że język Python jest najpopularniejszym językiem używanym do prac naukowych i akademickich [11]. Ze względu na wymienione cechy, również do tego projektu środowisko Python jest najlepszym wyborem wraz z bibliotekami:

- NumPy - oferująca rozbudowane funkcje matematyczne, obsługę wielowymiarowych tabel i macierzy;
- Pandas - wydajne w działaniu, relatywnie łatwe w użyciu narzędzie do analizy i manipulacji danymi tabelarycznymi;
- Scikit-learn - narzędzia do statystycznej analizy danych m.in.: klasyfikacji, predykcji;
- TensorFlow - kompleksowa biblioteka do uczenia maszynowego;
- PyAnsys - umożliwia korzystanie z technologii środowiska Workbench bezpośrednio z poziomu kodu programu;
- Matplotlib - obszerna biblioteka do tworzenia statycznych i interaktywnych wykresów i wizualizacji;

Tworzenie kodu musi odbywać się również w odpowiednim edytorze. Do pracy nad projektem wybrano JupyterLab. Jest to środowisko programistyczne pozwalające tworzyć notatniki zawierające kod i dane. Elastyczny interfejs pozwala konfigurować i organizować przepływ pracy poprzez możliwość dzielenia pliku na bloki kodu (ang. chunks), co znacząco ułatwia prace rozwojowe.

15.4 Koncepcja optymalizacji rozruchu turbiny parowej

Problem optymalizacji, przy uwzględnieniu wszystkich czynników technicznych opisanych w poprzednich pracach Doktoranta, sprowadza się z matematycznego punktu widzenia do maksymalizacji poniższej wartości chwilowej q:

$$q = \frac{1}{f_{dop} - f}$$

Gdzie f to chwilowe naprężenie zależne od zmiennych opisanych w Sek. 15.2.1 a f_{dop} to dopuszczalne naprężenie podane przez producenta. W tym miejscu należy rozróżnić dwa główne narzędzia matematyczne, jakie powinny zostać użyte w celu predykcji i optymalizacji. Pierwszym z nich powinien być uniwersalny model predykcyjny, dzięki któremu można przewidzieć wartości naprężeń f powstałych w turbinie na podstawie pozostałych zmiennych, mających wpływ na te naprężenia. Model taki musi odwzorowywać zachowanie układu w różnych stanach, a z matematycznego punktu widzenia, przy różnych wartościach zmiennych wpływających na tenże stan. Zachowanie turbiny jest procesem nieliniowym i wielowymiarowym, stąd właściwym narzędziem do zamodelowania takiego środowiska są sieci neuronowe. Problematyka predykcji została szerzej opisana w Sek. 15.4.2. Mając model, który odpowiada rzeczywistym zjawiskom wytrzymałościowym w miejscu krytycznym turbiny, można przystąpić do optymalizacji parametrów rozruchu, czyli maksymalizacji q. Istnieje wiele algorytmów optymalizacyjnych, jednak jednym z najwydajniejszych są algorytmy genetyczne, które działają na zasadzie losowej mutacji populacji. Problematyka optymalizacji została szerzej opisana w Sek. 15.4.3. W Sek. 15.4.1 natomiast opisano zagadnienia tzw. preprocessingu danych.

15.4.1 Tworzenie i praca ze zbiorem danych

Jak pisano we wstępie tej pracy, stworzony model MES jest bazą do stworzenia zbioru danych wejściowych dla programu optymalizującego rozruch turbiny parowej. Podstawą Machine Learningu jest liczność doświadczeń na których model może się uczyć jak i jakość poszczególnych rekordów danych. Jeden przykład uczący krzywej rozruchowej to zdecydowanie za mało by podjąć próbę stworzenia uniwersalnej funkcji. Z tego powodu pierwszym etap pracy nad zbiorem powinno być stworzenie skryptów i procedur poprzez środowisko Python, które zautomatyzuje produkcję przykładowych odpowiedzi modelu MES przy różnych stanach początkowych i wprowadzonych charakterystykach cech opisanych

w Sek. 15.2. Jakość modelu uczenia maszynowego w przeważającej większości zależy od ilości i jakości danych wejściowych do procesu uczenia i testowania. Negatywnymi zjawiskami, które mogą wpłynąć na przydatność danych są przede wszystkim błędy wynikające z nieprawidłowego działania przyrządów pomiarowych, bądź czynnika ludzkiego. Rekordy zawierające dane nieobarczone błędem ale znacząco odstające od spodziewanych wartości (z ang. outlier), wynikające z anomalii występujących podczas pomiarów rzeczywistych również nie wpływają pozytywnie na skuteczność modelu i powinny zostać usunięte ze zbioru.

W przypadku rozpatrywanego problemu i przyjętego podejścia, jakość danych nie jest znaczącym wyzwaniem w procesie tworzenia programu. Dane uzyskiwane z modelu MES nie są obarczone brakami wartości, czy zakłóceniami występującymi w rzeczywistych pomiarach.

