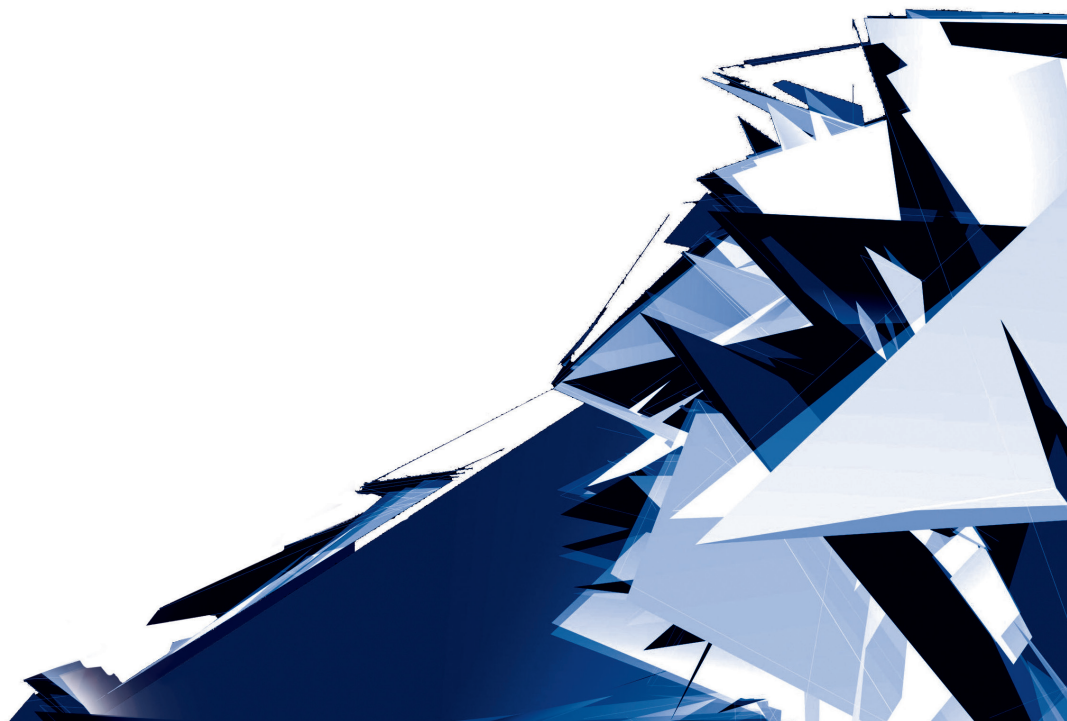


Technical Transactions

Czasopismo Techniczne

Volume 12

Year 2017 (114)



Chairman of the Cracow University of Technology Press Editorial Board
Przewodniczący Kolegium Redakcyjnego Wydawnictwa Politechniki Krakowskiej

Tadeusz Tatara

Editor-in-chief
Redaktor naczelny

Józef Gawlik
(jgawlik@mech.pk.edu.pl)

Scientific Council
Rada Naukowa

Jan Blachut – University of Liverpool (UK)
Wojciech Bonenberg – Poznan University of Technology (Poland)
Tadeusz Burczyński – Silesian University of Technology (Poland)
Massimo Corcione – Sapienza University of Rome (Italy)
Leszek Demkowicz – The University of Texas at Austin (USA)
Joseph El Hayek – University of Applied Sciences (Switzerland)
Ameen Farooq – Technical University of Atlanta (USA)
Zbigniew Florjańczyk – Warsaw University of Technology (Poland)
Marian Giżejowski – Warsaw University of Technology (Poland)
Sławomir Gzell – Warsaw University of Technology (Poland)
Allan N. Hayhurst – University of Cambridge (UK)
Maria Kuśnierowa – Slovak Academy of Sciences (Slovakia)
Krzysztof Magnucki – Poznan University of Technology (Poland)
Herbert Mang – Vienna University of Technology (Austria)
Arthur E. McGarity – Swarthmore College (USA)
Antonio Monestiroli – Polytechnic of Milan (Italy)
Marek Pabich – Lodz University of Technology (Poland)
Ivor Samuels – University of Birmingham (UK)
Mirosław J. Skibniewski – University of Maryland (USA)
Günter Wozny – Technical University in Berlin (Germany)
Roman Zarzycki – Lodz University of Technology (Poland)

Native Speakers

Weryfikacja językowa

Robin Gill
Justin Nnorom

Section Editor
Sekretarz Sekcji

Dorota Sapek
(dsapek@wydawnictwo.pk.edu.pl)

Editorial Compilation
Opracowanie redakcyjne

Aleksandra Urzędowska
(aurzedowska@pk.edu.pl)

Typesetting
Skład i łamanie

Małgorzata
Murat-Drożyńska

Design
Projekt graficzny

Michał Graffstein

Series Editors
Redaktorzy Serii

ARCHITECTURE AND URBAN PLANNING

Mateusz Gyurkovich
(mgyurkovich@pk.edu.pl)

CHEMISTRY

Radomir Jasiński
(radomir@chemia.pk.edu.pl)

CIVIL ENGINEERING

Marek Piekarczyk
(mpiekar@pk.edu.pl)

ELECTRICAL ENGINEERING

Piotr Drozdowski
(pdrozdow@usk.pk.edu.pl)

ENVIRONMENTAL ENGINEERING

Michał Zielina
(mziel@vistula.wis.pk.edu.pl)

**PHYSICS, MATHEMATICS
AND COMPUTER SCIENCES**

Włodzimierz Wójcik
(puwojcik@cyf-kr.edu.pl)

MECHANICS

Andrzej Sobczyk
(andrzej.sobczyk@mech.pk.edu.pl)

www.ejournals.eu/Czasopismo-Techniczne
www.technicaltransactions.com
www.czasopimotechniczne.pl

Contents

ARCHITECTURE AND URBAN PLANNING

Marcin Charciarek

Elementary architecture – between rationalism and poetic of form 5

Magdalena Jagiełło-Kowalczyk

Social, Economic, Spatial and Construction-related problems of Zwierzyniec.

The Four Seasons..... 15

Michał Krupa

The origins and history of the spatial development of Rabka-Zdrój during the medieval period – introduction to research 25

Michał Nessel

The relationship between cultural and geographic parameters in algorithmic design

– sample algorithm..... 35

Magdalena Rzeszotarska-Palka, Magdalena Czałczyńska-Podolska

Public engagement in the design process for landscape architecture objects..... 47

Elżbieta Waszczyszyn

The Children's University Hospital in Cracow. History and architecture..... 63

CHEMISTRY

Grzegorz Kurowski, Otmar Vogt, Jan Ogonowski

Paint-degrading microorganisms 81

Tomasz M. Majka, Agnieszka Leszczyńska, Krzysztof Pielichowski

Mechanical properties vs biotic degradation of polyamide/layered

silicates nanocomposites 93

CIVIL ENGINEERING

Dmytro Hladyshev, Roman Kinasz

Local action of loading on concrete cubes 113

Remigiusz Wojtal, Laurence R. Rilett

Calibration of microsimulation models of advance detection and warning systems

at signalized intersections..... 123

COMPUTER SCIENCES

Jerzy Orlof

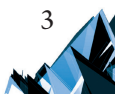
Point cloud based viewshed generation in AutoCad Civil 3D..... 143

ELECTRICAL ENGINEERING

Krzysztof Tomczyk, Marek Sieja

Assessment of propagation of modelling uncertainty by the procedures for determining

maximum dynamic errors..... 157



ENVIRONMENTAL ENGINEERING

Marta Bożym, Agnieszka Bok

Advantages and disadvantages of the solar drying of sewage sludge in Poland 171

Marek Kubala

The usefulness of cluster analysis in the analysis of data obtained in the monitoring of the water environment 181

Renata Sikorska-Bączek

The proposed design of a refrigeration system with an energy-efficient chiller with adjustable capacity for a fruit storage chamber 187

MECHANICAL ENGINEERING

Augustyn Lorenc, Małgorzata Kuźnar

The cost of risk evaluation in intermodal transport – variant analysis 197

Karol Miądlicki, Mateusz Saków

The use of machine vision to control the basic functions of a CNC machine tool using gestures 213

Agnieszka Żyra, Rafał Bogucki, Sebastian Skoczypiec

Austenitic steel surface integrity after EDM in different dielectric liquids 231

PHYSICS

Marcin Majka, Giacomo Gadda, Angelo Taibi, Piotr Zieliński

Effects of aortic valve diseases on pressure profiles in selected locations of the human arterial system 243

Marcin Charciarek (A-26@pk.edu.pl)

Institute of Architectural Design, Faculty of Architecture, Cracow University of Technology

ELEMENTARY ARCHITECTURE

– BETWEEN RATIONALISM AND POETIC OF FORM

ARCHITEKTURA ELEMENTARNA

– POMIEDZY RACJONALIZMEM A POETYKĄ FORMY

Abstract

The purpose of the article is to define the range of academic-didactic subject practiced in The Chair of Elementary Architecture of WAPK, known as Elementary Architecture – Between Rationalism and Poetic of Form. The wide academic range, as well as the aim of the research, allow only to present several aspects of defining contemporary architectural space. They aim to establish the origins, sources of architectural ideas in broad reference to structural rationalism and resulting from it poetic of form.

Keywords: elementary architecture, rationalism, poetic of form

Streszczenie

Artykuł jest próbą określenia zakresu tematyki naukowo-dydaktycznej uprawianej w Pracowni Architektury Elementarnej WA PK pod nazwą Architektura Elementarna – pomiędzy racjonalizmem a poetyką formy. Szeroki zakres badawczy i cel poszukiwań pozwalają jedynie na poruszenie niektórych aspektów definiowania współczesnych przestrzeni architektonicznej, których sensem jest ustanowienie źródła idei architektonicznej w szerokich odniesieniach do strukturalnego racjonalizmu i z niego wynikającej poetyki formy architektonicznej.

Słowa kluczowe: architektura elementarna, racjonalizm, poetyka formy

1. Rationalism and Intuition. Geometry

We are raised to respect the knowledge originating from reason. We are taught to choose things deriving from our experience, culture and knowledge. The rational mode of thinking, commonly known as the empiric view of the world around us, is limited to studying physical phenomena and facts that can be measured and proved. This intellectual attitude focuses on separate elements and isolated facts deriving from direct experience. Thinking is strictly limited to technical and practical processes that are best formed in theories and methodologies based on practice and experience. Friedrich Schelling wrote that if architecture was to be absolute art so it had to be identified with reason (without any indirectness). Architecture becomes beautiful only when it is independent from need. Architecture becomes beautiful when it is independent from itself [1, p. 272]. Architecture cannot represent the universe only through the form. It must represent the universe both in essence and in form. Rational ideal and perfection reveal before us the unlimited source of representation.

Kant took a different stand – The sources of knowledge are two fundamental components: intuition and thought. Our thinking is based on imagination. It means that it is based on our senses, so the only way to understand objects is to imagine them. Intellect alone is not able to perceive anything and senses cannot think. Only the combination of both can become knowledge. When we accept that thinking is the process of high-levelled imagining – only then we conclude that the whole knowledge is based on imagination [2].

This type of thinking looks for phenomena and experience that describe more than just the sum of parts. The main object of deliberation is not reality as itself, but the search for universal *Idea* for bigger *Content* that would glue everything together. This seems close to defining the problem by Etienne Louis-Boullée, Le Corbusier and Mies van der Rohe. For them, architecture as an expression was the metaphor of its own structure fulfilling the aim that was its representation. A building is understood as any figurative meaning incorporated in the physical structure through forms, essential as the whole and in its parts. On the other hand, interpretation following the poetic of architecture does not give us the certainty of judgment because it is based on the principle of similarity and distant references. This is the problem the poetic of form deals with.

Therefore, thinking of architecture must establish the elementary starting point for examining the essence of architecture that is rooted in any aesthetic rationalism. And the aesthetic is understood not only as the potential of geometry, but also as its poetic interpretation. Elementary means basic, original, inseparable (perhaps – modest) – the simplest in thinking and creating fundamental meanings to architecture – through its geometrical and material references. We discover it in specific designing ‘tactics’, but also in theories referring to meanings fundamental to the contemporary – from proto-modern ‘primitive hut’ of Abbé Laugier, stripped of ornaments ‘cave’ of Adolf Loos, Louis Kahn’s ‘pre-form’, Le Corbusier’s purity and Tadao Ando’s purity of forms.

Geometry is fundamental for building elementary meanings. Geometry – means the perfection of form and its remains in the history of architecture and culture. Geometry – enforcing a judgment that is reaching the explicit and clear relationship between the concept

and its practical realisation on paper. It seems that the creation of the consistent range of knowledge – from the rudiments of architecture to the advanced architectural formula might become a principle transmitted by credible and comprehensible language.

Elementary – also means designated by the consistent logical process of controlling in its essential rudiments the balance of the poles of architecture – general and detailed.

2. Limitation and Simplicity

It seems that the paradigm of limiting each structure and the form of this structure is still valid. Le Corbusier's purity, Mies's 'clarity of form', the reduction of contemporary minimalism and the rationalism of the Italian Novecento establish the continuity of architecture as the art of limitation. The idea of simplicity, developed at the beginning of the 20th century, did not impoverish life. It rather became the opportunity to create the path to fully experienced reality and to help us understand the process of defining architecture in its essential rudiments. As Gerrit Rietveld said: the biggest becomes accessible by the smallest [3, p. 33]. The similar idea of the process of the simplest perception of architecture is defined by John Pawson as the recurring ideal that divided various cultures searching for the path free from too many objects. Pawson writes: 'it is an elusive quality. Wide geographical and historical spread of various cultures interested in its attraction does not help define its essence' [4, p.8]. For the contemporary followers of reduction in architecture, the undeniable authority is a Benedictine architect Hans van der Laan. According to him, the experience of architecture is the strive for a clear division of a building into parts both in scale and detail. Hans van der Laan differed functional, expressive and monumental forms. They were adequate to three contexts of human life: nature, culture and liturgy. There should exist proper proportions and balance between them. He thought the great civilisations were mistaken to stress the expressive aspect of designing. This mistake resulted from decadence. Or it might have been an attempt to overcome the civilizational barrier.

Identical in this context is ideological, architectural 'patience' of Peter Zumthor. It defines vernacular effort, labour in which the aspect of industrial technology is clearly rejected by the architect. The model work – chapel in Wachendorf – seems the negation of the manifesto so as art does not need expression – so as the architect considered it harmful, because it disturbs and collides with true values of art, that are always totally material-formal. Exactly in this order.

3. Interpretations of the building material

The question of the rules of building has always been present in the theory of architecture. Raimund Abraham reminds that the essence of an architect's job is to understand how one stone is laid on the other. If architecture tries to organise space, so the structure organises matter to sustain the object's stability. That is why it is important which material is used and how it binds – they serve to prolong the life of the building. We accept that, maintaining that

the structural language is connected with architecture more than with other arts. The example of logic incorporated in triglyph, reflecting in the column-architrave system, is not only the feature characterising the principle of classical architecture. It also shows the complexity of architecture's technical aspect, as well as give us the idea of what is the limitation and solidity of structural type. Because of that, architecture seems not only the art of forms, but mainly the art of structured/ structural objects. Mortar between stones or shoring might define artistry, precision or their lack in the creator and his work. That is why we claim that the building of architectural work reflects the time and structure of the human mind.

If we accept the assumption that any architectural aesthetics are interested in the real object of aesthetic experience, we certainly accept the existence of material basis which creates this object. It is due to the material, that the unreal world of ideas and ideals, thoughts and words, concepts and drawings transfer art into the space of real values that become subject to final judgment and feeling. Similarly to paintings and sculptures, the transformation of concept into the real object gives architecture the meaning of art incorporated in the relation between shape and material. Therefore – idea and matter in architecture are inseparable – together with form and matter they are aesthetic unity. The sense of architectural art lies in the relations between the thoughtful configuration of ideas and the qualities of matter. Using the material in the simplest way, from general to detail, does not only search for the relations between the idea and matter of the work. It is rather a non-impulsive reach for the simple and communicative essence of architecture. According to this principle, architecture should be self-describing and introvert. Therefore, the language used by the creator of the building should strictly derive from logic, simplicity, geometry and figures – in other words – from the technique. Schelling wrote: ‘matter is absolute truth and reason is its essence’ [1, p. 525].

4. Models, Prototypes and Standards. ‘The Hut’ and ‘The Cave’

In the times when architecture was defined by the canon of ancient models, buildings created a consistent and clear field of knowledge completed by the collection of traditional meanings. For Leon Battista Alberti in *De re aedificatoria libri decem* or Andrea Palladio in *Studio elementare degli ordini di architettura* was obvious to use the template of details, proportions, orders derived from a Vitruvian past. Tables describing various ways of constructing and building the parts of colonnades were used to order materials in which they were to be modelled. Order meant the incorporation of the established principles in the reality of Renaissance.

Nowadays, nobody negates the fact that form and content are spread in architecture. But those who think that contemporary architecture lacks identity, and its postmodern embodiments change as quickly as in the kaleidoscope, are wrong. The multitude and richness of contemporary architecture allow us to repeat after Jorge Luis Borges, that, even though everybody whispers miserably that our century is incapable of spinning cohesive plots, it is because of these plots deciding about the shape of the whole, that our century might be superior to the past.

There are no doubts, that the history of architecture is full of examples proving genial inventiveness of its creators. On the other hand, most works of art are repetition. The sense of architectural idea is shown in the same way in the 'first', 'innovative' work, as well as in the sincere will of modelling followers. There are among them creators 'discovering' the new qualities of the model; or 'composers' transforming the established order – for them, art is the expression of general truths and solid fictions through the idealisation of form and matter. Architecture in the 21st century is a constant search for originality supported by geometrical models from the past, which define the present day of architecture. The relation idea-matter-architecture has not changed either. It still reveals before us limitless models for newly build objects.

Le Corbusier referred to 'poetic reaction' resulting from proper, wise and splendid play of blocks in the light. The architect saw the purity and precision of form and plan in prisms, pyramids, spheres and cylinders. The logic and peace of 'innate' in subconsciousness elementary blocks was the aim for total architecture revolution of the 20th century.

For Mies van der Rohe, the solution of the logic of architecture had to be simple, because it was supposed to establish the model/pattern of universal simplicity. He always claimed that there had to be such field where the answers – a priori – create one, closed, proper architectural system. The key to this analogy is not 'logical space' but rather Wittgenstein's principle of the search for clarity. The lack of ornament, together with precision, measures and proportions are the fundamental features of Rohe's theories and all his buildings, which allow us to find similarities to Wittgenstein's 'Logical-Philosophical Treatise'.

Louis Kahn finds 'order' between space and the technology of structure as inseparable connection and natural creating of rational beauty. To show the world, how the building is made, what are its parts, what it creates and how it 'works' – this is the harmony and the attempt to grasp the essence of architecture closed in expression '*Order is*'. The mode of ordering monolithic space through logic and the building method might have derived from the very beginning – the drawing as the 'frame' of the ideal relations of the whole to its parts and of parts to the whole. Kahn wrote: '[...] If we learned to draw just the way we built, from the bottom to the top, stopping the pencil to mark the places where the next stages of construction join together, the ornament would grow from love to expressing the method' [5, p.256][Ill. 1].