Skalowanie zmiennych

W drodze budowania modelu predykcyjnego algorytm uczący przypisuje każdej zmiennej objaśniającej wagę, która oznacza jak dużo dana cecha wnosi do oszacowania zmiennej objaśnianej. Można to interpretować jako ważność danej cechy. Fakt, że w przedstawionym problemie optymalizacji każda z cech opisuje inne zjawisko fizyczne i wyrażana jest w niemożliwych do porównania jednostkach oraz w zróżnicowanym zakresie wartości, powoduje że wprowadzone przez model wagi nie będą zapewniały takiego efektu jak mogłyby. W konsekwencji konieczne jest przeskalowanie wszystkich predyktorów do jednego zakresu, zachowując jednak wzajemne relacje pomiędzy poszczególnymi wartościami. Jednym z możliwych do użycia skalowań jest skalowanie min-max. Polega ona na odnalezieniu ekstremalnych wartości danej cechy zawartych w zbiorze i zastosowanie wzoru:

$$x'_i = \frac{x_i - \min x}{\max x - \min x}$$

Wartościom ekstremalnym przypisane zostaną: 0 dla minimum i 1 dla maksimum, relacja pomiędzy wartościami pośrednimi zostaje zachowana. Inną metodą jest standaryzacja. Jest to metoda powszechna w klasycznej statystyce, pozwala na określenie o ile od-

chyleń standardowych wartość odstaje od średniej.

$$x'_i = \frac{x_i - \bar{x}}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}}$$

W odróżnieniu od metody min-max, standaryzacja nie posiada z góry założonego przedziału, jednak jej zaletą jest możliwość łatwego rozpoznania wartości odstających.

Korelacja

Istnieją matematyczne miary współzależności dwóch zmiennych. Dzięki nim można określić wzajemne powiązania w danym zbiorze danych. Umożliwiają one wczesne wykrycie obiecujących zmiennych skorelowanych z wartością, którą chcemy modelować bądź też par cech, które wzajemnie wnoszą tę samą informację. W drugim przypadku możemy rozważyć pozostawienie jedynie reprezentanta grupy powiązanych ze sobą zmiennych, jako że zawsze preferowane są modele najprostsze. Najczęściej używanym współczynnikiem korelacji jest współczynnik korelacji liniowej Pearsona. Współczynnik ten jest ilorazem kowariancji i iloczynu odchyleń standardowych zmiennych. Pozwala on na wykrycie zależności liniowych pary zmiennych. Estymator tego współczynnika definiuje się w nastę-pujący sposób:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x}) - (y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

x_i i y_i to zmienne, dla których współczynnik jest liczony, natomiast $\bar{x}$ i $\bar{y}$ to średnie z tych wartości. Estymator zwraca wartość z przedziału $< -1, 1 >$ gdzie 0 oznacza brak korelacji. Nie ma jednej słusznej definicji mocnej i słabej korelacji – w niektórych dziedzinach wartość bezwzględna 0,3 będzie uznawana za istotną, a w przypadku rozpatrywania zjawisk fizycznych 0,9 może być za mała. W rozpatrywanym zagadnieniu nie ma potrzeby przeprowadzania testów korelacji wybranych cech. Doktorant posiada dużą wiedzę domenową, która pozwoliła wyłączyć z analizy pewne korelujące ze sobą parametry (opisano w Rodziale 15.2) a jak napisano powyżej, uzyskanie wartości estymatora nie pozwalają jednoznacznie stwierdzić, czy dana cecha może zostać pominięta. Jeśli w przyszłości podjęto by działania na rzecz rozbudowania programu o kolejne parametry, to wskazane byłoby przeprowadzenie testów kore-

lacji. Współczynnik Pearsona bazuje na odległości w przestrzeni pomiędzy pewną wartością a średnią. Średnia arytmetyczna jest czuła na wartości odstające oraz błędy grube. Współczynnik dziedziczy tę cechę i w przypadku dużej liczności tzw. outlierów w zbiorze będzie on zbiegał do zera. Jak napisano w tym rozdziale, zważywszy na sposób pozyskania danych dla rozpatrywanego zadania optymalizacyjnego, zagrożenie wystąpienia zestawu danych niskiej jakości jest mało prawdopodobne.

Podział danych

W przypadku uczenia nadzorowanego konieczne jest podzielenie zbioru danych (próby wejściowej) na części.

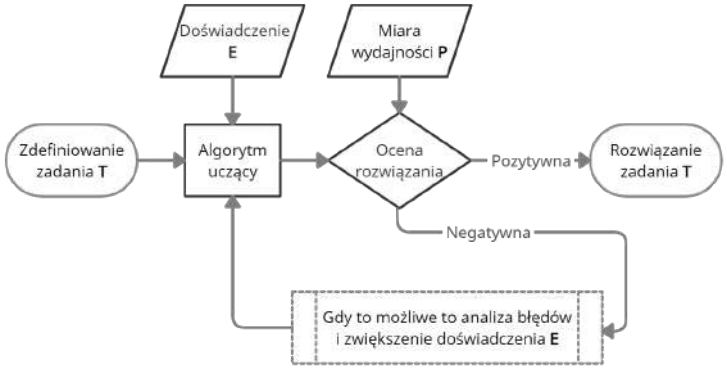
- treningowy – służy do utworzenia i dopasowania modelu. Liczność tej próby powinna być pomiędzy 40% a 70% próby wejściowej;
- walidacyjny – dalsza kalibracja modelu oraz monitorowanie jakości podczas jego budowy (20% - 30% próby wejściowej);
- testowy – ocena i porównanie modeli na niezależnym zbiorze, niewykorzystywanym w procesie tworzenia modelu (20% - 30% próby wejściowej).

Możliwe jest zrezygnowanie ze zbioru walidacyjnego ze względu na to, że funkcje zbioru walidacyjnego pokrywają się w pewnym stopniu z poszczególnymi funkcjami zbioru testowego i treningowego. Zastosowanie podziału na część treningową i testową jest konieczne w celu zbudowania poprawnego modelu. Podziału zbioru dokonuje się najczęściej poprzez proste losowanie.

15.4.2 Predykcja charakterystyki naprężeń

By rozwiązać zagadnienie parametrów rozruchu turbiny Doktorant postanowił wykorzystać zagadnienia sztucznej inteligencji - uczenia maszynowego. Uczenie maszynowe to dziedzina nauki, która umożliwia budowanie programu komputerowego na danych bez konieczności ich jawnego programowania. O uczeniu programu komputerowego można mówić gdy na podstawie doświadczenia E w odniesieniu do zadania T i pewnej miary wydajności P, jego wydajność

(zmierzona poprzez miarę P) wobec zadania T wzrasta wraz z nabywaniem doświadczenia E [12].



Rys. 15.4: Schemat ideowy uczenia maszynowego.

Uczenie maszynowe dzieli się na kilka typów [13]:

- Uczenie nadzorowane - używany zbiór danych został wstępnie oznaczony i sklasyfikowany przez użytkowników, aby umożliwić algorytmowi sprawdzenie dokładności jego działania;
- Uczenie nienadzorowane - używany jest surowy zbiór danych, a algorytm identyfikuje wzorce i zależności w danych bez pomocy użytkowników.
- Uczenie częściowo nadzorowane - zbiór danych zawiera ustrukturyzowane i nieustrukturyzowane dane, które prowadzą algorytm w drodze do samodzielnego wyciągania wniosków. Połączenie tych dwóch typów danych w jednym zbiorze umożliwia algorytmom uczenia maszynowego uczenie się dopasowania nieoznakowanych danych.
- Uczenie przez wzmocnienie - zbiorze danych zastosowano system "nagrody/kara", dzięki któremu algorytm otrzymuje informację zwrotną, aby uczyć się na podstawie własnych doświadczeń metodą prób i błędów.
- Głębokie uczenie (deep learning) - automatyczne uczenie się na podstawie zbiorów danych bez wprowadzania reguł lub

wiedzy człowieka. Wymaga to ogromnych ilości surowych danych do przetwarzania, a im więcej danych się otrzymuje, tym bardziej poprawia się model predykcyjny.

Do prac nad modelem predykcji najlepiej sprawdzi się uczenie nadzorowane dzięki danym dobrej jakości i ich odpowiedniemu przygotowaniu (opisane w Sek. 15.4.1). Typowy algorytm nadzorowanego uczenia maszynowego składa się z (w przybliżeniu) trzech elementów:

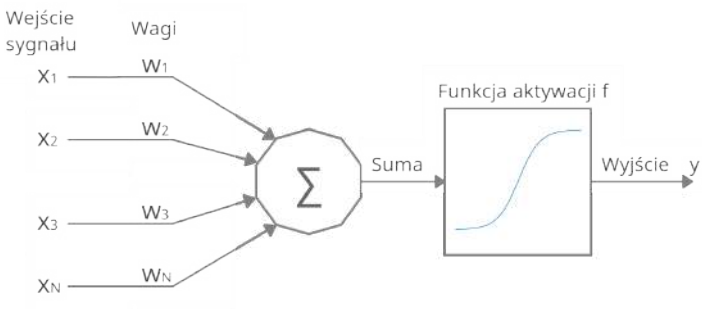
- Procesu decyzyjnego - formuła obliczeń lub innych kroków, który przyjmuje dane (doświadczenie E) i zwraca predykcję wartości objaśnianej.
- Funkcja błędu: Na podstawie zdefiniowanej metryki (miara wydajności P) następuje weryfikacja rozwiązania, jak dobre było przypuszczenie, porównując go ze znanymi przykładami.
- Proces aktualizacji lub optymalizacji: Algorytm analizuje i aktualizuje sposób, w jaki proces decyzyjny podejmuje ostateczną decyzję, tak aby następnym razem różnica pomiędzy wartościami przypuszczenia a przykładu była zminimalizowana.