Since the times of Kahn, Le Corbusier, Mies van der Rohe, the understanding of concrete, brick or steel has become the ability of immediate transformation of thought into form – where architecture as the art of free, yet ordered choice became the aesthetic domain pointing to that what is ideal in things. For all those mentioned above, 'elementary' meant the return to the beginning of form and their rational interpretation – those ancient ones using the perfection of shape and geometry, and the modern ones describing the abstraction of the source of architectural form. Terms: 'archetype', 'pre-form', 'model', 'prototype', 'paradigm', 'canon', were not only synonyms. They were rather the transmutation of the formula of contemporary meaning of architecture.

We search for the sources for elementary architecture in the phenomenon of the line of the horizon which divides the shape created 'beneath the earth' or 'beneath the sky'. The theories of architecture tell us about *the cave* giving shelter and safety – carved space linked to the earth or rock hiding before the viewer the material logic of the building. They also tell us about



the hut – the manifested form (from bone, wood or steel) revealing before the viewer the logic of matter and resulting from it created shape. The archetype of ‘the cave’ created by carving space refers to the meaning of darkness, emptiness, shutting off – it is stability and connection with the base, it is the archetype of home. The archetype of ‘the hut’ built from frame creates the world opened to the outside, it is a model of a wooden ark – it is a permanent attempt to universalise space and its dematerialisation [ill.2].

Both archetypes are the source of the genetics of this kind of architecture that is based on cyclical, subconscious image of materiality and immateriality of architecture. This binary world is also the place for Adolf Loos vision – there are two things that belong to architecture: a monument and a tomb. The rest should be expelled from the world of art.

5. Basic Meanings of City and Place. City As the Work of Art

Nowadays, we are designing our cities without a common idea of a plan. Under the pretext of building spatial order, a town/city is created without any references to the surroundings, without analysing urban space as an act of creating a work of art. The idea of a contemporary city as a geometrical composition is not ideology – it is the search for the elementary order, understood as the search for an alternative way of living in the city filled with compact buildings of the centre and the chaos of the outskirts. An elementary city, hides within itself – an analogue city, a modern city, a traditional city. Even though such a city seems to hide an obvious paradox, yet it is an idea incorporating the composition of the place – defined by the simplicity of geometrical units and determined by the shape of elementary meanings of urban planning. The sense of such city should lie in creating recognisable places without unnecessary needs of future transformations.

Traditionally ‘closed’ building quarters, as well as ‘open’ forms of the city architecture established in the 20th century by modernists, might be helpful in creating the model urban composition. Describing important local urban space according to the building hierarchy, proper scale, views, close-ups or adequate division of City and Place should identify the balanced relation between the built space and empty space (nature).

The Modern Movement focused on the housing industry, by rejecting the ideas of monument (Camillo Sitte) and concentrating on the connection between the city and nature (Ebenezer Howard). The fundamental question became new elementary parts: rejection of the street as a place, a building quarter and a monument as the basic elements of the city, accepting nature as context, determining the capacity of units creating/building the city. The Postmodern Movement rebuilt the relation of the positive space where the basics became again a square, a quarter and a street. The Contemporary Movement, inheriting after both modernisms, describes the state of decentralisation and tries to interpret the values of functionalism and postmodernism.

Each urban system described their own rules called elementary parts, but did not describe the most relevant point – the hierarchy of importance between them through recognising the figure of the city as the compositional priority. The question of identity still remains a problem

– the form of places and buildings representing the identity of the dwellers. We should look for clues in the term the architecture of the city.

Composing a city cannot be done without basic figures describing its outline and spatial shape: a quarter, a block, a housing estate, the centre-square, a courtyard, a park/garden, an avenue, a basin, a gate, a dominant, a wall, the stairs, a mount, a monument... etc. It seems that the most important among them is still the idea of the square. The square, more than any other places, pictures the building of the city as *the work of art*. The city shows its intention of representation through the square. The city cannot give up on the square, unless it means degradation and the loss of identity. The motivation for composing might also be any space determined by geometry play that creates the order of the solution of urban planning. A composition that takes into account the ideality of form – known as the art of building cities [6,p.63].

6. Not a Summary – 26 Aphorisms by Snozzi instead [7]

- 1) When I think: a man, I think: exploited.
- 2) Architecture cannot be created by revolution, and revolution is not enough to create architecture.
- 3) Do not avoid responsibility: work with form. That is where you find man.
- 4) Architecture comes from real needs, but it overcomes them.
- 5) If you want to discover architecture, look at the ruins.
- 6) An aqueduct becomes alive when it stops delivering water.
- 7) There is nothing to think about, there is everything to discover.
- 8) When the graduates stop serving the architectural studios, School will make a huge step forward.
- 9) Before a design becomes the instrument of change, it is mainly the tool of perception.
- 10) Architecture is an emptiness that must be defined by you.
- 11) Architecture is measured by an eye and a step..... a liner to a geodesist.
- 12) A building always starts from the fundamentals.
- 13) Looking for elasticity? Build walls always from stone.
- 14) Diversity is the prelude to monotony, if you want to avoid it, always repeat the same element.
- 15) Nature cannot stand anything but the truth. But I think it was already said by Adolph Loos.
- 16) A real lawn reaches the roots of the Earth.
- 17) Each intervention assumes destruction, destroy with a conscience.
- 18) What a waste of energy, why to pay for air conditioning, heating, electricity....
A window would be enough.
- 19) Not so long ago all forms of human settlement were 'geological maps'.
- 20) When you design a path, a street, a courtyard, a house, a quarter – think about a city.
- 21) When all red lights disappear from the city – you will be close to the solution.
- 22) When you design a street or a car park, think, a human is driving.

- 23) Thanks to human effort a city includes the fire of volcanoes, the sand of the desert, the jungle and the steppe, flora and fauna ... nature in its whole beauty.
- 24) An alpinist is happy in the mountains. He knows there is a city beyond the horizon.
- 25) A sailor is happy at the sea. He knows there is a city beyond the horizon.
- 26) ... But the main thing that counts is light!

Translation Berta Chojnowska

References

- [1] Schielling J.W., *Fiozofia sztuki*, Warszawa 1983.
- [2] Ungers O.M., *Projektowanie i myślenie. Obrazami, metaforami i analogiami*, [after:] O.M. Ungers, *Morphologie: City Metaphors*, Köln 2011.
- [3] Miłobędzki M., *Ima Summis*, Autoportret, No. 4 (43), Kraków 2013.
- [4] Pawson J., *Minimum*, London 1999.
- [5] Jencks Ch., *Ruch nowoczesny w architekturze*, Warszawa 1987.
- [6] Monestiroli A., *Tryglif i metopa. Dziewięć wykładów o architekturze*, transl. U. Pytlowany, A. Porębska, Politechnika Krakowska, Kraków 2009.
- [7] Fonti A., Mameli M., *Luigi Snozzi. Un'autobiografia architettonica. Aforismi. 1973–1975*, Milano 2012.

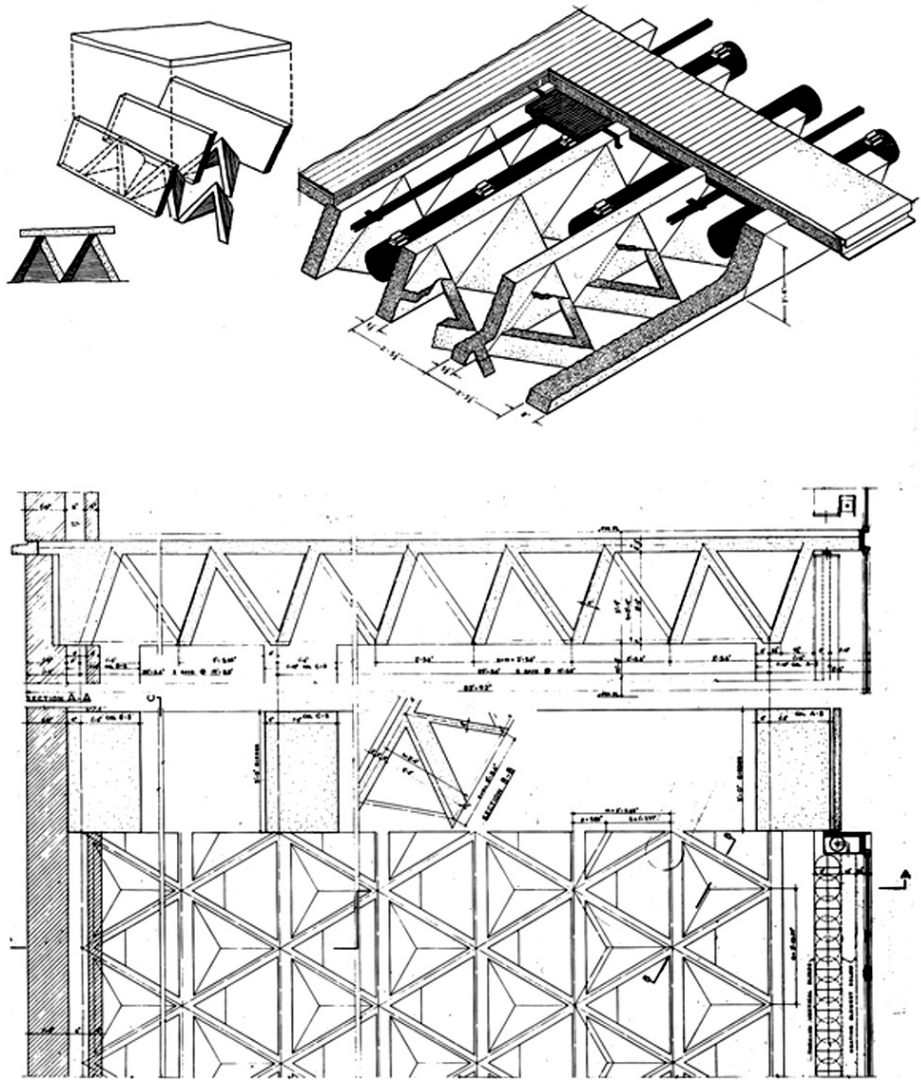


Fig. 1. L. Kahn, Scheme of reinforced concrete slab, Yale Art Gallery, New Heaven 1974

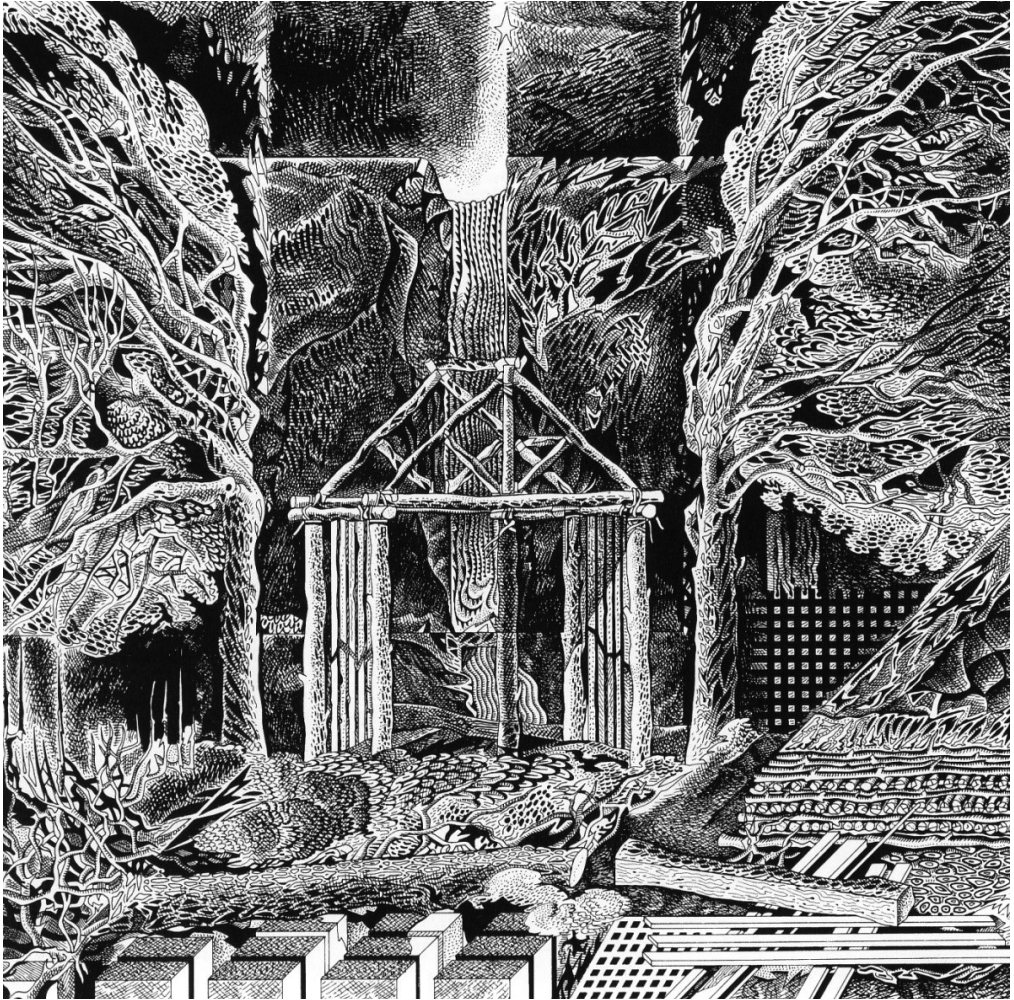


Fig. 2. F. Purini, *Trees Retreat after the Arrival of the Hut*, 1984

Magdalena Jagiełło-Kowalczyk (magdajagiellok@interia.pl)
Institute of Urban Design, Faculty of Architecture, Cracow University of Technology

SOCIAL, ECONOMIC, SPATIAL AND CONSTRUCTION-RELATED PROBLEMS OF ZWIERZYNIEC. THE FOUR SEASONS

PROBLEMY SPOŁECZNE, GOSPODARCZE, PRZESTRZENNO-INWESTYCYJNE
ZWIERZYŃCA. CZTERY PORY ROKU

Abstract

The publication presents the social, economic, spatial and construction-related problems that are being faced by the commune of Zwierzyniec. It discusses the effects of the cooperation of the commune with the Faculty of Architecture of the Cracow University of Technology. These effects, expressed in the form of diploma engineer's design theses of the students of the Faculty of Architecture, are to promote Zwierzyniec and its activation during all of the seasons of the year. One of the important, identified problems of the area of Roztocze in question is that all activity "freezes" there during the autumn and winter seasons.

Keywords: Social, economic, spatial, construction-related problems, activation

Streszczenie

Publikacja przedstawia problemy społeczne, gospodarcze i przestrzenno-inwestycyjne, z jakimi boryka się gmina Zwierzyniec. Omawia efekty współpracy gminy z Wydziałem Architektury Politechniki Krakowskiej. Efekty te, wyrażone w postaci projektów dyplomowych inżynierskich studentów Wydziału Architektury, mają służyć promocji Zwierzyńca i jego aktywizacji we wszystkich porach roku. Jednym bowiem z istotnych, zidentyfikowanych problemów omawianego terenu Roztocza jest jego „zamieranie” w sezonie jesienno-zimowym.

Słowa kluczowe: Problemy społeczne, gospodarcze, przestrzenne, inwestycyjne, aktywizacja

1. Introduction

The founding of Zwierzyniec is dated to the year 1593 [1]. It was then that a hunting bestiary was established at the site, surrounded by a 30-kilometre long fence, along with a manor made of larch. This gave the locality its name. At the beginning of the XIX century, the Zamoyski family relocated the Authority of the Zamoyski Family Fee Tail to Zwierzyniec. A complex of classicist halls that had been built for this purpose was to be its headquarters. This is also the time when the houses for the administrators of the fee tail were built. Foreign specialists from England, Italy and Germany were hired to work in the first agricultural equipment factory in Poland - built in the year 1804 – and the still functioning brewery from 1806. During the November Uprising Zwierzyniec was the site of the camping and movement of military forces, and the January Uprising has had one of its victorious battles in nearby Panasówka. In 1901 Józef Piłsudski, then a runaway from Petersburg, found help and shelter here. Thanks to the help of the fee tail's foresters, he made his way illegally across the border with Galicia on the Tanew River. During the turn of the XIX and XX centuries, Zwierzyniec became an attractive tourist resort. The fence of the hunting bestiary was completely dismantled in 1905. In September 1939, Zwierzyniec found itself along the combat trail of Stefan Rowecki and Stanisław Sosabowski, while during the occupation, it was the site of a prospering centre of the 9th infantry regiment of the Home Army Legions. The Nazi German prisoner of war camp for French captives was transformed into a transition camp for the pacified and forcefully relocated the population of the Zamojszczyzna. Ordinate Jan Zamoyski, along with his wife, involved themselves in the effort to save children from this camp. During the end of 1939, the supervision of the Zamoyski Family Fee Tail was taken over by the occupation authorities, with the Red Army entering this area in 1944. Even then, the 59,054-hectare fee tail was a large, well-organised manorial enterprise, bringing in quite a profit. On the 6th of September 1944, by decree of the PKWN regarding agricultural reform, the 350-year-old Zamoyski Family Fee Tail was dissolved. Zwierzyniec gained town rights in 1990. This charming, forest town, often called the Pearl of Roztocze [2] combines natural and cultural qualities. The four centuries of Zwierzyniec's history were, fortunately for the town, dominated by the Zamoyski Family Fee Tail. It is thanks to the Fee Tail that the cultural development of these lands had not destroyed, but rather underlined the natural qualities that were assigned for protection by creating the Roztocze National Park in 1974. The wonderfully wild green areas, both natural and manmade water layouts are matched with the material traces of the past. In this field, Zwierzyniec can pride itself in its small Baroque church, built on one of the islands of the lake in the years 1741–1747 by Teresa and Tomasz Zomoyski in thanks for the birth of their long-awaited son, the historical buildings of the Zamoyski Family Fee Tail, the XIX century brewery or the exceptional district of small timber houses called Borek, which was built during the period of the interbellum. The town, simply breathing history, flowing through the beautiful landscape of Roztocze, is, however, facing numerous problems, most of which have been identified, but still wait to be solved. Can the cooperation with the Faculty of Architecture of the Cracow University of Technology help, if only partially, to achieve this goal?

2. Identifying the main problems of the commune of Zwierzyniec

A series of main problems that affect the authorities and the inhabitants of the commune of Zwierzyniec has been identified [3]. The problems can be divided into three groups: problems that affect the social sphere, economic problems and spatial and construction-related problems (existing facilities, roads and technical infrastructure, public spaces, environmental and cultural problems).

In the **social sphere**, we can see the emergence of issues regarding the material wellbeing of inhabitants, shelter, security, healthcare, the life and care for disabled persons, personal development, access to culture, sports, recreation, social contacts, a feeling of identification with the town and the community, access to communication and information from the media and consumer needs, that require addressing and that decisions be made.