Narzędzie matematyczne, które podlega uczeniu maszynowemu i znajdzie wykorzystanie w omawianym problemie ze względu na bardzo wysoką skuteczność w rozwiązywaniu problemów aproksymacji zjawisk nieliniowych to sztuczne sieci neuronowe (SSN bądź z ang. ANN – Artificial Neural Network). Sztuczne sieci neuronowe inspirowane są biologicznymi sieciami neuronów tworzących mózg. SSN opiera się na zbiorze połączonych jednostek lub węzłów zwanych sztucznymi neuronami. Każde połączenie, podobnie jak synapsy w biologicznym mózgu, może przekazywać swój wyjściowy sygnał do innych neuronów.

Sztuczny neuron odbiera sygnał x, a następnie przetwarza i może przesyłać sygnał podłączonym do niego neuronom.

Pojedynczy neuron – perceptron, przyjmuje N sygnałów wejściowych które mogą być liczbą rzeczyw

$$o = f(u) = f(\sum_{i=1}^{N} w_i x_i)$$



Rys. 15.5: Schemat sztucznego perceptronu.

Wyjście każdego neuronu jest obliczane przez pewną nieliniową (skokową) funkcję f z sumy u iloczynu jego wejść i wag. Funkcja ta nazywana jest też funkcją aktywacji.

Na ich podstawie wyznacza się sygnał wyjściowego o. Uczenie przedstawionego modelu perceptronu odbywa się poprzez znajdowanie odpowiedniego wektora wag poprzez aktualizację wag w kolejnych iteracjach. Poniżej przedstawiono aktualizację poprzez tzw. regułę perceptronową [12].

$$w_i^{t+1} = w_i^t + \Delta w_i^t$$
$$\Delta w_i^t = \eta(y - o^t)x_i$$

Aktualizacja wag może obywać się ze względu na to, że dysponujemy danymi x oraz y, gdzie y to rzeczywista wartość zmiennej objaśnianej Indeks t jest oznaczeniem numeru iteracji. η to parametr sterujący gwałtownością uczenia, który z reguły powinien być mały [12].

Inną regułą aktualizacji wag jest reguła delta (zwana również Adeline). Reguła delta bazuje na wyznaczeniu błędu średniokwadratowego SSE (ang. Sum Squared Error) danej iteracji i wyznaczeniu gradientu zmian czyli pierwszej pochodnej względem wag w. Poniżej przedstawiono implementację perceptronu z funkcją aktywacji sigmoid oraz regułą delta adjustacji wag.

Oprócz funkcji aktywacji sigmoid używa się wielu innych funkcji

```
import numpy as np
import pandas as pd

def sigmoid(x):
    return 1/(1+np.exp(-x))

def sigmoid_derivative(x):
    return x * (1 - x)

def network(training_inputs, training_outputs, n_iter):
    weights = 2 * np.random.random((3,1))-1
    for it in range(n_iter):
        input_layer = training_inputs
        outputs = sigmoid(np.dot(input_layer, weights))
        error = training_outputs - outputs
        derivative = sigmoid_derivative(outputs)
        adjustments = error * derivative
        weights += np.dot(input_layer.T, adjustments)
    return outputs

np.random.seed(1)
outputs = network(training_inputs, training_outputs, 100)
```

Rys. 15.6: Implementacja perceptronu z funkcją aktywacji sigmoid i aktualizacją wag poprzez metodę spadku gradientu w języku Python z użyciem NumPy.

– *tanh,ReLU,maxout,ELU*, które oferują działanie sygnałowe, czyli implikują gwałtowną zmianę wartości poprzez niewielką zmianę argumentu.

Wielowarstwowe sztuczne sieci neuronowe to zagadnienia matematyczne, które składają się z wielu zestawów neuronów nazywanych warstwami, gdzie pierwsza nazywana jest wejściową i służy do wprowadzenia danych. Kolejne warstwy nazywane są ukrytymi. Każda z warstw połączona jest tylko i wyłącznie z kolejną warstwą w strukturze sieci. Ukryte warstwy nie przyjmują danych spoza sieci, jedynie od swoich poprzedników. Sygnał opuszczający warstwę ukrytą nie generuje ostatecznego sygnału wyjściowego. Odpowiedzialna za to jest warstwa wyjściowa. Każda z warstw może posiadać wiele neuronów, które mogą służyć jako wejścia dla tych samych neuronów w kolejnych warstwach.

Wielowarstwowe sieci neuronowe pozwalają na rozwiązanie zadań nieliniowych.

Poprawność działania sieci neuronowej z nauczaniem nadzorowanym określa się bliźniaczo jak dla każdego przypadku aproksymacji funkcją danych (zadanie regresji) – poprzez wyznaczenie jednej lub kilku z miar:

- kwadratu korelacji R^2;
- błędu średniokwadratowego*RMSE*;
- średnioprocentowego błędu bezwzględnego *MAPE*;
- średniego błędu bezwzględnego *MAE*.