The problems associated with the material and residential wellbeing of inhabitants have been defined as follows: an insufficient amount of jobs for persons with a higher education, the low economic standing of a large part of the inhabitants, poor use of social economy instruments, the lack of the capability to provide shelter to persons and families in need and the occurrence of difficulties for the residents of multi-family residential buildings belonging to the Housing Cooperative that are rooted in the high energy demands of these buildings.

In regard to security, incidents of disrupting public peace and some felonies were identified, mainly during the tourist season. Excessive and disrupting transit traffic through the voivodship road that crosses the city, as well as on some of the crossings in Zwierzyniec, in addition to the noise and vibrations generated by the wide-track railway line, is a different kind of danger.

In terms of healthcare and the life and care for those in need, the necessity to carry out feeding programmes in preschools, schools and the cafeteria has been identified. Limited access to rehabilitation and the problem with the care for the rising number of disabled and senior persons that require specialist care is also making itself evident. This is associated with a larger number of persons who are unable to function in the job market.

The town has an insufficient offering in terms of spending one's free time, sports and recreation. The offering of the only cinema in Zwierzyniec has been found to be of low quality and scope due to ageing technical equipment. An unsatisfactory offering in terms of culture for both residents and tourists has been identified. This comes down to the conclusion that the potential of the locality in terms of creating a cultural, recreational and tourist offering is made use of in an insufficient manner. In the sphere of education, it has been determined that schools are inappropriately equipped with educational aids.

A series of problems exist in the sphere of social contacts, group ties and the feeling of identification with the town and the community. A low amount of involvement in local democracy mechanisms by the inhabitants is observable. Other elements that have been identified are: a low level of the integration of the residents of the community and passivity regarding participation in various cultural, economic and social events or initiatives, a low level of the residents' participation in the commune's management, a very small amount of persons willing to participate in voluntary work, the passivity of a part of those persons who

receive social benefits in undertaking efforts that would improve their material situation, a low level of physical activity among the residents.

Regarding the ease of contact and access to information, the lack of coverage by mobile phone networks and Internet broadband service providers in some areas is a large problem, as well as the lack of optical fibre networks. This causes difficulties in communication and those associated with the lack of access to information.

In terms of consumer needs, a limited availability of basic services and an insufficient offer in terms of small gastronomy and a lack of a marketplace organised to a modern standard.

The key sectors that influence the **economic development** of the region are: the timber industry (woodcutting, furniture production, palette production), the forestry industry, tourism, food processing (brewery) and agriculture.

These fields have problems regarding the small scale of production and the seasonal nature of woodcutting. The following have been observed: a low competitive position of the Forestry Services Plant, the unregulated legal status of the buildings of the brewery, underdeveloped, fragmented, low-volume agriculture, the seasonal nature of the demand for gastronomy-type service. Tourism in the region has problems that have been identified as a lack of categorised tourist facilities, a lack of leading tourist facilities, a tourist season that is too short, a narrow, seasonal offering, a low number of foreign tourists, the lack of an integrated promotion of the community's offering in terms of tourism.

There are serious problems with the employment market. These include: the low mobility of the residents that generates unemployment, the incompatibility of the qualifications of the unemployed to the needs of the job market, the low activity of a part of the unemployed in the search for and taking up of employment, insufficient professional qualifications of a part of the persons who provide services in tourism.

One obstacle in the economic development of the community is the low amount of available buildable areas, a fragmentation of the ownership structure, an insufficient number of property developers, very limited opportunities in terms of attracting a large investor and the poor effects of the efforts of non-government organisations in this field, the low activity of non-government organisations in securing outside funding, a low level of cooperation between non-government organisations and commune authorities and between non-government organisations themselves.

In terms of **existing facilities**, we can observe problems such as: the high energy demand of the residential buildings of the Housing Cooperative, the lack of protected social housing, the insufficient facility base and technical equipment of the Cultural Centre and the Public Library in Zwierzyniec, village clubhouse buildings that require modernisation/redevelopment, the poor technical condition of a part of the historical buildings.

The topic of **roads and technical infrastructure** is associated with: the lack of sewerage in a part of the locality, the necessity to construct a local waterworks network, the poor technical condition of some commune and powiat-grade roads, the partial lack of illumination, an insufficient amount of sidewalks, bike lanes and walkways, the necessity to modernise pedestrian and road bridges on commune roads, the necessity to build local roads in areas where the amount of buildings is rising, an insufficient amount of parking lots and parking spaces.

There are architectural barriers for disabled persons in **public spaces**, an insufficient number of public toilets is observable, as well as a poorly developed tourist, sports and recreation infrastructure (few swimming spots, the lack of a swimming pool, etc.), in addition to the imperfection of the state of the architectural and natural development, chaotic built-up areas that negatively affect the landscape.

There are several **environmental problems** that appear in the area of the community of Zwierzyniec: illegal waste dumping sites, secondary succession in wastelands, the overgrowing of unused farmland, the lack of or an insufficiently developed sewerage system in some parts of the community and an outdated sewerage system in a part of Zwierzyniec, the negative influence of the wide-rail railway line that generates noise and vibrations, coal-powered boiler rooms, a waterworks network that partially requires modernisation, the lack of the technical means to process organic waste, a low degree of the use of renewable energy sources, dangers arising from intense, heavy transit traffic that goes through the centre of Zwierzyniec.

Cultural problems associated with construction include: the disappearance of culture and local customs; an insufficiently broad offering of cultural events; the degradation of a part of the historical buildings; legal limitations regarding care for historical buildings and unsolved property issues.

3. Architecture – *The Four Seasons*

Six engineer's diploma design theses were developed as a part of an agreement between the commune of Zwierzyniec and the Faculty of Architecture of the Cracow University of Technology [4]. The students visited Zwierzyniec in order to inspect the site in person and perform urban design analyses. During these visits, they had the opportunity to get to know the town in depth and get a feel for its atmosphere. They met with representatives of the commune office and the local community. They have also been presented with the problems that the commune is struggling with and that it needs to face. The students pointed out the low degree of industrialisation and urbanisation of Zwierzyniec and its focus on agriculture and forestry. They remarked that the town is alive mostly during the summer season, which is when cultural events are organised [5, 15, 16] In terms of an offering for tourists, there are kayak trips, as well as the opportunity to experience the Eastern Poland Bike Trail which runs through here. During winter, however, the town becomes inactive. There is a lack of facilities that would activate the local community and strengthen the residents' feeling of identity, but that would also stimulate the inflow of tourists. Tourism can be seen as the source of the social and economic development of the region, which can improve the material wellbeing of its residents. Zwierzyniec owes its charm chiefly to its massive forestation and the beautiful natural landscape of the Roztocze National Park. However, many green areas are undeveloped and devastated. The students have performed circulation, waterway, greenery and historical building placement analyses. They have performed an evaluation of the assets of the town's fragments, drawn conclusions and developed design guidelines. In cooperation with the town authorities, it was decided that designs of service and residential buildings – that could

facilitate an improvement of the town's image, make it more attractive and active throughout all the seasons of the year – would be developed for recommended plots. The entire project has thus been called “The Four Seasons”.

The first design involves the construction of a summer amphitheatre for around 800 spectators and the redevelopment of its surroundings [6]. The site, with a total surface area of 1,07 ha, is located in the centre of Zwierzyniec, in the northern part of the water and palatial complex near Wachniewskiej Street and is adjacent to Lake II. It is a flat area free of buildings, in the southern part of which, near the water, there runs an embankment. The area is covered with low-lying littoral greenery. According to regulations in the Local Area Development Plan, all real estate development on this site requires the approval of the Voivodship Conservator of Historical Buildings [7]. This place is to become the town's main attraction, which would activate the local population and stimulate tourism. The concerts that would take place here during summer, as well as other music events, would have a proper and interesting place to be organised. The music would be heard from the spectator stands, but also from the opposite shore of the lake, or during a romantic boat ride on the water. The rich functional programme of the backrooms for the artists and spectators would also stimulate social contact in this area. A light, “winged” form was developed, which features a proposal of using a roof covering system made from ETFE air cushions. According to the author, the shaping of modern forms in a historical environment enriches the existing spatial and cultural context.

The topic of the second work is the building of a worker's hotel in Zwierzyniec, which also provides rooms for tourists to spend the night in [8]. The building is to be located on a site with a surface area of 2755 m² in the western part of the town near Zdrowotna Street. A project located at this site requires the approval of the Voivodship Conservator of Historical Monuments [7]. The design features a building composed of two intertwining, mutually contrasting shapes. The modernist form of the entry section has been contrasted with the classic, two-storey building covered with a gabled roof. The footprint surface area is 598 m², while the entire project includes a terrace that has been designed to the west, from which there is a view of the fields and forests of Roztocze. The main part of the hotel is covered by a timber, latticework cladding, which is offset from the wall and backlit, which is to provide the effect of lightness and the illusion of it being from a fairy tale, in reference to the charms of “Magical Roztocze” [9]. The hotel is adapted to being used by disabled persons. The technical solutions implemented in the building guarantee its energy efficiency, while “green” parking spaces on a comparably small plot provide the proper amount of biologically active area.

A harbour with accompanying infrastructure has been designed to bring order to the space near the western shore of the lake on the Wieprz River [10]. The plot on which the construction of a harbour has been planned is located on Rudka Lake in Zwierzyniec. Currently, it is the site of an earthen parking lot. During the summer, kayaks and water bikes are chaotically anchored near the shore of the lake. The design includes the construction of a new, functional structure with a regular shape. The building is oriented in parallel to the shoreline. The building's distinct quality are timber pillars which are displayed on the facade. They were meant as a reference to the tree trunks of the trees that grow in the Roztocze National Park. There is a multifunctional hall dedicated to events associated with this place in

the main part of the building. There is a room for lifeguards, as well as a number of backrooms and technical rooms nearby. On the ground floor, there is a hall dedicated to the storage, repair and conservation of sailing equipment. The building has a glazed, double-skin facade. Thus, the rooms have an unblocked view of the lake and the natural surroundings of Roztocze. The facade is free from load-bearing walls. This is because of the adoption of a post-and-beam structure based on concrete pillars.

The next subject is the development of a conceptual design of a covered swimming pool with a recreational section, gastronomy section, technical and backroom section and a parking lot [11]. The swimming pool has been placed in Zwierzyniec, near Armii Krajowej Street. The site is not large, which results in limitations regarding the introduction of a larger area of greenery. This is compensated by the park which is located on the opposite side of the street. The swimming pool is to be generally accessible and adapted to the needs of disabled persons. The massing of the building is minimalistic, with numerous glazed elements, the largest of which are located near the pools. The people who are going to use the pool will be able to not only enjoy their swim, but also to look at the park through the transparent glass walls. Each facade is decorated by a latticework structure made from stainless steel which has a natural-like form. The main inspirations in the design of the decorative cladding of the building were the park, the forests and the entire natural landscape of Roztocze. During the day, the sun will be able to shine through between the “branches” of the facade inside the building, creating fantastic compositions of light and shadow, while during the night, the light from the interior will be able to travel outside of the building, bringing to mind clearance in the forest. The colours of the building were kept in the currently fashionable tones of grey, white and black. The structural skeleton is composed of reinforced concrete elements. A roof composed of glued laminated timber has been designed above the swimming pool hall. The transparent walls have a post-and-beam structure, while the glass is fitted using spider-type fastenings. The interplay of the facades which resemble branches or trunks of trees, as well as nests made by birds, provides a feeling of lightness and spices up the simple shape of the building. In the opinion of the author of the design, the swimming pool, with its geometric form and facades inspired by the nature of Zwierzyniec, is going to elevate the touristic qualities of this picturesque region even more, activating the town in the summer, spring, autumn and winter.

The design of covered ice skating rink features the construction of a building that would activate the population of the town - which is the most active during the summer- in the winter [12]. The site, with a total surface area of 4,54 ha, located in the centre of the town of Zwierzyniec contains within its borders the palace and park complex of the Zamoyski Family Fee Tail Authority. The part chosen for the project constitutes an undeveloped plot in the direct vicinity of Zwierzyńczyk [13]. The Local Spatial Development Plan designates this site with the symbol A70-IK, and all real estate development associated with it should be approved by Voivodship Conservator of Historical Buildings [7]. The design is also meant to be a tourist magnet, which would be an element that could drive the development of the region. The answer to this challenge came in the form of a design of an all-year ice skating rink. Its massing is inscribed into the surrounding landscape and its facade opens up towards the Zwierzyńczyk. The interior intertwines with the exterior, architecture mixes with nature.



The user of the ice skating rink has a view of the water: as it forms waves in tune to the gusts of wind, cradled by greenery in the summer, and bound in ice in the winter. The massing, according to the author's concept, is to be in the shape of a chunk of ice. The building, apart from the ice skating rink hall and social and administrative rooms, also has a spectator area for around two hundred spectators. The structure of the building is based on a combination of reinforced concrete and glued laminated timber.

The last design is the multi-family residential building with services on the ground floor. [14] The site is located in Zwierzyniec in the vicinity of Jasna Street, on a narrow plot, the surroundings of which are made up of other multi-family buildings of a similar size. The access to all amenities is provided from the side of Jasna Street. According to the Local Area Development Plan [7], the site is assigned for the construction of a residential building. A four-storey multi-family residential building with services on the ground floor has been designed. The proposed services include a fitness club, which would be very welcome here, as well as a hairdressing salon and a coffee shop. The three remaining stories contain apartments, which are divided into two types: small, single level ones with a surface area of around 70 m² and larger, two-level ones with a surface area of around 130 m². The apartments are functional, and the building, in accordance with the principles of sustainable design, is to be energy efficient, which is going to be beneficial to the wallets of the residents. The form of the building, as well as its facades, were inspired by local timber architecture, as well as the neighbouring surroundings. In order to achieve a greater lightness of form, the massing was divided into smaller parts using glazed staircases. The colours black and white were used, while the roof was "rotated" in order to give the building a more modern, dynamic appearance.

4. CONCLUSION

The entire region of Roztocze is a true landscape jewel of Poland, one that has not been fully discovered yet. We can calm down here, experience nature and simply find rest. Zwierzyniec is a charming town that blends into the nature of this area. It awakens during spring, lives over the summer, falls asleep during autumn and freezes during winter. To a visitor who seeks peace, an escape from the hustle and bustle of the cities, from the problems and frantic pace of life, the town's stagnation and lethargy are oftentimes seen as its advantage. However, the people who visit this town for longer, for touristic reasons, quickly begin to search for entertainment and propositions of active recreation. In the summer, these requirements are seemingly met, however, the tourist offering can be made more attractive for this time as well, in order to adapt it to a larger number of consumers and, first and foremost, propagate it. Autumn, winter, even spring in Zwierzyniec await an awakening. The influx of tourists and the activation of the residents themselves would undoubtedly influence the economic development of the town. New jobs would be adapted to all of the year's seasons. An improvement in the living conditions of residents, as well as their identification with the image of an attractive tourist town, should also prevent the outflow of people in search of work. The authorities of Zwierzyniec, when identifying its problems, underlined the necessity

to activate the region. Cooperation with the Faculty of Architecture of the Cracow University of Technology is based on preparing an offer for the town under the common title “The Four Seasons”. Their construction would, without a doubt, lead to a revival of Zwierzyniec. For now, however, the students’ designs can serve as an element of the promotion of the town.

References

- [1] <http://www.zwierzyniec.info.pl> (access: 17.03.2017).
- [2] www.magicznyszwierzyniec.pl (access: 18.03.2017).
- [3] www.inwestycje.zwierzyniec.vipserv.org/strategia/prob.pdf (access: 18.03.2017).
- [4] Agreement no. A-0/46/2016 of 15th of December 2016 regarding the cooperation of the Town and Commune of Zwierzyniec, represented by Commune Mayor Jan Skiba, with the Thaddeus Kościuszko Cracow University of Technology, represented by the Dean of the Faculty of Architecture prof. Jacek Gyurkovich.
- [5] Summer Film Academy <http://laf.net.pl> (access: 18.03.2017) <http://www.kultura-zwierzyniec.pl/rockowisko-2016> (access: 18.03.2017).
- [6] Diploma design thesis author: Joanna Skoczeń, promotor: prof. Magdalena Jagiełło-Kowalczyk, auxiliary promotor: D.Sc. Ph.D. Bogdan Siedlecki, Krakow 2017
- [7] Local Area Spatial Development Plan, <http://www.zwierzyniec.info.pl> (access: 29.03.2017).
- [8] Diploma design thesis author: Marcin Musiałek, promotor: prof. Magdalena Jagiełło-Kowalczyk, auxiliary promotor: D.Sc. Ph.D. Jarosław Huebner, Kraków 2017.
- [9] <http://www.magiczneroztocze.pl> (access: 29.03.2017).
- [10] Diploma design thesis author: Monika Kraus, promotor: prof. Magdalena Jagiełło-Kowalczyk, auxiliary promotor: D.Sc. Ph.D. Bogdan Siedlecki, Krakow 2017.
- [11] Diploma design thesis author: Magdalena Sendor, promotor: prof. Magdalena Jagiełło-Kowalczyk, auxiliary promotor: D.Sc. Ph.D. Przemysław Markiewicz, Krakow 2017.
- [12] Diploma design thesis author: Konrad Piwowar, promotor: prof. Magdalena Jagiełło-Kowalczyk, auxiliary promotor: D.Sc. Ph.D. Bogdan Siedlecki, Krakow 2017.
- [13] Zwierzyniec was established as a part of the project titled *Odtworzenie zabytkowego układu wodno-pałacowego w Zwierzyncu wraz z zagospodarowaniem turystycznym*, RPO Województwa Lubelskiego na lata 2007-2013, Oś Priorytetowa VII: Kultura, turystyka i współpraca międzyregionalna, <http://www.magiczneroztocze.pl> (access: 29.03.2017).
- [14] Diploma design thesis author: Małgorzata Kozela, promotor: prof. Jagiełło-Kowalczyk, auxiliary promotor: D.Sc. Ph.D. Przemysław Markiewicz, Kraków 2017.
- [15] <http://www.kultura-zwierzyniec.pl/kategoria/warsztaty-wokalne> (access: 18.03.2017).
- [16] <https://zwierzyniec.e-biuletyn.pl/upload/pliki/0uchwała223.pdf> (access: 18.03.2017).



Michał Krupa (michalkrupa@zeriba.pl)
Faculty of Architecture, Cracow University of Technology

THE ORIGINS AND HISTORY OF THE SPATIAL DEVELOPMENT
OF RABKA-ZDRÓJ DURING THE MEDIEVAL PERIOD
– INTRODUCTION TO RESEARCH

GENEZA POWSTANIA I HISTORIA ROZWOJU PRZESTRZENNEGO
RABKI-ZDROJU W OKRESIE ŚREDNIOWIECZA
– WSTĘP DO BADAŃ

Abstract

This article concerns the origins and the history of spatial development of Rabka-Zdrój during the medieval period. It discusses the literature of the subject as well as research methods thanks to which it was possible to prepare prolegomena to the research on the history of Rabka within the given period. During the medieval period, Rabka was a small village owned by the Cistercian monastery in Szczyrzyc. The impulse for the development of settlement in this area must have been provided by salt deposits, known since the 12th century, which were primarily used by the monastery. Only with time they started to be explored by subsequent owners of the village. The spatial layout of Rabka developed basing on the local and long-distance routes existing in the Middle Ages, which ran through the village. The article addresses the issue of the development of the village, which suddenly transformed in the 19th century, becoming one of the most important Polish health resorts.

Keywords: Rabka, Middle Ages, history, spatial layout

Streszczenie

Niniejszy artykuł dotyczy genezy powstania oraz historii rozwoju przestrzennego Rabki-Zdrój w okresie średniowiecza. Omówiono w nim literaturę przedmiotu, a także metody badań, dzięki którym udało się opracować prolegomenę do badań nad dziejami Rabki w założonym okresie. Rabka w średniowieczu była niewielką wsią należącą do klasztoru cystersów w Szczyrzycu. Impulsem do rozwoju osadnictwa na tym terenie były zapewne znane od wieku XIII złoża soli, z których pierwotnie korzystał klasztor. Dopiero z czasem były one eksplorowane przez kolejnych właścicieli wsi. Układ przestrzenny Rabki rozwinął się na podstawie istniejących w średniowieczu traktów ponadlokalnych oraz lokalnych wiodących przez wieś. Przedmiotowy artykuł przybliża zagadnienie rozwoju miejscowości, która w wieku XIX uległa naglej metamorfozie, stając się jednym z najważniejszych polskich uzdrowisk

Słowa kluczowe: Rabka, średniowiecze, historia, układ przestrzenny

1. Introduction

Rabka-Zdrój is one of the best known Polish health resorts. It is located on the river Raba, between the Gorce and Beskid Wyspowy mountain ranges. As a village, it was established during the Middle Ages, but before, the area had been known as a salt exploitation site.

This article addresses the issue of the history of the village and its spatial development during the medieval period.

Despite its eventful history and interesting origins, Rabka does not have its monograph yet. The only scientific writings concerning the history of Rabka itself are: unpublished *Study conservation materials for the complex of the health resort and the former village* by Z. Beiersdorf and B. Krasnowolski from 1977 [1], and the article by the same authors entitled *Spatial development and architecture of Rabka* [2]. The remaining literature of the subject concerns the history of settlement in the Podhale region, the Cistercian monastery in Szczyrzyc to which Rabka belonged during the medieval period, and salt exploitation in lesser Poland during the Middle Ages. One should mention here the works by: A. Keckowa entitled *Mining work in Rabka in the 2nd half of the 16th century* [9, p. 7–34]; K. Potkański, whose findings on the subject of settlement development in Podhale were published in *Posthumous writings* [15]; J. Grzesiowski and J. Piotrowicz entitled *Salt in Lesser Poland in monastic endowments* [5]; A. Jodłowski entitled *The issue of salt exploitation in the Krakow area during the antiquity and the early medieval period* [8, p. 155] and S. Zakrzewski *The oldest history of the Cistercian monastery in Szczyrzyc* [19, p. 1–75].

The article in question was based on the following research methods: internal desk research, comparative analyses of archive cartography, and field research. Internal desk research involved primarily preliminary research of archives and literature, which was carried out in appropriate archives and libraries. The National Archives were analysed paying special attention to both historic manuscripts and printed documents (Section I – Old Polish municipal and land records, family archives and collections) as well as town plans (Section V – cartographic materials and specification sheets). Preliminary research was also carried out in the Austrian Military Archive (Kriegsarchiv) in Vienna, the Central Archives of Historical Records in Warsaw, the Branch Archive of the National Heritage Institute in Krakow; in the Archive of the Lesser Poland Voivodeship Monument Protection Office in Nowy Targ; the Jagiellonian Library and the Władysław Orkan Museum in Rabka-Zdrój. Obtained archive materials, in the form of manuscripts were translated from Latin into Polish, which allowed for putting forward new hypotheses related to particular stages of the village development and its characteristics; while archive cartography was compared and analysed paying special attention to determining the spatial form of medieval Rabka. Besides the internal desk research and comparative research, detailed field research was also carried out involving the analysis of the current functional-spatial structure of the health resort with special attention to its historical shape.

2. Rabka during the Middle Ages

It is difficult to determine the prehistory of the area where the village of Rabka developed because of a lack of archaeological research data. The archaeological photograph of Poland has not yet been prepared for the area of Rabka, and scientific studies concerning the prehistory of the Podhale and Orava regions do not contain information about the described area. One can only surmise that the land, like the entire Podhale region, was settled relatively late, which resulted from its lack of fertile soil and unfavourable climate. Some researchers believe that the first unorganised migration of settlers into the above-mentioned area took place in the 13th century [16].

The first source information related to later Rabka date back to the 13th century. In 1234, Duke Henryk Brodaty (Henry the Bearded) in the document granting ... *Theodore, the Voivode of Krakow, the right to settle Germans in some forests located by the rivers Dunajec, Rogoźnik and other* mentioned the name *fluvium Slona* which can be identified with the name of the stream *Slona* or *Słonka* in Rabka [5, p. 96; 11].

Another vital document, considering the origins of Rabka, dates back to the year 1254. In the document, entitled *Clemens et Marcus filii condam Marci palatini Cracoviensis, omnes donationes a partuosuo comite Theodoro coenobio Cisterciensium Ciriciensis factas, assentiente Boleslao duce Cracoviae et Sandomiriae, renovant*, Duke Bolesław Wstydlivy (Boleslaw the Chaste) confirmed the endowments for the Cistercian monastery in *Szczyrzyc* [11], mentioning e.g. lands in the vicinity of the river *Ponecza* that can be interpreted as the Poniczanka River, and the name *sal de Rapschycza* which must have referred to the already known salt deposits in Rabka. From the above document, one can conclude that the area of the later village must have belonged to the Cistercian monastery in *Szczyrzyc* during the early medieval period.

More information essential for these considerations is included in the work *Liber Beneficiorum dioecesis cracoviensis...* from the years 1470–1480, written by Jan Długosz, in which in the section entitled: *Cziczczense. Quintum monasterium Ordinis Cisterciensis in dioecesi episcopatus Cracoviensis consistons* the chronicler wrote about the estates belonging to the Cistercian monastery in *Szczyrzyc* [4, p. 436]. He mentioned several places, including those located in the Podhale region: Nowy Targ, Szaflary, Waczmuth, Klikuszowa, Poręba, Niedźwiedź, Podobin, Olszówka, Mszana, Kasina, Ogonówka, Dębno, Ostrowsko, Ostrowiec, Wilcze Pole, Ludźmierz, Krauszów, Rogoźnik and Rabka where the monastery obtained salt [19, p. 50]. It is worth noticing that in the text by Jan Długosz there appeared the name *Sal in Rabka* which the chronicler did not label as a *villa* (village) – as he did other places e.g. Ludźmierz. Therefore, it is rather difficult to decide when the first settlement began to crystallise here. However, one can assume that the deposits of salt and brine [8, p. 155], known since the mid-13th century, must have been a key factor contributing to the fact that the first settlers began to arrive here.

Considering the issues related to the settlement development in Rabka, one can surmise that the village grew slowly, just like others belonging to the Cistercians. It was connected with the fact that the area had previously been completely uninhabited, most probably owing to difficult climate and poor-quality soil [13, p. 176–177].

The Cistercian Order from *Szczyrzyc* conducted their urbanisation activities in the Podhale region until 1382 when King Louis I of Hungary divested the monastery of its estates

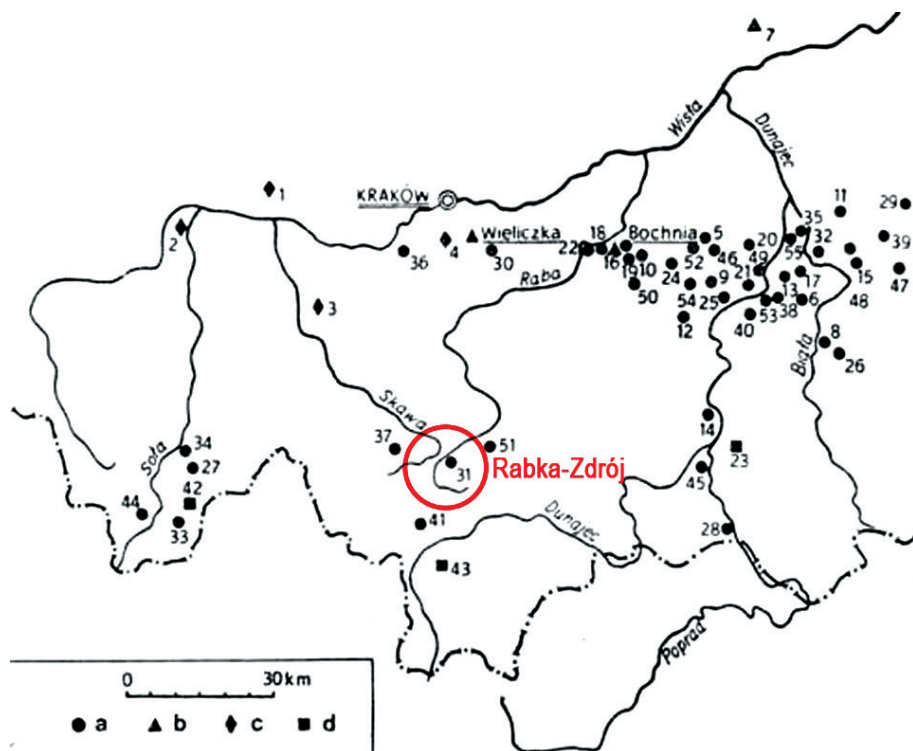


Fig. 1. Map of salt springs in western Lesser Poland during the early Middle Ages with marked springs in Rabka (Source: [8, fig. 1, p. 140])

located in the Podhale region. The simplest explanation of the fact is that, in those times, kings (particularly King Kazimierz Wielki – the Great) frequently divested the church of its lands and included them in the royal demesne [1]. In this case, however, S. Zakrzewski, who carried out research in the Szczyrzyc archives, believes that the actual reason why King Louis I the Hungarian divested the monastery of its lands was their mismanagement. According to Zakrzewski, the King’s decision was strongly influenced by one case when the abbot of Szczyrzyc let the castle in Szaflary to a Jew who counterfeited coins there [19, p. 51]. In the following years, in answer to the abbot’s appeals, the King returned two villages: Ludźmierz and Krauszów, to the monastery. He did not return the other, as he died in 1382 [19, p. 51].

During the reign of Władysław Jagiełło, Rabka was handed over to Zygmunt, the Vogt from Dębowy Dział (Dębowiec), the testimony of which is a note in the court records of the former Krakow Lands, where it was written that, in the year 1403, he sold the villages of Rabka and Kasinka to Jan Ligęza: “Zygmunt, the current vobt from Dębowiec, his two villages called Rabka and Kasina which he possessed in accordance to what is written in the document of King Kazimierz, he sold and handed over for 1000 ‘grzywnas’ of Prague grosz, to Jan Ligęza, Voivode of Łęczycza, with all the rights and property rights and appurtenances, pledging in accordance to the legal formula that he reserved nothing for himself forever” [18]. In the books, there is a reference to the “document of King Kazimierz” that might have been a document of the first foundation of Rabka, and which has not been found until the present day.

Thus, since the year 1403, Rabka was the property of Jan Ligęza of the Półkozic coat of arms, who might have died in 1419. He was the Voivode and starost of Łęczycza. In archive documents, he was also called Jan from Niewiarow; Jan from Niegowic; Jan from Przeclaw and Jan from Bobrek. The lands belonging to Jan Ligęza were scattered. Besides Rabka he owned e.g.: the already mentioned Kasina, Jurków, Niewiarów and Niegowic, Sławkowice, Kamyk, Witowice, a part of Zręczycze, Brzeźnicę, a part of Brniki and Świdówka, the estate in Gorzkow, Bobrek, a part of Libiąż, Chrzanów, Czernichów, Prokocim, Przeclaw and endowment on the Vogt property in Dębowiec near Jasło, which he must have purchased from Zygmunt to follow Rabka and Kasina [14, p. 316–317].

Another mention of Rabka dates back to the year 1440. It is known that, at the time, the village belonged to the descendants of Jan Ligęza, who went by the name of Niewiarowski [7, p. 61 (561)]. In the court records of the former Krakow Lands there appeared Jan from Niewiarow who „Appearing in court [...] he stated and committed himself to [...] consenting to court introduction to possession of Rabka property and, like he committed himself in writing, to pledge 58 ‘grzywnas’. That interest on the capital he vowed to pay to the noble Anna, the wife of Mirosz from Zagajow, and he should give her that court introduction to possession of Rabka property...” [6, p. 463].

Yet another, probably the most important medieval document concerning Rabka, is the one dated to August 16, 1446. It refers to the renewal of the privilege for the foundation of the Rabka village for Andrzej and Piotr, sons of Jakusz, a free peasant from Olszówka, made by Mikołaj, the squire from Niewiarow. In the privilege, one can read e.g. a suggestion that in the mid-15th century the area of Rabka was abandoned, which was to be altered by the village foundation – “to people abandoned places with numerous settlers”. It means that, previously, a settlement must have existed here, which became depopulated for unknown reasons. Here, one should recall the above-mentioned “*lost*” document of the first foundation of Rabka from 1364, in which King Kazimierz Wielki allowed Mikołaj from Uście to found, according to the Magdeburg law, villages located between the rivers: Raba, Słonka, Łętownia and Kasina [17, p. 344]. It is worth noticing that, in 1360, probably the same Mikołaj founded the village of Lubień near Myślenice [10, p. 30], which suggests that the area was within the reaches of the urbanisation programme carried out by the King. It also has to be emphasised that the vicinity of Rabka as the land rich in salt deposits, which was known since the 13th century, might have been of special economic significance for the King. During his reign, the Royal Treasury acquired as much as 36% of its income from the salt mines [5, p. 73]; therefore, the desire to urbanise the surroundings of Rabka seems rather obvious.

The new foundation charter of Rabka was granted to brothers Andrzej and Piotr, from Olszówka. The Olszówka in question must have been the one currently located in the Limanowa County, the Mszana Dolna district, established in the year 1388 by King Władysław Jagiełło. In reference to the foundation act of Rabka, one has to emphasise that, according to the described document, the village of *Raba* was to have been laid out within precisely outlined limits: “...from the upper river Raba flowing below to the river called Biały Potok, and from that river i.e. the stream running below to the river called Kamienny Potok and from



there moving forward to the river known as Potok, and then going down from that river to the river or stream called Gliniasty from the upper Gliniasty (stream) to the summit of (Mount) Skomielna and from Mount Skomielna to Mount Lubański and from Mount Lubański to half-way up Mount Lubań and from half-way up the mentioned mountain to the stream called Miedziany from Miedziany Potok Mountain to the Skalisty stream and from that same stream climbing up to the Słony Potok stream and from it going up to the river called Poniczna and from the same river to the main river Raba...". It is also known that Rabka was founded according to the Magdeburg rights. Within it, ten Franconian lans were measured out, from which one lan "...in a better place and land..." was to be given to the church "... in the hope of our salvation and to develop for the church the community of the said village..."; three lans were to be bequeathed to the aldermen of Rabka, while the next one was intended for a *skotnica* i.e. the place where cattle was driven or pasture. It should also be mentioned that, at that time, 1 Franconian lan equalled around 25 ha [18, p. 59–63]. According to the discussed privilege, locators were also "entitled to build one mill with one wheel, as it had been for a long time or where it would seem better and more beneficial". The quoted fragment again suggests that the settlement used to exist once before ("as it had been for a long time"). From Mikołaj, the squire of Niewiarow, locators received an additional privilege to set up a mill with two waterwheels: "If there were not enough water, let them build two mills, but each of those with one wheel". He also allowed them to dig two ponds in Rabka and to build two inns. The act in question also ensured income for the brothers Andrzej and Piotr in the form of „every sixth denarius from rent, every third denarius from court fines [...] from stalls of craftsmen and cobblers and bread and meat stalls". Other bequeaths for locators referred to the right of "hunting, catching all game, fowl and fish in the rivers within [...] the village". The brothers, who were to serve as aldermen, were also entitled to judge and condemn persons to blame according to the German law. Moreover, the discussed document repealed "all Polish laws and customs" which might be contrary to the German law. In order to accelerate the urbanisation of Rabka and to encourage newcomers to settle in the newly founded village, Mikołaj, the squire from Niewiarow, exempted them from all taxes, duties and obligations to him for the period of 25 years. After 25 years, the aldermen of Rabka and their successors were obliged to pay two 'grzywnas' in Polish coinage for military expeditions. Inhabitants were to pay "from each lan every year on St. Martin's day [...] eight 'skojec' in the same coin" and "a tithe to the one who is entitled to it by the law, [...] custom and norm observed for years in Nowy Targ".

Summing up the foundation charter, it can be surmised that, like it was suggested above, it was the second foundation after the unsuccessful one from the times of Kazimierz Wielki.

About 40 years later, Rabka was already in the hands of the Jordan family from Zakliczyn, of the Trąby coat of arms. According to the sources, it took place around the year 1484, which means that it occurred during the lifetime of Jan who died in 1507, as well as of his brother Mikołaj who died in the year 1521. Literature of the subject differs as to which of the Jordans actually purchased Rabka because of the lack of virtually any source documents concerning the transaction.

Describing the history of the place, one cannot ignore the issue of the parish existing there. It was established in Rabka as late as the year 1557, thanks to the foundation of Spytek Jordan

who had the first church dedicated to St. Mary Magdalene erected here. Previously Rabka belonged to the parish in Mszana Dolna, and later in Łętownia [12, p. 23]. The relatively late creation of the parish in Rabka confirms the thesis that the village grew very slowly, and the inflow of a larger number of settlers took place only in the 16th and the following centuries [13].

The moment when Rabka was taken over by the Jordan family, symbolically ends the medieval history of the village. Moving on to summarising and outlining initial findings related to the spatial development of the village of Rabka, one has to say that salt deposits were the main factor influencing settlement development in this area. Yet, despite that factor, the settlement process was rather slow, and the first attempt at founding the village, carried out by King Kazimierz Wielki, did not succeed. The failure might have resulted e.g. from the fact that the mountainous land with infertile soil was not sufficiently attractive for settlers.

In the foundation charter from August 16, 1446, just like in the majority of foundation documents, there is no information about the shape of the established village.

A hypothesis could be put forward that, originally, the village of Rabka developed along two routes. One of them was long-distance and led from Mszana Dolna to Nowy Targ via Rabka. The other was a local road and led to the site where salt was obtained, which must have been in Słone (the east part of the present-day spa).