15.4.3 Optymalizacja

Optymalizacja jest procesem wyznaczania najlepszego rozwiązania według określonego kryterium. Mnogość algorytmów optymalizacji skłania do przyjrzenia się kilku najszerzej stosowanym. Algorytm Wspinaczki (ang. Hill Climbing) [14] to technika optymalizacji matematycznej należąca do rodziny przeszukiwania lokalnego. Jest to algorytm iteracyjny, który rozpoczyna pracę w dowolnym punkcie, a następnie próbuje znaleźć lepsze rozwiązanie funkcji f(x), gdzie x jest wektorem wartości ciągłych lub dyskretnych. Poszuki-wane jest lokalne minimum funkcji f(x), poprzez wprowadzanie przyrostowych zmian w optymalizowanym rozwiązaniu. Jeśli zmiana ta przyniesie lepsze rozwiązanie, w nowym rozwiązaniu wprowa-dza się kolejną przyrostową zmianę, aż do momentu, gdy nie będzie można znaleźć zmiany, która poprawiłaby wartość f(x). Wtedy mówi się, że x jest "lokalnie optymalny". Algorytm wspinaczki niekoniecznie znajdzie maksimum globalne, ale może zbiegać do maksimum lokalnego. Problem ten nie występuje, jeśli heurystyka jest wypukła. Ponieważ jednak wiele funkcji nie jest wypukłych, algorytm może często nie osiągać globalnego maksimum. Problem ten starają się rozwiązać inne algorytmy przeszukiwania lokalnego, takie jak Stochastyczny Algorytm Wspinaczki (ang. Stochastic Hill Climbing), Spacer Losowy (ang. Random Walk) czy Symulowane Wyżarzanie (ang. Simulated Annealing). Poniżej przedstawiono przykładową implementację Algorytmu Wspinaczki w języku Python z użyciem

biblioteki NumPy.

```
def hill_climb(membership, solution):
    counter = 0
    while True:
        counter += 1
        neighbours = []
        for num_index in range(len(solution)):
            first_part = solution[0:num_index]
            change = solution[num_index:num_index+1]
            last_part = solution[num_index + 1:]
            if change[0] != membership[1]:
                change_add = change + 1
                inc = np.concatenate([first_part,change_add,last_part], axis = 0)
                neighbours.append(inc)
            if change[0] != membership[0]:
                change_sub = change - 1
                dec = np.concatenate([first_part,change_sub,last_part], axis = 0)
                neighbours.append(dec)
        actual = fitness_function(solution)
        best = actual

        for neighbour in neighbours:
            new_cost = fitness_function(neighbour)
            if new_cost < best:
                best = new_cost
                solution = neighbour
        if best == actual: break
    return solution
```

Rys. 15.7: Algorytm Wspinaczki w języku Python z użyciem NumPy.

Stochastyczny Algorytm Wspinaczki jest ulepszoną wersją Algorytmu Wspinaczki. Podczas, gdy algorytm w wersji podstawowej zawsze wybiera najbardziej stromą część funkcji, wersja stochastyczna wybiera losowo jedno wzniesienie funkcji, a prawdopodobieństwo wyboru może się zmieniać w zależności od stromości funkcji w tym punkcie.

Kolejnym wartym uwagi algorytmem jest Symulowane Wyżarzanie (ang. Simulated Annealing) [15]. Jest to probabilistyczna technika wyznaczania przybliżenia globalnego optimum dla danej funkcji. W szczególności jest to metaheurystyka służąca do aproksymacji globalnej optymalizacji w dużej przestrzeni przeszukiwania problemu optymalizacyjnego. W przypadku problemów, w których znalezienie przybliżonego optimum globalnego jest ważniejsze niż znalezienie dokładnego optimum lokalnego w określonym czasie, Symulowane Wyżarzanie może sprawdzić się bardzo dobrze.

Symulowane Wyżarzanie jest oparte na praktykach metalurgicznych, w których materiał jest podgrzewany do wysokiej temperatury, a następnie chłodzony. W wysokich temperaturach atomy mogą się nieprzewidywalnie przemieszczać, często eliminując zanieczyszczenia podczas chłodzenia materiału do postaci czystego kryształu. Można to odtworzyć za pomocą algorytmu optymalizacji Symulowanego Wyżarzania, którego „stan energetyczny" odpowiada bieżącemu rozwiązaniu.

W algorytmie definiuje się temperaturę początkową, często równą 1, oraz temperaturę minimalną, rzędu 10^{-4}. Bieżąca temperatura jest mnożona przez pewien ułamek α i w ten sposób zmniejszana, aż osiągnie temperaturę minimalną. Dla każdej różnej wartości temperatury uruchamiamy procedurę optymalizacji rdzenia stałą liczbę razy. Procedura optymalizacyjna polega na znalezieniu sąsiedniego rozwiązania i zaakceptowaniu go z prawdopodobieństwem $e^{f(c)-f(n)}$, gdzie c jest rozwiązaniem bieżącym, a n jest rozwiązaniem sąsiednim. Sąsiednie rozwiązanie jest znajdowane przez zastosowanie niewielkiego zakłócenia w bieżącym rozwiązaniu. Ta losowość jest przydatna, aby uniknąć częstego problemu heurystyk optymalizacyjnych, jaką jest wpadanie w pułapkę lokalnych minimów. Akceptując potencjalnie mniej optymalne rozwiązanie niż to, które aktualnie posiadamy i przyjmując je z prawdopodobieństwem odwrotnym do wzrostu kosztu, algorytm ma większe szanse na zbliżenie się do globalnego optimum. Projektowanie funkcji sąsie-dztwa jest dość skomplikowane i musi być przeprowadzane indywidualnie dla każdego przypadku. Przykładową implementację Symulowanego Wyżarzania w języku Python zamieszczono poniżej.