It is difficult to state when precisely the road to Chabówka and Skawa emerged from the route leading from Mszana Dolna to Nowy Targ. It might already have occurred during the medieval period, since Skawa, like Rabka, used to belong to the Mszana parish then. The fork of the road mentioned here was located in the southern section of the village, close to the ford on the river Raba. On this site, the market square evolved in the shape of a fork, even today known in Rabka as the “main square”. Such a spatial layout of Rabka from the 18th century was marked on the Mieg’s Map. The Mieg’s Map i.e. the topographic map of the Kingdom of Galicia and Lodomeria from the years 1779–1783 (*Originalaufnahme des Königreiches Galizien und Lodomerien*) is a large-scale (scale 1:28 800) map of the Kingdom of Galicia and Lodomeria drawn by the Austrian Staff of the General Quatermaster under the supervision of officer Friedrich von Mieg [3, p. 91–112]. The map of Galicia was sketched at the last possible moment before the approaching radical economic and social changes of the 19th century. Because of its precision and the time when it was created, it is a valuable source depicting the landscape of the still Old-Poland period.

An analysis of the history of the Rabka village offers a basis for a hypothesis that its spatial shape did not change much between the medieval period and the 19th century when Rabka became a spa. Thus, it can be assumed that the image of Rabka from Mieg’s map does not differ radically from the one of the village during the high Middle Ages, naturally with the exception of two objects marked on the map: the parish church of St. Mary Magdalene and the manor, which did not yet exist in the medieval period.

This article ought to be concluded by a statement that the health resort in Rabka, which was established in the 19th century, crystallised on the site of a small medieval village that basically did not evolve into a larger spatial layout until the spa, which required laying out a new urban structure, was founded. The structure was realised to the east of the village of Rabka, in the vicinity of brine springs.





Fig. 2. Fragment of Mieg's map from the years 1763–1787 representing Rabka
(Source: Copy of the map [in:] Archive of KHAUiSzP WA PK. Original [in:] Kriegsarchiv, s.v.)

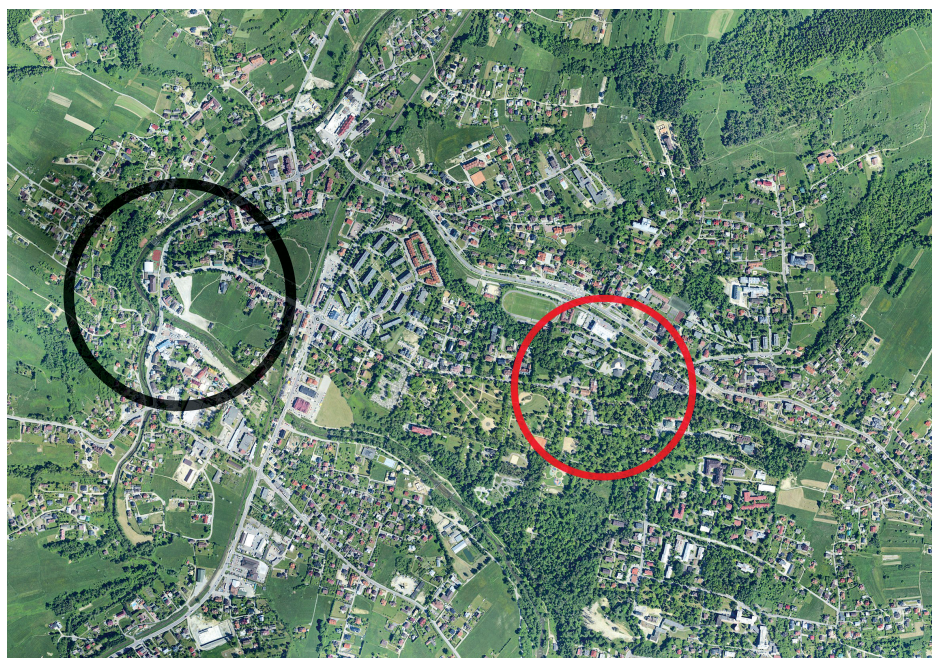


Fig. 3. Rabka on an aerial photo. Black colour marks the area where the village of Rabka developed in the Middle Ages. Red colour marks the area where the health resort was founded in the 19th century (Photo by W. Gorgolewski 2017)

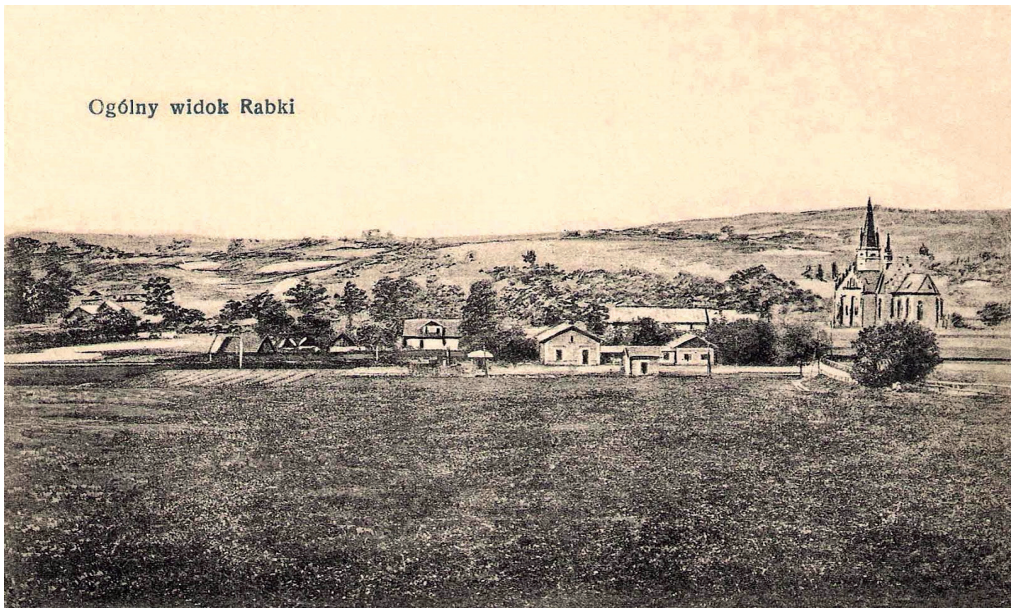


Fig. 4. Rabka at the beginning of 20th century, view from the east.
The historic postcard (Source: Author's archive)

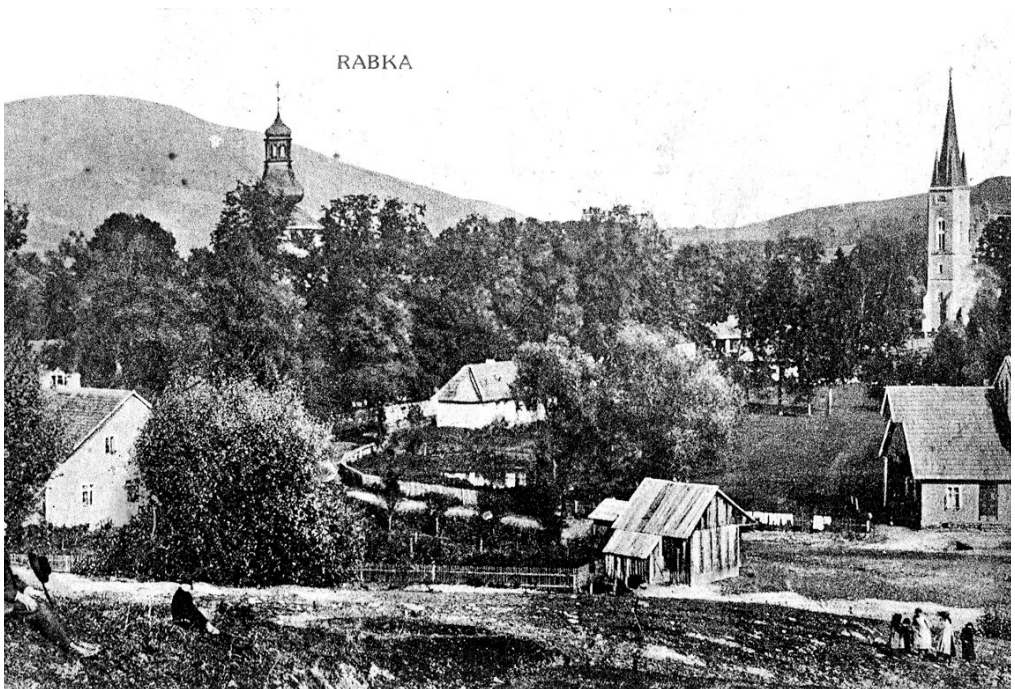


Fig. 5. Rabka at the beginning of 20th century, view from the west.
The historic postcard (Source: Author's archive)

References

- [1] Beiersdorf Z., Krasnowolski B., *Rabka (woj. Nowosądeckie). Studialne materiały konserwatorskie dla zespołu uzdrowiska i dawnej wsi*, typescript, P.P. Pracownie Konserwacji Zabytków, Kraków, Pracownia Dokumentacji Naukowo-Historycznej, Kraków 1977.
- [2] Beiersdorf Z., Krasnowolski B., *Rozwój przestrzenny i architektura Rabki*, Teka Komisji Urbanistyki i Architektury, v. XIV, 1982.
- [3] Bukowski W., Janeczek A., *Mapa józefińska Galicji (1779–1783) w przededniu edycji. Przedmiot i założenia programu wydawniczego*, *Studia Geohistorica. Rocznik Historyczno-Geograficzny*, No. 1, 2013.
- [4] Długosz J., *Joannis Długossii Senioris Canonici Cracoviensis Opera omnia / Liber beneficiorum dioecesis cracoviensis nunc primum e codice autographo editus*, v. 9, part 3, Publ. Ex typographia Kirchmajeriana, Kraków 1864.
- [5] Grzesiowski J., Piotrowicz J., *Sól małopolska w nadaniach i przywilejach dla klasztorów*, [in:] *Studia i materiały do dziejów żup solnych w Polsce*, Muzeum Żup Krakowskich, Wieliczka 1965.
- [6] Helcel A. Z., *Starodawne Prawa Polskiego Pomniki*, v. 2, Kraków 1870, No. 1010 (863).
- [7] *Herbarz Polski Kaspra Niesieckiego S. J.*, Wyd. Jan N. Bobrowicz, v. VI, Lipsk 1841.
- [8] Jodłowski A., *Problem eksploatacji soli w okolicach Krakowa w starożytności i we wczesnym średniowieczu*, [in:] *Archeologia Polski*, v. XIV, i. 1.
- [9] Keckowa A., *Prace górnicze w Rabce w 2 połowie XVI wieku*, [in:] *Studia z dziejów górnictwa i hutnictwa*, v. 9, Wrocław-Warszawa-Kraków 1965.
- [10] *Kodeks Dyplomatyczny Małopolski*, v. 1, F. Piekosiński, Kraków 1876, p. 21, doc. no 15; v. 1.
- [11] *Kodeks Dyplomatyczny Małopolski*, v. 1, F. Piekosiński, Kraków 1876, p. 46, doc. no 40.
- [12] Kowalczyk K., *Miejsca kultu religijnego w Rabce Zdroju*, Instytut Teologiczny Księża Misjonarzy, Kraków 2003.
- [13] Kumor B., *Powstanie i rozwój sieci parafialnej*, "Prawo Kanoniczne", Publ. Akademia Teologii Katolickiej w Warszawie, Year V, Warszawa 1962, No. 3–4.
- [14] *Polski Słownik Biograficzny*, v. XVII, PAN/PAU, Kraków-Warszawa 1972.
- [15] Potkański K., *Pisma pośmiertne*, Poznańskie Towarzystwo Przyjaciół Nauk, Poznań 2004.
- [16] Rutkowska-Płachcińska A., *Sądeckczyzna w XIII i XIV wieku. Przemiany gospodarcze i społeczne*, Zakład Narodowy im. Ossolińskich, Wrocław 1961.
- [17] *Słownik Geograficzny Królestwa Polskiego i innych Krajów Słowiańskich*, B. Chlebowski, W. Walewski (ed.), Nakł. W. Walewskiego, Warszawa 1889, v. IX.
- [18] Stamm E., *Miary powierzchni w dawnej Polsce*, [in:] *Rozprawy wydziału Historyczno-Filozoficznego*, edition II, v. 45, PAN, Kraków 1936.
- [19] Zakrzewski S., *Najdawniejsze dzieje klasztoru Cystersów w Szczyrzycu (1238–1382). Przyczynek do dziejów osadnictwa na Podhalu*, *Rozprawy Akademii Umiejętności. Wydział Hist.-Filoz.*, ser. II, v. 16, Kraków 1902.

Michał Nessel (mnessel@pk.edu.pl)

Division of Descriptive Geometry, Technical Drawing & Engineering Graphics, Institute of Construction Design, Faculty of Architecture, Cracow University of Technology

THE RELATIONSHIP BETWEEN CULTURAL AND GEOGRAPHIC PARAMETERS
IN ALGORITHMIC DESIGN – SAMPLE ALGORITHM

ZALEŻNOŚĆ MIĘDZY PARAMETRAMI KULTUROWYMI I GEOGRAFICZNYMI
W PROJEKTOWANIU ALGORYTMICZNYM – PRZYKŁADOWY ALGORYTM

Abstract

The article presents selected issues related to the role of cultural and geographic parameters in architectural and urban design. The author defines the groups of parameters and analyses a relationship between them and its influence on design process. The relationship between the groups of parameters is presented at the example of evolutionary algorithm that searches for proper location for Roman military camp. Presented algorithm is an example of possible usage of the relationship between the groups of parameters in modern algorithmic design where cultural factors are used for evaluation of geographic parameters.

Keywords: cultural and geographic parameters, algorithmic design

Streszczenie

Artykuł przedstawia wybrane zagadnienia związane ze znaczeniem parametrów kulturowych i geograficznych w projektowaniu architektonicznym i urbanistycznym. Autor definiuje grupy parametrów oraz analizuje zachodzącą między nimi zależność i jej wpływ na proces projektowania. Relacja pomiędzy grupami parametrów została przedstawiona na przykładzie algorytmu poszukującego lokalizacji dla historycznego obozu rzymskiego. Prezentowany algorytm jest przykładem możliwości zastosowania zależności pomiędzy grupami parametrów we współczesnym projektowaniu algorytmicznym, gdzie czynniki kulturowe mogą zostać wykorzystane w celu ewaluacji parametrów geograficznych

Słowa kluczowe: parametry kulturowe i geograficzne, projektowanie algorytmiczne

1. Introduction

The aim of the research is to analyse selected issues related to the relationship between cultural and geographical parameters and its influence on the design process. Due to expansion of parametric and algorithmic design better understanding of fundamental parameters and relationships between them is a key to more effective, conscious and successful design [1, 2].

A starting point for the research was a question: whether cultural or geographic parameters play the main role in the design process. Janusz Ballenstedt in a book: “Architektura. Historia i teoria” emphasises the role of geographic parameters in the development of architectural styles and in design [3]. From the other side, the importance of cultural factors is widely debated by Patrik Schumacher, for example in the paper: “The Stages of Capitalism and the Styles of Architecture” and in the interview “On Styles, Society, and Architecture’s Communicative Capacity”, where he states that architecture “...is a dynamic layer of a multi-layered society” [4, 5]. The definition of architecture strongly connects the design with human factors. Due to this fact, we have decided to focus not on the role of each group separately but on the relationship between them. Such approach leads to more natural observations. Human beings are always a part of design process, so the groups of parameters always come together.

Being involved in research in the field of evolutionary computing in design we have observed a correlation between the relationship characteristic for the group parameters and fitness function, which is an important element of evolutionary algorithms [6, 7]. The research is focused on the possibility of using cultural parameters as a starting point for the definition of fitness function in evolutionary computing for evaluation of geographic parameters in algorithmic design.

In this paper, we describe the location of the research, define and explain the groups of parameters and finally, we present and visualise the key relationship at the example of an evolutionary algorithm that searches for the location for historical Roman military camp [8–10].

The basic tools for final visualisation are: Rhinoceros 5, Grasshopper and Galapagos, the tools that are commonly used for algorithmic design and evolutionary computing in architecture and urban design [11–13].

2. Location of the research

The research area covers the coast of Adriatic Sea between the towns of Makarska and Venetia, now parts of Croatia, Slovenia and Italy. Over the centuries the region was ruled by many different cultures and countries, like: Illyrian tribes, ancient Rome, Slavic tribes, Venetia, Hungary, Austria, Yugoslavia with Soviet Union influences and finally Italy, Slovenia and Republic of Croatia in the southern part of the coast [14]. The variety of cultures that were creating an architectural landscape of this region at similar geographic parameters was crucial in the selection of the location for this research. It allows to isolate the influence of cultural parameters on the design process and to specify their role and relationship with geographic parameters more precisely.

3. Groups of parameters

Architectural and urban design is determined by vast number parameters. The analysis of each parameter separately would be very difficult or even impossible and would lead to unintelligible results of research. For that reason, the parameters were divided into two main groups: cultural and geographic. Definition and nature of this division is an important element of the research.

The group of cultural parameters is defined as a set of parameters related to human activities and associated with entities that have true impact on the design process, like architects, designers, construction engineers, builders, but also officials, investors and representatives of the ruling class. The group consists of following factors: ethnic origin, cultural background, customs, beliefs, aesthetic background, relation to the monuments, political and economic situation, level of technology and civilization, individual features, etc.

Human activity factors associated with location only, like: existing buildings, societies and their historical background are classified as cultural context and are included in geographic parameters group. This is crucial for understanding the relationship between the groups of parameters that is presented in this paper. The reason for placing cultural context in the group of geographic parameters is the fact that cultural background of entities being involved in design process may be different than cultural background of the location.

The group of geographic parameters is defined as a set of parameters related to location of the project, like: terrain, climate, accessibility of materials, the presence of reservoirs and waterways, flora, fauna, cultural context, etc.

4. Relationship between cultural and geographic parameters

During the field research, we stated that at selected locations we can observe a variety of different design solutions at very similar geographic parameters, for example: Town hall in Piran rebuilt under Austrian rules side by side with Venetian house built under Venetian rules, The Roman Arena in Pula and nearby modernistic building with the remnants of Roman Octagonal Mausoleum, Pre-Romanesque St Donatus' Church in Zadar built exactly at the place of Roman Forum, etc. (Fig. 1).

It means that the diversity of design in such cases is determined mostly by cultural parameters. Due to this fact, we suggest that geographic parameters do not have a direct impact on design. It does not mean that they do not have an impact at all. Their importance is determined by cultural parameters as defined in this article. In other words, the key is the relation of designers and other entities that have a true impact on the design process to geographic parameters associated with given location. For example, the relation of the designer, which is determined by cultural parameters, to the cultural context associated with a specific location can lead to different decisions, like: protect or destroy existing buildings. In such case, cultural context, which is classified as a geographic parameter, doesn't determinate design solutions and the final decision is determined by cultural parameters.

Basing on the described relationship we observed that the value of constant set of geographic parameters may vary depending on cultural factors. This led us to the conclusion that cultural parameters may be used for the evaluation of geographic parameters in evolutionary computing in algorithmic design.

5. Basic algorithm

To visualise and test suggested relationship between groups of parameters, we have created an algorithm that uses evolutionary computing and cultural factors for evaluation of geographic parameters (Fig. 2). The aim of the algorithm is to search for a proper location for historical Roman military camp in given terrain.

5.1. Overview

Galapagos, an evolutionary computing tool for Rhino, generates the first randomised population of locations. Each location is evaluated by fitness function, which bases on cultural parameters. Better locations are chosen for reproduction, while worse ones are eliminated from the population. During the reproduction process, Galapagos generates a new generation of locations. Evaluation, selection and reproduction processes are repeated until a good location is found.

5.2. Parameters

Geographic parameters for the algorithm are reduced to curves that represent the shoreline and levels. The terrain data used in the simulation represents a part of the Adriatic coast between the towns of Červar and Zelena at Istrian Peninsula in Croatia (Fig. 3a). The area covers the centre of Poreč town, where Roman Castrum was built during II century BC. Thanks to this fact we can compare the location chosen by the algorithm with the true location of Roman Castrum in this area.

It needs to be noted that it is difficult to define some of the cultural factors as simple parameters in algorithmic design due to their complexity. For that reason, we use a cultural pattern of Castrum and its properties for the definition of cultural parameters' group. It should be mentioned that those properties are a result of more basic parameters, for example, related to the political and economic situation, size of military troops and strategy.

Cultural parameters are reduced to the shape and size of the Roman military camp and its desirable accessibility: protected by the shape of shoreline but easily accessible by Roman military troops (flat terrain). Those factors are a starting point for the definition of a fitness function that evaluates the locations generated during evolutionary computing process.

5.3. Genome

The genome for evolutionary computing is defined as X position, Y position and rotation of the shape of Roman Castrum. In Grasshopper and Galapagos each of genes is represented by a slider. The ranges of X position and Y position sliders are set to correspond with position and size of the curves that represent the terrain.

5.4. Fitness function

The parameters of shape and size of Roman military camp are passed to fitness function as a rectangle with specified dimensions. Accessibility factors are defined as an algorithm. The rectangle is divided into twelve equal segments. The algorithm measures the distances between division points and curve that represents shoreline. Each of distances is squared and those values are summed together being a starting point for fitness value. Galapagos is set to search for the minimal value of fitness function, what means that the lower fitness value the better location.

The algorithm then detects if division points are located on land or on sea by checking the side of shoreline on which each of the points is located. Fitness value is multiplied by the number of points located on sea + 1.

Finally, the algorithm detects intersections between area of the rectangle and curves that represent levels. Fitness value is multiplied by total length of curves that are a result of the intersection divided by 20 and increased by 1.

5.5. Galapagos settings

Galapagos is set to search for the minimal value of fitness function as described in point 5.4 Fitness function. We use the population of 200 locations with an initial boost of 2. Maintain parameter that defines the percentage of individuals that can be carried over to a new generation, is set to 10% and inbreeding parameter is set to -50%. Those settings increase the diversity of individuals in a population and are crucial for the probability of finding the best location.

5.6. Results of computing

During the computing process algorithm analyses, many of locations and detects that the most valuable ones are located on the peninsula of Poreč town, what corresponds with the true location of Roman castrum in this area. Depending on the randomly generated first generation algorithm finds the almost exact location of historic Roman castrum or a similar one (Fig. 4).



6. Additional simulations

To verify the algorithm, two more tests have been done. In one of them, we change the terrain data by adding additional levels at peninsula where the historical military camp was located (Fig. 3b). The other assets and parts of the algorithm are the same as in the case of the basic one described in previous division. In this case the algorithm detects that the most valuable location is placed at flat terrain by the shoreline close to Mali Maj town (Fig. 5a). The locations at the peninsula, where Poreč town is located, are omitted as well as other locations where the terrain is not flat.

In the case of second alternative test, we use original terrain data, the same as in the case of the basic algorithm (Fig. 3a). However, we change a part of the fitness function, which is related to accessibility factors and evaluates the levels. The algorithm detects intersections between the area of the rectangle and curves that represent levels and measures total length of curves that are a result of the intersection. The length is divided by 20 and increased by 1 as in basic algorithm. However, the fitness value is not multiplied, but divided by this value.

It changes the meaning of desired accessibility of location from the one suitable for Roman castrum to the one characteristic for a medieval castle. We still use the shape of Roman military camp, so it doesn't exactly represent the algorithm suitable for medieval cultural factors. However, even such modification of basic fitness function affects the results of evolutionary computing. A tiny peninsula, which is located in the southern part of the terrain, is now chosen as best location (Fig. 5b). The location is well protected by the shoreline and additionally, there is a small hill that makes this location more valuable to the modified version of the algorithm. A peninsula, where Poreč town is located, is omitted due to flat terrain.

7. Conclusions

The simulations presented in the article confirm the correlation between fitness function and the relationship between cultural and geographic parameters. It should be emphasized that the basic version of algorithm precisely points to the location where historical Roman castrum was located. The accuracy of the computing is very high even despite the fact that the geographic parameters were reduced to terrain only. It should be mentioned, that true process of finding the location for the Roman military camp was much more complex and based on a larger number of parameters including the cultural context. However, despite of the simplifications the algorithm finds correct location.

It should be also noted that modifications in fitness function affect the values of geographic parameters and the final results of computing. That confirms the observation described in division 4 that the value of geographic parameters in design may depend on cultural factors. As Janusz Balenstedt stated, "nature creates the opportunities and human is the one who can choose out of them ... human is free to choose, but not the humankind" (translation: author) [3]. The relationship presented in this paper is essential at the level of the design process; however, it is important to be aware of the fact that it does not cover the whole complexity of the relationships between the culture and nature.

The presented algorithm by itself is an example of a practical implementation of the results of this research in design. By changing the fitness function, the algorithm can be adapted to current cultural factors and expectations and used in the modern algorithmic design. An additional conclusion is that evolutionary computing can be an effective tool that supports research in the field of architecture and urban design as presented in the article.

References

- [1] Schumacher P., *The Parametricist Epoch: Let the Style Wars Begin*, AJ The Architects' Journal, 2010, No. 16, Vol. 231, May 06.
- [2] Helenowska-Peschke M., *Parametryczno-algorytmiczne projektowanie architektury*, Politechnika Gdańska, Gdańsk 2014.
- [3] Ballenstedt J., *Architektura. Historia i teoria*, Wydawnictwo Naukowe PWN S.A., Warszawa 2000.
- [4] Schumacher P., *The Stages of Capitalism and the Styles of Architecture*, ASA web-magazine & (forthcoming): UED (Urban Environment Design) Magazine – International Edition, Beijing.
- [5] Schumacher P., *On Styles, Society, and Architecture's Communicative Capacity*, Architettura PostDecostruttivista Vol. 2 – La Maniera Biomimetica, Deleyva Editore, Roma-Monza, 2016.
- [6] Godlberg D.E., *Genetic Algorithms in Search, Optimization, and Machine Learning*, MA: Addison-Wesley Longman Publishing Co. Inc., Boston 1989.
- [7] Krawiec K., *Evolutionary feature programming, Cooperative learning for knowledge discovery and computer vision*, Wydawnictwo Politechniki Poznańskiej, Poznań 2004.
- [8] Książek M., *Materiały pomocnicze do studiów w zakresie historii urbanistyki*, Politechnika Krakowska, Kraków 1996.
- [9] Słodczyk J., *Historia planowania i budowy miast*, Uniwersytet Opolski, Opole 2012.
- [10] Wróbel T., *Zarys historii budowy miast*, Zakład Narodowy Imienia Ossolińskich, Wrocław-Warszawa-Kraków-Gdańsk 1971.
- [11] Rhino 5 Features, <https://www.rhino3d.com/features> (access: 18.05.2017).
- [12] Grasshopper, <http://www.grasshopper3d.com> (access: 18.05.2017).
- [13] Galapagos, <http://www.grasshopper3d.com/group/galapagos> (access: 18.05.2017).
- [14] *Wielka Encyklopedia PWN*, Warszawa 2001.



Fig. 1. Examples of buildings designed at similar geographic parameters and different cultural ones: a) Town hall in Piran rebuilt under Austrian rules and b) Venetian house in Piran built under Venetian rules, c) The Roman Arena in Pula and d) modernistic building in Pula with the remnants of Roman Octagonal Mausoleum, e) Pre-Romanesque St Donatus' Church in Zadar built on the Roman Forum, f) Remains of Roman buildings built in to the foundation of St Donatus' Church (M. Nessel, 2017)

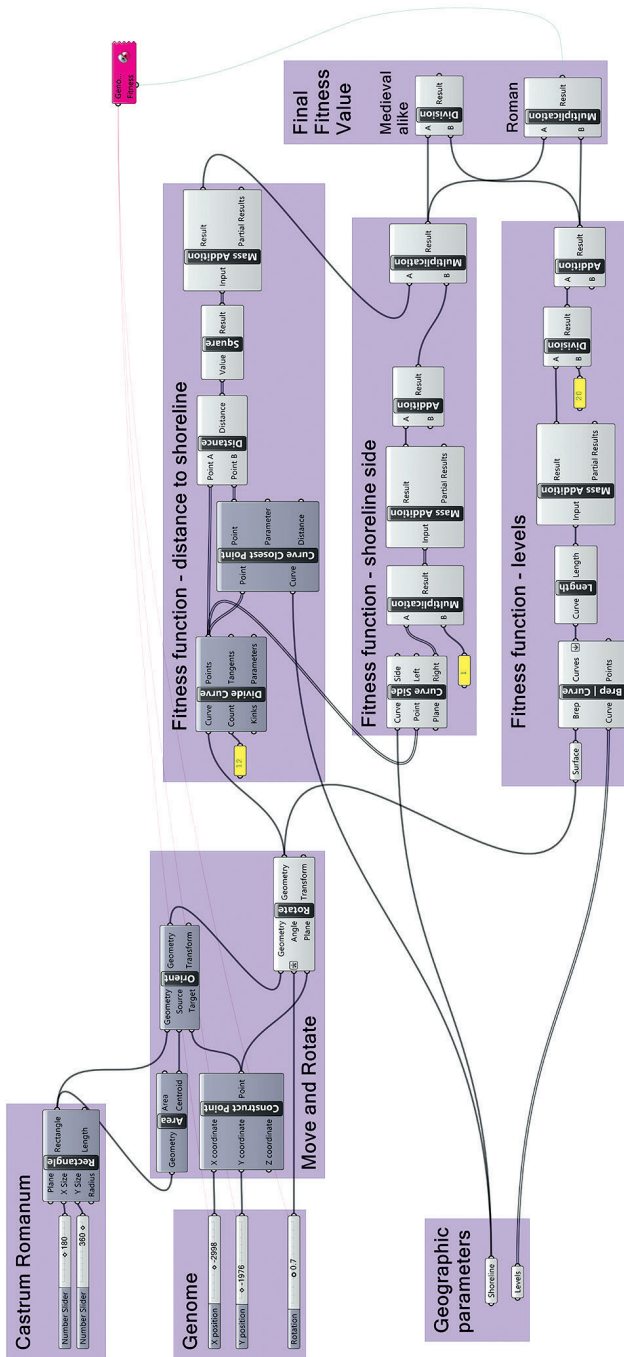


Fig. 2. Basic algorithm as described in division 5. (M. Nessel, 2017)



Fig. 3. Terrain data: a) original one used in basic algorithm, b) modified terrain data used in first additional test with additional levels at peninsula of Poreč town. (M. Nessel, 2017)

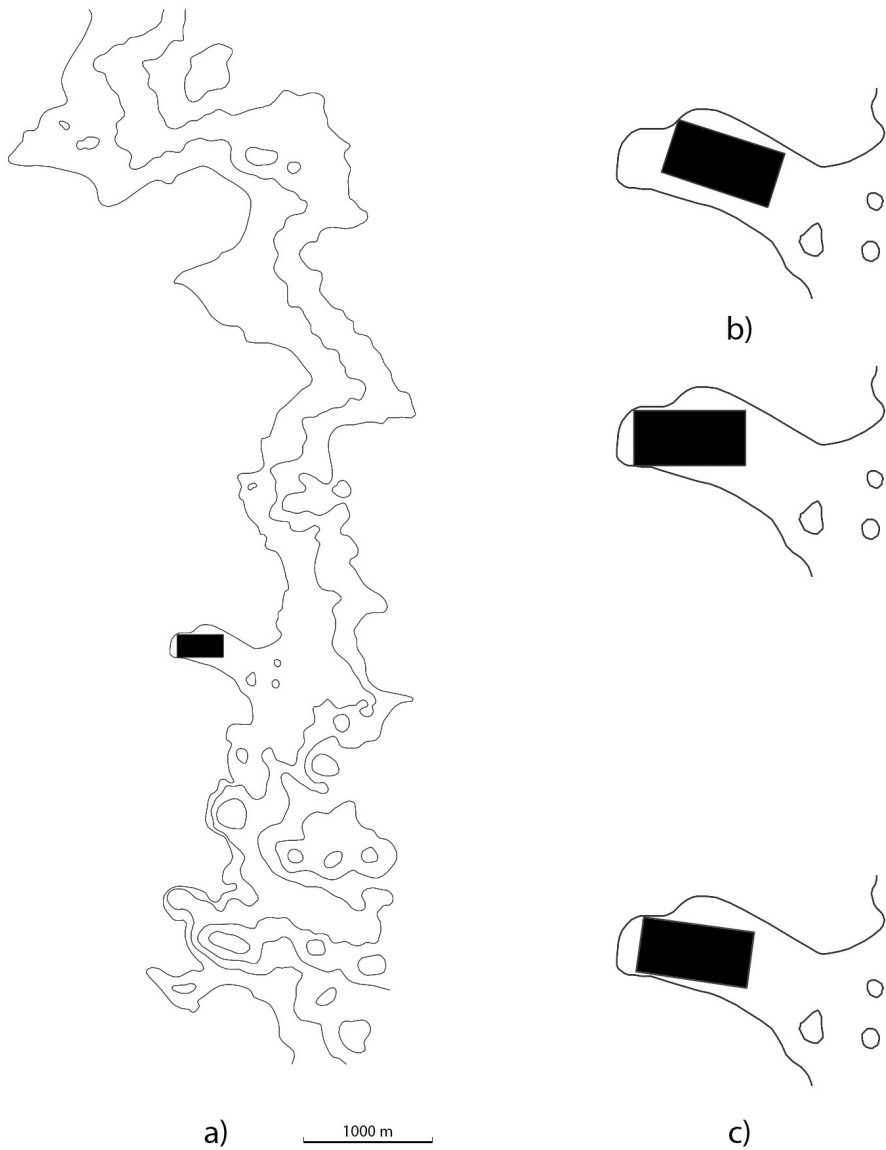


Fig. 4. Results of basic test: a) location for Roman castrum found by basic algorithm, b) close up of alternative locations found in different runs of basic algorithm, c) approximate location of historical Roman castrum (M. Nessel, 2017)

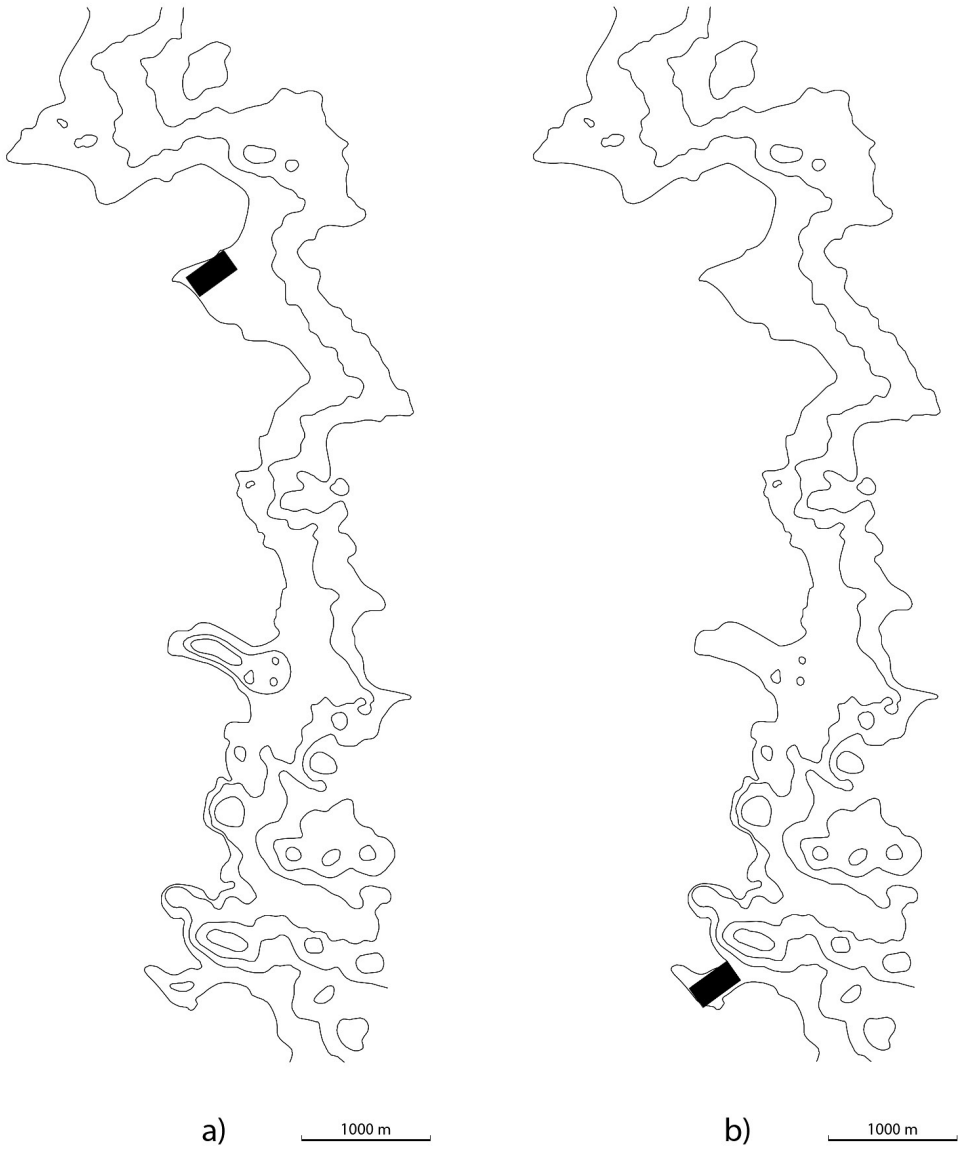


Fig. 5. Locations for Roman castrum: a) found in first additional test with modified terrain data, b) found in second additional test with modified fitness function (M. Nessel, 2017)

Magdalena Rzeszotarska-Palka (mrzeszotarskapalka@zut.edu.pl)

Department of Landscape Design, Faculty of Environmental Management and Agriculture,
West Pomeranian University of Technology in Szczecin

Magdalena Czałczyńska-Podolska

Department of Contemporary Architecture, Theory and Methodology of Design, Faculty of
Civil Engineering and Architecture, West Pomeranian University of Technology in Szczecin

PUBLIC ENGAGEMENT IN THE DESIGN PROCESS FOR LANDSCAPE ARCHITECTURE OBJECTS

ZAAANGAŻOWANIE SPOŁECZNE W PROCESIE PROJEKTOWANIA OBIEKTÓW ARCHITEKTURY KRAJOBRAZU

Abstract

Although the issues of social involvement in the landscaping design process are essential for creating a people-friendly environment, in Poland, it is still insufficiently identified. The purpose of the article is to present selected examples of various levels and methods of public involvement in the landscaping design process in comparison to the theoretical principles of public participation in the architectural design. The analysed examples were selected in terms of differences in public participation levels, social dialogue tools used and communication with the designer, as well as due to a different functional program. Social involvement should not end at the stage of pre-design consultations. It is necessary to aim at its maintenance and strengthening also at the design stage, and, above all, during implementation and further use of the space.