Wydajnym i skutecznym algorytmem optymalizacji są algorytmy genetyczne [16] (ang. Genetic Algorithms). Algorytm genetyczny to heurystyka wyszukiwania inspirowana teorią ewolucji naturalnej Karola Darwina. Algorytm ten odzwierciedla proces doboru naturalnego, w którym najzdolniejsze osobniki są wybierane do reprodukcji w celu wytworzenia potomstwa w następnym pokoleniu. Proces selekcji naturalnej rozpoczyna się od wyboru najzdolniejszych osobników z danej populacji. Dają one potomstwo, które dziedziczy cechy rodziców i będzie dołączone do następnego pokolenia.

```python
def simulated_annealing(sol_len = 12, domain = [0,9], temperature = 1, alpha = 0.99):
    solution = np.random.randint(domain[1], size = sol_len)
    step = 1
    count = 0
    while temperature > 10**-4:
        temp_solution = solution.copy()
        index_to_change = np.random.randint(sol_len)
        change = np.random.choice([-step, step])

        if temp_solution[index_to_change] == domain[0]:
            temp_solution[index_to_change] += step
        elif temp_solution[index_to_change] == domain[1]:
            temp_solution[index_to_change] -= step
        else:
            temp_solution[index_to_change] += change

        loss = np.random.uniform(0,1)
        cost = fitness_function(solution)
        temp_cost = fitness_function(temp_solution)
        probability = math.e ** ((-temp_cost - cost)/temperature)
        if (temp_cost < cost or loss < probability):
            solution = temp_solution
        temperature *= alpha
        count += 1
    return solution

simulated_annealing()
```

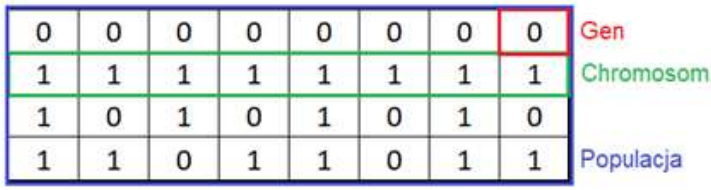
Rys. 15.8: Implementacja Symulowanego Wyżarzania w języku Python z użyciem NumPy.

Jeśli rodzice mają lepszą „kondycję", ich potomstwo będzie lepsze od rodziców i będzie miało większe szanse na przetrwanie. Proces ten powtarza się, a na końcu zostanie znalezione pokolenie z najzdolniejszymi osobnikami. Nie zawsze jednak z najlepszych rodziców rodzi się najlepsze potomstwo. Czasami skrzyżowanie słabszych jednostek powoduje powstanie zdecydowanie lepszego potomstwa, co jest uwzględnione w algorytmie. Algorytm genetyczny przebiega w pięciu fazach:

1. populacja początkowa;
2. funkcja fitness;
3. selekcja;
4. krzyżowanie;
5. mutacja.

Proces rozpoczyna się od zbioru osobników, który nazywamy populacją. Każdy osobnik jest pewnym rozwiązaniem problemu (w większości przypadków nieoptymalnym), który chcemy rozwiązać. Osob-

nik jest charakteryzowany przez zestaw parametrów (zmiennych) zwanych genami. Geny są łączone w ciąg, tworząc chromosom (rozwiązanie). W algorytmie genetycznym zestaw genów danego osobnika jest reprezentowany za pomocą łańcucha znaków, w postaci alfabetu. Zazwyczaj stosuje się wartości binarne (ciąg 1 i 0). Mówimy, że kodujemy geny w chromosomie.



Rys. 15.9: Schemat podziału zbioru na chromosomy.

Funkcja fitness określa, jak sprawny jest osobnik (zdolność osobnika do konkurowania z innymi osobnikami). Każdemu osobnikowi przypisuje ona wynik kondycji, a prawdopodobieństwo, że dany osobnik zostanie wybrany do reprodukcji, zależy od jego wyniku tejże kondycji.

Ideą fazy selekcji jest wybór najzdolniejszych osobników i przekazanie przez nich swoich genów następnemu pokoleniu. Dwie pary osobników (rodzice) są wybierane na podstawie ich wyników kondycji. Osobniki o wysokiej kondycji mają większe szanse na bycie wybranymi do reprodukcji.

Krzyżowanie jest najważniejszą fazą w algorytmie genetycznym. Dla każdej pary rodziców, którzy mają zostać połączeni, wybierany jest losowo punkt krzyżowania spośród wszystkich genów.