Keywords: public engagement, public space, landscape architecture

Streszczenie

Problematyka zaangażowania społecznego w procesie projektowania obiektów architektury krajobrazu, choć kluczowa dla kreacji środowiska przyjaznego dla człowieka, w realiach polskich jest jeszcze niedostatecznie rozpoznana. Celem artykułu jest przedstawienie na wybranych przykładach różnego stopnia i sposobu zaangażowania społecznego w proces projektowania obiektów architektury krajobrazu na tle teoretycznych zasad partycypacji społecznej w projektowaniu. Analizowane przykłady zostały dobrane zarówno pod kątem różnic w poziomach partycypacji społecznej, wykorzystywanych narzędzi dialogu społecznego i komunikacji z projektantem, jak i ze względu na różny program funkcjonalny projektowanych obiektów. Zaangażowanie społeczne nie powinno kończyć się na etapie konsultacji przedprojektowych. Dążyć należy do jego podtrzymania również na etapie projektowania, a także realizacji i dalszego użytkowania przestrzeni.

Słowa kluczowe: partycypacja społeczna, przestrzeń publiczna, architektura krajobrazu

1. Introduction

The issues of public engagement in the design process for landscape architecture objects, even if crucial for the creation of a human-friendly environment, in Polish reality, are still insufficiently researched. In the effect, the accomplishment of theoretical public participation rules encounters large practical difficulties. What seems problematic is not only the choice of type of participation, but also tools connected therewith. All too often the public participation is understood and accomplished in form of informing the community about planned actions, but without granting it a possibility to shape the final decision. Still, public participation in architecture can concern both the active participation in decision making and an active participation in the realisation of space. Properly used tools, which are adjusted to the scale of action and its place, prove extremely important for effective design and investment processes.

2. Aim and methods

The aim of the paper is to present, with use of selected examples, the different degree and ways for social engagement in the process of designing objects of landscape architecture confronted with public participation in theory. The analysis was based on examples of designs by the Department of Landscape Design of the West Pomeranian University of Technology in Szczecin, concerning areas located in the city of Szczecin.

The analysed examples were chosen both because of their differences in levels of social participation, and in the tools utilised for social dialogue and communication with the designer, and also because of their different functional programs. The examples were chosen due to the fact that they were all small objects of landscape architecture that strongly focus the local communities around them. Designs of integrative spaces for local communities utilise both environmental and educational themes, and also references to the identity of the place that the inhabitants identify themselves with.

The analysis included:

- ▶ the course of the design process in the decision-making process,
- ▶ the mode in which the local community was engaged,
- ▶ elements utilised in the design that enhance the degree of social involvement.

3. Public participation in theory

The notion of public participation denotes the participation of citizens in managing public issues of their respective communities. Engagement in the issues of the community is made possible after meeting some conditions, which include: identification of the most substantial problems, learning the inhabitants' preferences regarding the way they should be solved and their will to participate, selecting leaders, explaining ways for their solution and gaining acceptance of the inhabitants for the adopted solutions, inclusion of the inhabitants in the

action [5, p. 41]. There are 5 classic tools for public participation: elections, referenda, social consultations, administrative proceedings and direct action.

In the context of shaping space, public participation plays the role of a dialogue between the architect and user, but also of the integrative element for the local community. "Planning spaces are this area, in which the presence of direct democracy (e.g. social consultations) is particularly justified as setting zoning regulations concerns vital interests of both the whole local community and its respective member" [10, p. 43]. In fact, it seems that public participation can be accomplished with different methods and with different degrees of engagement of the community. Participation of inhabitants in the spatial planning process is linked with the mental or social need to participate in the creation of one's own environment and need for active stance towards it. If this need cannot be satisfied, it becomes the source of socially negative feelings, such as indifference and frustration [1, pp. 5–50] that can be further deepened.

Carole Pateman indicates three main types of participation that differ in the degree they engage the participants: the pseudo-participation (only individuals with whom the actions are consulted are engaged), partial participation (the community engages in the decision-making process but only small portion thereof influences its final shape) and full participation (every member of the community has equal influence on the outcome of decision making) [6]. Jules Pretty quotes a more specific typology, listing as many as eight types of participation [10]. These are:

- ▶ manipulative participation – an apparent action that only engages representatives of the community who have no actual influence on the decisions made;
- ▶ passive manipulation – the people are informed about planned action, but their opinion is not taken into account;
- ▶ participation for information sharing – exchange of information depends on professionals, participation in decision making takes place by answering questions (e.g. in surveys) without the possibility to influence the course of the decision-making process;
- ▶ participation through asking questions – engagement takes the form of participation in a survey prepared by the external company that is used to define the problems and solutions, with the possibility of altering the solutions in the light of answers received. The opinions gathered must be included in the decision-making process;
- ▶ participation by material benefits – participation is linked with particular benefits, e.g. the inhabitant engage in works in exchange for food;
- ▶ functional participation – participation in the framework of a group formed to set the goals for action;
- ▶ interactive participation – participation includes plan analysis and formation of new ideas, frequently with the use of interdisciplinary work methods, and the group formed has an influence on the final decisions;
- ▶ self-mobilisation – participation in order to trigger change, in form of independent initiative of inhabitants acting collectively.

The methods used within public participation include mostly different types of workshops that engage local communities, sometimes taking the form of a series of events, during



Table 1. Selected social dialogue tools [13]

Dialogue tools	Disadvantages of the tool	Advantages of the tool
Drawing workshops	<ul style="list-style-type: none"> – no possibility to acquire ready solutions (children are unable to design that space) 	<ul style="list-style-type: none"> – group integration, engagement of children, – building emotional relations with space, feeling of being its host, <ul style="list-style-type: none"> – designer engagement the feeling of the purposefulness of the task acquired by the designer
Direct interview with teachers and management	<ul style="list-style-type: none"> – repetition of well-known schemes and functional solutions 	<ul style="list-style-type: none"> – discussion possibility choice of optimal solution through negotiation, development of the joint concept, – increase of designer’s awareness of user needs
Survey based on open-ended questions	<ul style="list-style-type: none"> – lack of possibility to develop the common concept through negotiations, – the threat of omission or simplification of information already at survey preparation stage and later in data processing, – suggesting answers 	<ul style="list-style-type: none"> – acquiring information about user needs, – increase of designer’s awareness of user needs
Direct conversations with residents	<ul style="list-style-type: none"> – lack of possibility to develop common concept through negotiations 	<ul style="list-style-type: none"> – possibility to add details to data acquired in the survey, – increase of designer’s awareness of user needs, – getting to know the future users of space better
Survey based on close-ended questions	<ul style="list-style-type: none"> – lack of possibility to develop the common concept through negotiations, – threat of omission or simplification of information already at survey preparation stage and later in data processing, – suggesting ready answers, – limited possibility of leaning future users 	<ul style="list-style-type: none"> – increase of designer’s awareness of user needs

which the inhabitants are familiarised with the design criteria. Forming group unity is also an important element of workshops [9]. “The choice of an appropriate dialogue tool still requires the knowledge of its advantages and limitations” (Tab. 1.) [13].

Unfortunately, in the context of spatial planning, social participates frequently take the form of participation of the community in making decisions about the space through information only. When it is lead in different ranges and scales of architecture [12], it can still concern active participation in decisions and execution. In this context, social participation becomes one of the ways the idea of homeliness can be accomplished [7, 8] and also positive

change – both in a special and in a social context, enhancing the identification of users with space and influencing the individualisation of architecture [5].

4. Social participation in Poland

In Polish law, the regulations pertaining to a social participation in the processes of shaping public spaces are included in the Act of 27 March 2003 on spatial planning and management of space [14]. These procedures concern most of all:

- ▶ formal applications for zoning studies and development directions of the municipality and local zoning regulations at the beginning of the procedure for making its draft (or draft of its changes) by an executive authority of the municipality (art. 11 p. 1 of the aforesaid Act – in the case of study, art. 17 p. 1 in the case of Local Zoning Regulations) and their consideration by those authorities (art. 11 p. 3 – study, art. 17 p. 3 Local Zoning Regulations [LZR]);
- ▶ communicating the new LZR study, after design, consideration and agreements to the local community by making it accessible and opening public discussion on the solutions adopted in the draft municipal study or draft LZR (art. 11 p. 10 of the aforesaid Act in the case of study & art. 17 p. 10 in the case of LZR);
- ▶ making notes on the draft study and LZR (art. 11 p. 11 of the aforesaid Act in the case of a study and art. 17 p. 11 – LZR) and their consideration by municipal authority in the final stage of planning procedure (art. 11 p. 12 for the study and art. 17. P. 12 for the LZR).

Thus, the residents have the possibility to participate in area planning of the community by submitting notes and participating in public discussions on different stages of preparation of Local Zoning and the Study of Conditions and Directions for Zoning Management. Still, as Katarzyna Kikosicka observes, it is worth to consider that the law does not state how the results of public discussion are to influence the zoning design, and the main problems connected with the participation of a community in zoning development include a lack of sufficient information about start of work on local zoning, and the form of planning documentation that is difficult to understand for citizens [3].

The quite precise formulation, in Polish law, of principles of social consultations at the level of zoning regulations of the community, is not continued on the level of designing individual public spaces, including landscape architecture objects. Social consultations depend on the good will of the investor (most frequently the commune, less frequently a private investor) and are currently rather rare. Unfortunately, in Poland, we still observe significant aversion and scepticism towards the participation of the society in design processes, which frequently is the consequence of the concept of limitation of designer's freedoms [8].

Among the examples of good practices that engage the local community in the decision-making process pertaining to space, we can list actions by Pracownia Przestrzeni Publicznej – Institute of Design Kielce during their work on City Salon in the Kielce Market Square. In 2013, the Institute of Design organised meetings with Kielce residents, asking them to choose functions that should appear in this place during the so-called “civic game” that utilised – to

some extent – the “Planning for real” consultation method¹. As one of the designers, Michał Gdak, states: “They could participate in three different ways. First of all, we placed a mock-up of the market square with empty spaces that the residents could fill with pictograms representing the respective functions and we then analysed which of them were most frequently repeating. Then, we gave them the possibility to arrange the space on their own, with use of wooden pawns that represented the respective functions and activities. Thirdly, the residents could also cast cards with the desired function in a collection box installed in the centre of the town. We also gave them the possibility of designing their own functions and activities that they wish to use in the market square. Thanks to that information acquired from the inhabitants, we formed project assumptions that included one main theme in form of re-introduction of greenery in the Market space, but also activities such as: sitting, sipping, surfing, reading, playing, talking, walking, meeting” [After: 15]. The tool applied in this case for participation through asking was given an attractive form of a city game, with its results included in the design stage, and today, we can already confirm that it enabled the “taming” of the public space that was, to that date, dehumanised, and now, the inhabitants are keen to spend their time in it.

5. Course of the design processes with different social engagement levels on selected examples

5.1. School garden at the Public Primary School in Głębokie in Szczecin

In the case of objects with a strictly defined and limited number of users, good communication between the designer and future users is possible, and can be based on information gathered and their needs, and inducing the will to participate in the construction of design and cooperation.

The Public Primary School in Głębokie was launched in 1948. First, it was located in a residential building. But it was not long that the Ministry of Education granted funds from its experimental schools fund to construct a building to house the school. The school was designed by architect Janowski, and patterned on the model of Swedish schools, constructed as a single level building adjusted in its scale to the detached homes forming its surrounding. The new building, completed and launched in 1960, was quite a sensation with its pavilion design. Each of the classrooms was built as a pavilion housing a classroom, dressing room and classroom garden (Fig. 1). Both its location (villa residential area, the proximity of a forest) and architectural solutions made it possible for the school to uniquely accomplish the idea of integrating school that is accessible to everyone and promotes ecological knowledge.

¹ „Planning for real” is a method for social consultations that utilizes meetings during which participants, provided with appropriate materials, prepare large, 3D spatial models of given area. Participants use the ready model to put cards that present local problems that require solving (e.g. the accessible infrastructure, healthcare, public transport, etc.) on it. When putting their cards on the model the inhabitants should at the same time state a suggestion concerning the solutions connected with the respective place. Then, in small groups, priority actions are chosen and on their basis, the plans of action are drawn [After: 14].



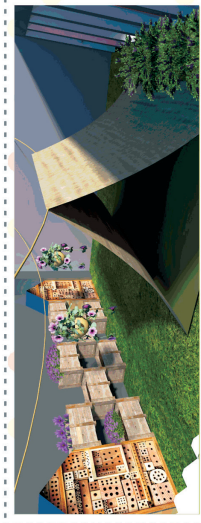
Fig. 1. The Public Primary School in Głębokie, Szczecin (photo by M. Czałczyńska-Podolska)

The area surrounding the school and the class gardens became probing grounds for development of the concept of development of garden space that would best fit the aims of the school: integration and ecological education. Local, communal character of school allowed for pre-design consultations with teachers, meetings with parents and observations of children, linked with drawing workshops. Drawing workshops, during which the children have drawn works concerning the school grounds enabled gathering information about the development elements that were important for children, with plants, especially the species that were known and liked by children due to their ludic properties, proving the most important ones. Preparation of own visions became a major activating factor for children, encouraging them to cooperate in the completion of class gardens and works in the schoolyard. Meeting with parents, which was used to present the variants for development of class gardens, supported with information about the children's vision, formed the basis to include parents in organisational end execution works. Awareness of purpose of work (class gardens are formed for a set group of children who will use them for several years) increased the chances of success of the investment.

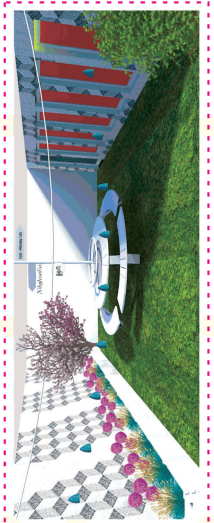
What was formed were integrating gardens that differed in form and character, used didactic elements, mainly from the field of ecology and development and were formed with the close participation of children, forming the pride of classes and the whole school community and an element that builds and integrates the school community (Fig. 2). Engagement of children and parents in decision-making and execution also contributed to the development of class communities (Fig. 3). The gardens make the school stand out from others, forming an important aspect of its identity, and the edible plants used in them additionally enhance the effect of homeliness of space.

Kolorowy świat wyobraźni. Koncept zagospodarowania ogrodu przyklasowcyh

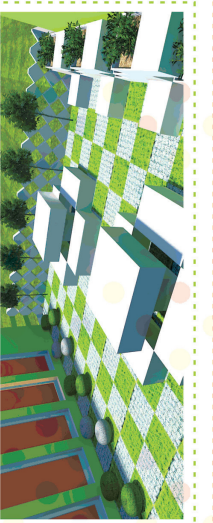
W tym celu, wyobraźni dzieci, w projekcie przedstawiono detaliczne koncepcje kształtowania ogrodu. Dzięki kolorowym płytom i kwiatom, który dzieci mogą złożyć, powstanie im na deskach oddechanie zmysłu do sztuki i sztuki, dominując kolorystyka jest ład. Słonecznik pomysłowe są takie, zmieniają formę, w tym celu, aby dzieci mogły zobaczyć, jak zmieniają się kształty i kształty w otoczeniu. Dzieci mogą być zaangażowane do różnych form bajkowych, takich jak: "Hogwart", "Narnia", "Władcy Pierścieni", "Władcy Smoków", "Złoty Kominek", "Hogwart", "Narnia", "Władcy Pierścieni", "Władcy Smoków", "Złoty Kominek", "Hogwart", "Narnia", "Władcy Pierścieni", "Władcy Smoków", "Złoty Kominek".



Projekt oparty został na modernizacyjnej formie nawiązującej do bryły budynku. Prosta kubatura przetrzeźnia, łącząca w sobie proste elementy geometryczne, które zostały wyodrębnione z otoczenia. W tym celu, aby dzieci mogły zobaczyć, jak zmieniają się kształty i kształty w otoczeniu. Dzieci mogą być zaangażowane do różnych form bajkowych, takich jak: "Hogwart", "Narnia", "Władcy Pierścieni", "Władcy Smoków", "Złoty Kominek", "Hogwart", "Narnia", "Władcy Pierścieni", "Władcy Smoków", "Złoty Kominek".



W tym celu, wyobraźni dzieci, w projekcie przedstawiono detaliczne koncepcje kształtowania ogrodu. Dzięki kolorowym płytom i kwiatom, który dzieci mogą złożyć, powstanie im na deskach oddechanie zmysłu do sztuki i sztuki, dominując kolorystyka jest ład. Słonecznik pomysłowe są takie, zmieniają formę, w tym celu, aby dzieci mogły zobaczyć, jak zmieniają się kształty i kształty w otoczeniu. Dzieci mogą być zaangażowane do różnych form bajkowych, takich jak: "Hogwart", "Narnia", "Władcy Pierścieni", "Władcy Smoków", "Złoty Kominek", "Hogwart", "Narnia", "Władcy Pierścieni", "Władcy Smoków", "Złoty Kominek".



W tym celu, wyobraźni dzieci, w projekcie przedstawiono detaliczne koncepcje kształtowania ogrodu. Dzięki kolorowym płytom i kwiatom, który dzieci mogą złożyć, powstanie im na deskach oddechanie zmysłu do sztuki i sztuki, dominując kolorystyka jest ład. Słonecznik pomysłowe są takie, zmieniają formę, w tym celu, aby dzieci mogły zobaczyć, jak zmieniają się kształty i kształty w otoczeniu. Dzieci mogą być zaangażowane do różnych form bajkowych, takich jak: "Hogwart", "Narnia", "Władcy Pierścieni", "Władcy Smoków", "Złoty Kominek", "Hogwart", "Narnia", "Władcy Pierścieni", "Władcy Smoków", "Złoty Kominek".



Agata Chochol, Kinga Dłabańska – koncepcja zagospodarowania terenu Publicznej Szkoły Podstawowej nr 21 w Szczecinie, w oparciu o idee outdoor education – 2017/2018 – Promotor: dr inż. arch. Magdalena Czajczyńska-Podolska

Fig. 2. Concept for development of class gardens at the Public Primary School in Głębokie – project completed by students of landscape architecture under the supervision of M. Czajczyńska-Podolska



Fig. 3. Gardening works in one of class gardens (photo by M. Czalczyńska-Podolska)

5.2. The Wojtek Bear Square in Szczecin

In the case of larger communities that include whole housing estates, the design-development process is linked with a selection of clear leaders of the enterprise, most frequently its inventors. The design-development process requires both working with the group of mostly interested people, and use of such methods of dialogue that will enable participation and engagement of the broader community.