W niektórych nowo powstałych potomkach niektóre z ich genów mogą być poddane mutacji z małym prawdopodobieństwem losowym. Oznacza to, że niektóre bity w ciągu bitowym mogą zostać odwrócone. Mutacja ma na celu utrzymanie różnorodności w populacji i zapobieganie przedwczesnej konwergencji. Algorytm kończy działanie, jeśli populacja jest zbieżna (nie wytwarza potomstwa, które znacząco różni się od poprzedniego pokolenia). Mówi się wtedy, że algorytm genetyczny dostarczył zestaw rozwiązań na-

szego problemu. Przykładową implementację Algorytmu Genetycznego w języku Python zamieszczono poniżej.

```python
def gene_mutation(domain, step, solution):
    index_to_change = np.random.randint(len(solution))
    if np.random.uniform(0,1) > 0.5:
        if solution[index_to_change] + step <= domain[1]:
            solution[index_to_change] += step
    else:
        if solution[index_to_change] - step >= domain[0]:
            solution[index_to_change] -= step
    return solution

def genes_crossover(domain, solut_1, solut_2):
    dividing_index = np.random.randint(len(solut_1))
    solut_1_new = solut_1[:dividing_index] + solut_2[dividing_index:]
    return solut_1_new
```

```python
def genetic(domain, fitness_function, population_size = 50, step = 1,
            mutation_prob = 0.3, elitism_fract = 0.3, generations_num = 100):
    sol_len = 12
    population = [np.random.randint(domain[1]+1, size = sol_len) for _ in range(population_size)]
    elitism_num = int(elitism_fract * population_size)
    population_costs = [fitness_function(solution) for solution in population]
    np_arr = np.array([population, population_costs], dtype = "object").T

    for _ in tqdm(range(generations_num)):

        sort_mask = np_arr[:,1].argsort()
        np_arr = np_arr[sort_mask][:elitism_num]
        new_population = []

        while np_arr.shape[0] + len(new_population) < population_size:
            if np.random.uniform(0,1) < mutation_prob:
                m = np.random.randint(elitism_num)
                to_mutate = np_arr[m,0]
                mutated = gene_mutation(domain, step, to_mutate)
                mutated = hill_climb(domain, np.array(mutated))
                new_population.append(mutated)
            else:
                to_cross_1, to_cross_2 = np.random.randint(0, elitism_num, size = 2)
                crossed = genes_crossover(domain, list(np_arr[to_cross_1,0]), list(np_arr[to_cross_2,0]))
                crossed = hill_climb(domain, np.array(crossed))
                new_population.append(crossed)

        new_costs = [fitness_function(solution) for solution in new_population]
        new_np_arr = np.array([new_population, new_costs], dtype = "object").T
        np_arr = np.concatenate([np_arr, new_np_arr])

    sort_mask = np_arr[:,1].argsort()
    np_arr = np_arr[sort_mask]
    return np_arr
```

Rys. 15.10: Przykładowa implementacja Algorytmu Genetycznego w języku Python z użyciem NumPy.

Ponieważ algorytm genetyczny jest algorytmem wydajnym i skutecznym, zdecydowano, że jest on najlepszym wyborem do przeprowadzenia optymalizacji parametrów rozruchu turbiny. Proces optymalizacji powinien minimalizować różnicę pomiędzy aktualną wartością naprężeń w krytycznej lokalizajcji podczas rozruchu, wynikającej z odpowiedzi modelu sieci neuronowych, a naprężeniami dopuszczalnymi producenta – co opisuje wskaźnik q przedstawiony na początku Sekcji. Jednocześnie należy uwzględnić

pewien margines bezpieczeństwa pomiędzy tymi wartościami oraz zadeklarować, że różnica musi być większa od 0 (ograniczenie wskaźnika q na przykład do 1 pozwala zachować margines 1 MPa). Wynika z tego, że algorytm optymalizacyjny może zostać uruchomiony dopiero po skonstruowaniu modelu sieci neuronowych i musi działać bezpośrednio na jego odpowiedziach.

15.5 Wnioski

Uczenie maszynowe jest jednym z najbardziej istotnych elementów kolejnej rewolucji przemysłowej i towarzyszącej jej analityki. Koncentruje się na wykorzystywaniu dużych zestawów danych, wykorzystując metody statystyczne by identyfikować wzorce, iteracyjnie szkoli się modele do klasyfikowania lub predykcji w ramach projektów eksploracji danych różnych zjawisk fizycznych.

Koncepcja wykorzystania przedstawionych w niniejszej pracy rozwiązań matematycznych zakłada stworzenie sieci neuronowej, która symulowałaby zjawiska fizyczne zachodzące w rozpatrywanej turbinie parowej na podstawie zadanych cech wejściowych. Opracowany model jest niezbędny do poprawnego działania algorytmu genetycznego. Algorytm ten po procesie uczenia byłby w stanie zwracać optymalne krzywe rozruchowe dla aktualnego stanu począ-tkowego.

Korzyści płynące z zastosowania algorytmu optymalizacyjnego w porównaniu do istniejących systemów spodziewane są być największe właśnie w stanach pośrednich pomiędzy projektowanymi krzywymi rozruchowymi. Krzywe rozruchowe określają bowiem warunki przeprowadzenia optymalnego pod względem czasu rozruchu dla konkretnego stanu początkowego (np. 8h od odstawienia dla stanu gorącego, czy 48h od odstawienia dla stanu gorącego). Natomiast w rzeczywistości rozruchy są przeprowadzane ze zróżnicowanych stanów, dla których to algorytm optymalizacyjny będzie mógł określić inną wartość gradientów, które to doprowadzą do rozruchu w czasie szybszym, przy jednoczesnym zachowaniu żywotności.

Zaletą przedstawionej koncepcji optymalizacyjnej jest to, że algorytm optymalizacyjny jest odrębnym systemem od istniejącego systemu ograniczeń termicznych turbiny. Dzięki algorytmowi optyma-

lizacyjnemu operator mógłby otrzymać optymalne krzywe naboru obrotów/ mocy/ temperatury na obiekcie, bezpośrednio przed planowanym rozruchem dla rzeczywistych wartości początkowych, aby możliwie skrócić czas rozruchu, jednak turbina dalej chroniona jest przez istniejący i zwalidowany wieloletnią eksploatacją system ograniczeń termicznych, który to byłby systemem nadrzędnym do algorytmu optymalizacyjnego. W wyniku tego ryzyko podczas testów wdrożeniowych proponowanego rozwiązania ograniczone byłoby do minimum.

Kolejną zaletą proponowanej koncepcji jest fakt, że po dopracowaniu kodu algorytmu genetycznego stałby się on uniwersalny dla każdej innej konstrukcji turbiny parowej. Konieczna do zmiany byłaby sieć neuronowa, gdyż odpowiedź układu (wartość naprężeń) byłaby inna dla innej konstrukcji. Niemniej jednak należy zauważyć, że proces generacji zestawów danych z modelu MES oraz algorytmy uczenia sieci neuronowej również pozostałyby niezmienne. Z tego powodu stworzenie systemu dla innej turbiny, oprócz budowy modelu MES byłoby zautomatyzowane.

15.5.1 Podziękowania

Praca została wykonana w ramach III edycji programu „Doktorat Wdrożeniowy" ustanowionego komunikatem Ministra Nauki i Szkolnictwa Wyższego.

Bibliografia

[1] K. Dominiczak, R. Rządkowski, W. Radulski, and R. Szczepanik. *Sieci neuronowe w systemie ograniczeń termicznych wirnika turbiny parowej.* ITE - PIB, 2015.

[2] K. Dominiczak. *Mechanizmy zużycia i żywotność turbin parowych. Wybrane zagadnienia inżynierii mechanicznej.* 8th International Conference of Steam Turbine Users, 2013. praca zbiorowa pod redakcją Magdaleny Mieloszyk i Tomasza Ochrymiuka.

[3] K. Dominiczak, M. Drosińka-Komor, R. Rządkowski, and J. Głuch. Optimisation of turbine shaft heating process under steam run-up conditions. *Archives of thermodynamics*, 41(4), 2020.

[4] J. Przytulski. Mechanizmy zużycia i żywotność turbin parowych. In praca zbiorowa pod redakcją Magdaleny Mieloszyk i Tomasza Ochrymiuka, editor, *Wybrane zagadnienia inżynierii mechanicznej*, pages 179–198. Wydawnictwo Instytutu Maszyn Przepływowych PAN, Gdańsk, 2020.

[5] J. Przytulski. Obliczenia zmęczenia niskocyklowego wirnika turbiny parowej metodą elementów skończonych pod kątem optymalizacji rozruchów. In *Selected problems in mechanical engineering.*

[6] Warszyński M Bukowski, L. and. *Niezawodność maszyn i urządzeń hutniczych.* OW FORWARD, 1994.

[7] T. Chmielniak. *Obciążenia cieplne turbin parowych.* WNT, 1990.

[8] A. Rusin, M. Banaszkiewicz, M. Lipka, H Łukowicz, and W. Radulski. *Continuous control and optimization of thermal stresses in the process of turbine start-up.* Wiedeń 26-29 Maj.

[9] C. Althoff. *Programista samouk – profesjonalny przewodnik do samodzielnej nauki kodowania.* Helion, 2018.

[10] Python - programming language. https://en.wikipedia.org/wiki/Python_(programming_language). Data dostępu 03.03.2022.

[11] Top 10 reasons why python is so popular with developers. https://www.upgrad.com/blog/reasons-why-python-popular-with-developers/. Data dostępu 03.03.2022.

[12] A. Géron. *Uczenie Maszynowe z użyciem Scikit-Learn i Tensor Flow.* Helion, 2020.

[13] M. Tamir. What is machine learning (ml. https://ischoolonline.berkeley.edu/blog/what-is-machine-learning/. Data dostępu 03.03.2022.

[14] P. Samui and D. Tien Bui. *UHandbook of probabilistics Models.* Butterworth-Heinemann, 2019.

[15] S. Marsland. *Machine Learning – An algorythmic Perspective.* Chapman and Hall, 2009.

[16] R. Hurbans. *Algorytmy sztucznej inteligencji.* Helion, 2021.