Representatives of Estate Board of the Krzekowo-Bezrzecze Estate came with the initiative of new development for a square located near the crossing of ulica Żołnierska and ulica Żyzna, inspired by Wojtek the Bear, who served with soldiers of the gen. Anders army all the way from Iran, through the battle of Monte Cassino, up to Scotland. The tradition of the place was the source of the idea. The square is close to ul. Żołnierska [“Soldiers St.”] and ul. Gen. Sosabowskiego, and the NATO Corps is also based nearby. Thus, the statue of the bear was to become the characteristic element of development and the military themes were used both as elements of small architecture, playground devices and details. On the basis of several consultation meetings, a development concept for the square was developed in form of type of military playground – an educational play area with an integrative character that was to fulfil recreational functions for the whole community of the estate. The prepared design variations were made available on the webpage of estate board, thus making it possible to conduct a plebiscite among the residents. The chosen concept was developed in detail and subjected to further consultations. Among the development details proposed were, among others, an educational path with the statue of Wojtek the Bear, and a field for construction and movement plays, sprinklers,

grass-covered hills resembling trenches, a family meeting ground under an openwork structure and a flower meadow that introducing garden character in the space, also grants its homeliness (Fig. 4.). The whole was in a form that resembles military style (Fig. 5.).



Fig. 4. Development concept for Wojtek the Bear Square – military playground – project for Board of Krzekowo-Bezrzecze Estate in 2014, authors: M. Czalczyńska-Podolska, J. Lubarska, S. Mazurek



Fig. 5. Development concept for Wojtek the Bear Square – military playground – project for Board of Krzekowo-Bezrzecze Estate in 2014, authors: M. Czalczyńska-Podolska, J. Lubarska, S. Mazurek

The concept that appeared on the webpage of the Board, but also on wyborcza.pl News Portal was met mainly with supporting opinions. Unfortunately, one of the communities that neighbours the square directly was against its new development and the meeting of the Board and the protesting community did not end with conclusive decisions. The protesters supported their stand with fear of new users of the square and playground. In the effect, further works on the project were halted, but most probably they will be continued in future.

5.3. Gen. F. S. Sosabowski Square in Szczecin

The idea of the new development of Gen. Sosabowski Square in Szczecin is also an example of grassroots mobilisation, forming an independent initiative of residents active in the Board of Krzekowo-Bezrzecze Estate. For the last ten years, the Board of the Estate was actively pursuing actions that were to commemorate the person of general S. F. Sosabowski. The initiative started in 2007 with the “Honor Generała” [Honor of the General] movie by TVP Szczecin journalist and board member - Joanny Pieciukiewicz. The next action of the Board, commemorating the war hero came in 2008 with the Junior High School no. 27 and the newly built street, forming the extension of Gen. Taczaka, being named after general Sosabowski. There is a close relative of Sosabowski living in Szczecin - Wiesław Użarowski, who also engaged in the ceremonial opening of new street and proposed to commemorate the Polish hero with a monument in form of a rock with a plaque. The monument was placed in a square that to that date was run down, between Łukasińskiego, Bracka and Gen. Sosabowskiego streets.

In 2012 the board members initiated activities that aimed at the creation of a common recreational space for all residents of the estate. They turned to the Department of Landscape Design of ZUT in Szczecin for aid in designing several concepts for the development of that



Fig. 6. Development project for Gen. F.S. Sosabowski Square in Szczecin – project for Board of the Krzekowo-Bezrzecze Estate, 2013, authors: M. Rzeszotarska-Palka, M. Kościńska, J. Lubarska, S. Mazurek

square, which were to form the basis for further consultations with residents. Pursuant to suggestions of board members general Sosabowski was to form an inspiration for the creation of design concepts, and the commemorative rock one of the main compositional elements. Three development concepts were designed: one utilised the motive of a parachute in the design of flowerbeds on the lawn, another one the motive of general's epaulets in the surface engraving. These designs were then presented at a meeting of the Board of Krzekowo-Bezrzecze Estate that was open for residents. Some 40 people participated, and after it was ended, the designs were presented on the Estate's official Facebook profile, on paper, and in the Szczecin edition of *Gazeta Wyborcza*, also in a radio broadcast and an article on the webpage of *Radio Szczecin*. Vote for the best design was conducted in three forms: on Facebook, on the webpage of *Gazeta Wyborcza*, and by e-mails sent directly to the Board. On 9 April 2013 the vote was ended, with a total of 2840 votes cast (according to the register of residents of the Szczecin City Council there are some 4000 residents in the estate). After adding all votes, it turned out that concept no. 3 with a total of 1323 votes was most popular, with concept no. 2 gathering 1193 and 324 votes cast for concept no. 1. On the basis of results of residents votes for the development concept of one third of the total square area a technical design was developed, also subjected to multiple consultations with board members of the Krzekowo-Bezrzecze estate board (Fig. 6).

Another investment-promoting action was the action conducted by the Board together with Junior High School no. 27 in form of planting crocuses in shape of a parachute in the square. Between the end of 2013 and 2014, the obelisk commemorating gen. Sosabowski was



Fig. 7. Gen. F.S. Sosabowski Square in Szczecin, June 2017 (photo by M. Rzeszotarska-Palka)

moved, raised and lighted with the active participation of residents. The parachute ornament made of crocuses was already visible in March 2014 and the following years, up to 2016, when the Estate Board finally managed to gather financial resources (PLN 164 000) and the investment was completed by Zakład Usług Komunalnych in Szczecin (Fig. 7, 8).

The main idea of the completed project was to get a coherent public space, both as a symbol of identity of inhabitants and allowing the organisation of events addressed to all residents of Szczecin, and soldiers of the NATO Corps stationed in nearby barracks, through the introduction of function commemorating general Sosabowski. The investment included:

- ▶ relocation of the existing commemorative monument in a place that allowed its proper exhibition – on the axis of the whole development in the southern part of square, on a new pedestal,
- ▶ locating a paved square around the commemorative stone, as dominant element of development,
- ▶ introduction of the motive of general's epaulets in the square surface with use of concrete blocks and plants adjusted to movement of pedestrians (Irish moss),
- ▶ designing a new footpath along the axis of development – main alley running from the square with the commemorative rock to a small piazza with magnolia surrounded with a stone bank.



Fig. 8. Gen. F.S. Sosabowski Square in Szczecin, June 2017 (photo by M. Rzeszotarska-Palka)

6. Conclusion

The conducted analysis of the course of the design process, engagement mode for the local community in the decision-making process and use of design elements that enhance the degree of social engagement proved that:

- ▶ the social engagement in the design process is crucial for the success of the investment;
- ▶ a particularly beneficial situation takes place when the design processes are initiated by the local community, or the future users;
- ▶ the extent and degree of consolidation of the local community plays an important role in the whole consultation and design processes, influencing their final outcome;
- ▶ appropriate choice of development elements (including plants) may enhance the social engagement both in the design processes and during the subsequent stage of use of space.

Social engagement should not end on the pre-design consultations stage. Efforts should be made to support and enhance it also on the design stage, and most of all during its construction and further use of space. The environmentally friendly and educational elements used in the design for school gardens largely contributed to the strengthening of bonds of the school community with space and the feeling of belonging to it. Another factor that positively influences the engagement of local community is a clear definition of the identity of the place, as it was the case of both squares in the Krzekowo-Bezrzecze estate. The grassroots initiative of inhabitants, inspired by the eventful life of General Sosabowski and the figure of Wojtek the Bear contributed to the construction of projects selected by popular vote and thoroughly consulted with the board members.

In the case of works on the design of the Wojtek the Bear square, the recalling of the tradition of location, cooperation with representatives of local residents and consultations in the design of spatial concept proved insufficient to finalise and construct it. The Gen. Sosabowski square project had a different outcome. Determined and long-lasting actions of board members that promoted both the General and the idea of creating a common place for residents in the square named after the general succeeded. A space was thus created that both the residents of the estate and students of the Gen. Sosabowski High School can identify with. The square that was run down to the date became a meeting ground both during official celebrations commemorating the hero with participation of students and the soldiers stationed in the nearby NATO barracks, and also in actions organised by Estate Board such as: common crocus planting, care for seasonal plants in flowerbeds or installations of Christmas decorations.

The school gardens, realised on one part with the cooperation of children and parents, and on the other, utilising the educational and pro-environmental elements in the design, which produce strong bonds with space and between the members of school community, prove that both the decision-making and construction process and the narrative layer of the spatial concept (accomplished by themes that are understood and wanted by users, who identify with them) form an element that is crucial for a successful completion of the development.

References

- [1] Dylewski R., *Głos w dyskusji redakcyjnej*, Człowiek i Środowisko, Vol. 2, 1981.
- [2] Fiedziukiewicz K., *Udział społeczeństwa w planowaniu przestrzennym. Aspekty prawne*, <http://pspe.gridw.pl/?id=11> (access: 11.04.2017).
- [3] Kikosicka K., *Partycypacja społeczności lokalnej w planowaniu przestrzennym (przykład gminy Dąbrowice)*, Acta Universitatis Lodzianis Folia Geographica Socio-Oeconomica, Vol. 16, 2014, 98–113.
- [4] *Komunikacja i partycypacja społeczna*, J. Hausner (ed.), Małopolska Szkoła Administracji Publicznej Akademii Ekonomicznej, Kraków 1999, http://www.msap.uek.krakow.pl/doki/publ/LGPP_Komunikacja_i_partycypacja_spoleczna.pdf (access: 11.04.2017).
- [5] Kosk K., *Social Participation in Residential Architecture as an Instrument for Transforming Both the Architecture and the People who Participate in It*, Procedia Engineering, Vol. 161, 2016, 1468–1475.
- [6] Pateman C., *Participação e Teoria Democrática*, Paz e Terra, Rio de Janeiro 1992.
- [7] Pawłowska K., *Idea i metody partycypacji społecznej w architekturze krajobrazu. Zarządzanie Krajobrazem Kulturowym*, Prace Komisji Krajobrazu Kulturowego, Komisja Krajobrazu Kulturowego PTG, Sosnowiec, Vol. 10, 2008, 617–624.
- [8] Pawłowska K., *Przeciwdziałanie konfliktom wokół ochrony i kształtowania krajobrazu. Partycypacja społeczna, debata publiczna, negocjacje*, Wydawnictwo PK, Kraków 2008.
- [9] Portschy S., *How To Design Together? – Community Architecture* <http://www.fulbright.web2.vhost.hu/book6/portschyszabolcs.pdf> (access: 11.04.2017).
- [10] Pretty J., *Participatory learning for sustainable agriculture*, World Development, Vol. 23 (8), 1995, 1247–1263.
- [11] Siemiński W., *Cele i zasady partycypacji społecznej w planowaniu przestrzennym – przegląd literatury*, „Człowiek i Środowisko”, Vol. 31, 2007, pp. 37–59, https://www.igpim.pl/publikacje/str07_1-2/Siemiński.pdf.
- [12] Skalski K., *Humanistyka i projektowanie*, Arkady, Warszawa 1990.
- [13] Rzeszotarska-Pałka M., Czalczyńska-Podolska M., *Social dialogue – its implications in teaching*, [in:] *Creation/Reaction*, D. Goodwin, B. Kotzen, E. Wall, T. Waterman (ed.), University of Greenwich, London 2017, 1049–1064.
- [14] Ustawa z dn. 27 marca 2003 r. o planowaniu i zagospodarowaniu przestrzennym (Dz.U. Nr 80, poz. 717).
- [15] Interview with M. Gdak, <http://projekt-wbz.com.pl/a/169,ozywic-miejska-pustynie-rozmowa-z-michalem-gdakiem-wspolautorem-miejskiego-salonu-w-kielcach> (access: 11.04.2017).
- [16] Planning for real, <http://urbnews.pl/planning-real-wazne-narzedzie-w-planowaniu-na-szczebli-lokalnym> (access: 11.04.2017).

Elżbieta Waszczyszyn (ela.waszczyszyn@pk.edu.pl)

Institute of History of Architecture and Monument Preservation, Faculty of Architecture,
Cracow University of Technology

THE CHILDREN'S UNIVERSITY HOSPITAL IN CRACOW. HISTORY AND ARCHITECTURE

UNIWERSYTECKI SZPITAL DZIECIĘCY W KRAKOWIE. HISTORIA I ARCHITEKTURA

Abstract

In the common consciousness, the heritage of architecture and urban planning from the period of the People's Republic of Poland is perceived as a legacy of the totalitarian political system, and consequently underestimated and unprotected. One of the interesting buildings coming from this period, the beginnings of which reach back to the 1960s, is the Children's University Hospital in Cracow. In the process of its construction innovative – at the time – architectural and material-related solutions were adopted. The spatial and aesthetic values of the hospital have been preserved to this day, despite its degradation, which has been progressing for some time now and which is to be stopped by renovation and modernisation works undertaken in the hospital. It testifies to its timeless cultural value, both in terms of its formal solutions and materials used, and in terms of its conceptual contents, which are the output of and testimony to the 'difficult' epoch when the hospital was erected.

Keywords: hospital architecture, children's university hospital

Streszczenie

Dziedzictwo architektury i urbanistyki z okresu trwania Polskiej Republiki Ludowej w powszechnej świadomości postrzegane jest jako spuścizna totalitarnego systemu politycznego, a co za tym idzie jest często niedoceniane i nieobjęte ochroną. Do ciekawych obiektów, pochodzących z tego okresu, którego początki sięgają lat 60. XX wieku, należy Uniwersytecki Szpital Dziecięcy w Krakowie. W trakcie jego budowy zastosowane zostały nowatorskie, jak na ówczesne czasy, rozwiązania architektoniczne i materiałowe. Do dzisiaj zachowane zostały walory przestrzenne i estetyczne szpitala pomimo postępującej od pewnego czasu jego degradacji, której przeciwdziałać zaczęły podejmowane ostatnio na jego terenie prace renowacyjne i modernizacyjne. Dowodzi to jego ponadczasowej wartości kulturowej, tak w sensie formalno-materiałowym, jak i treści ideowej, będącej dorobkiem i świadectwem „trudnej” epoki, w której szpital powstał.

Słowa kluczowe: architektura szpitalna, uniwersytecki szpital dziecięcy

1. Introduction

In the common social consciousness, Polish architecture and urban planning from the period of 1945–1989 are perceived as the legacy of the People's Republic of Poland, a result of the influence of the ideology of the totalitarian political system, and consequently has been – and still is – often underestimated and unprotected at all. Irrespective of its difficult origins, these valuable buildings should be part of the resources of the Polish cultural heritage and should be covered with relevant protection.

The complex of the Children's University Hospital in Cracow is a valuable structure the beginnings of which reach back to the 1960s. The spatial and aesthetic values of the hospital have been preserved to date, and it still extends its structures and remains an important medical educational, and research centre [1, p. 786; 13, p. 326]. Today, it has nearly 500 beds and it serves children from Malopolska region and from adjacent provinces. Patients from infants to 18-year-olds are treated in 24 wards. The hospital offers treatment in all children's paediatric and surgical specialities, thanks to which it secures full comprehensive treatment. The facility has an entire range of diagnostic and rehabilitation opportunities, as well.

The oldest part of the hospital was designed by Władysław Poray-Biernacki [1, p. 786; 13, p.326], an American architect of Polish descent, and is a very interesting example of the American functionalism from the 1960s imported to the Polish architecture and urban planning. One encounters here innovative – at the time – architectural, spatial, and material-related solutions. For this reason, the protection should cover the initial form and elevations of the facility, as well as its urban composition.

With time, however, the hospital started to slowly degrade, as a result of its technical and functional wear and tear. Constantly undertaken renovations, especially the extensive programme of modernisation of the entire hospital complex implemented in recent years, aim to compensate for this situation. The spatial and aesthetic values of the hospital have been preserved to date. This fact testifies to its cultural value, both in terms of its form and materials used, and in the conceptual contents, which are an output of and the testimony to the epoch when the hospital was erected. Its architecture attempts to combine arts and pure science, but at the same time, it is a place where the reality of illness tries to encounter a healing force of the hospital space.

One question remains: is it a complex which is sufficiently adjusted to the needs of contemporary hospital management? Modernisations forced by the development of medicine consume enormous costs. One could repeat after an English architect, John Weeks: "Hospitals built in the current way are frighteningly long-lasting..." [27; 28]. The durability of a building does not secure the unchangeability of the way it will be used. Structures with rigid spatial layouts are hard to adapt to changing functional requirements [19, p. 253–259]. A hospital, and a children's hospital in particular, should constitute a combination of the state-of-the-art medical technology and the best care possible over a young patient, creating a truly healing space, where even staying in it should foster healing.

2. Historical background

2.1. The first seat of the Children's University in Cracow

Cracow is a city where, as early as in the 18th century, practical classes in paediatrics for future physicians were held at the ward of foundlings of St. Lazarus Hospital at Kopernika street. In 1835, the Obstetrics Institute, under the supervision of Józef Kwaśniewski, was established. In the second half of the 19th century, nearly every single field of medical sciences had developed its contemporary foundations. Their development brought about the number of new chairs and departments of Collegium Medicum of the Jagiellonian University [26; 18, p. 104–105; 21, p. 480]. Amongst them, one could also mention the Chair of Paediatrics – separated from the Chair of Obstetrics and Gynaecology – which on the basis of an agreement signed between the Ministry of Enlightenment and the Provincial Department established a separate Paediatric Clinic in 1873. Since the very moment of its establishment, located in very modest premises at St. Lazarus Hospital at Wesoła street, it had to struggle with considerable difficulties referring to the premises allocated to it [26; 18, p. 104–105; 21, p. 480; 7, p. 171].

The need to create a new, specialist children's hospital in Cracow, which would also play the role of a university hospital, had been advocated for a long time then by the then head of the clinic, Professor Maciej Jakubowski [8, p. 45–52]. Thanks to his earnest efforts, on 1 May 1876 an edifice of St. Louis Hospital was put into use on the corner of Kopernika street and Krzyżowa street (today Strzelecka street) [8, p. 172; 11, p. 4; 9, p. 9; 26].

The new edifice was a foundation of the Children's Hospital Care Society in Cracow, established in 1872 upon the initiative of Jakubowski and Princess Marcelina Czartoryska, née Radziwiłł [11, p. 4; 26]. Over a relatively short time the Society collected a considerable amount of money, reaching the sum of 40,000 guildens, and for 9,000 guildens purchased a plot of land in the suburbs called Wesoła, opposite St. Lazarus hospital, and quite close to the other university clinics [10, p. 27; 26]. Right after the transaction, the Society turned to several architects with a request to draw up sketches to the design of the hospital. A work by Vincent Schell was regarded as the most suitable one. Having introduced some minor changes, Antoni Łuszczkiewicz [9, p. 9; 26], a Cracow-based architect, prepared detailed plans.

Right after the opening of the hospital, it turned out that halls for contagiously ill children had not been properly isolated and there was a threat that infections could spread within the hospital. It was decided to solve this problem by building a new pavilion only for patients suffering from infectious diseases. Professor Jakubowski proposed to build a detached, one-floor building within the territory of the hospital garden. He even submitted sketches prepared by Antoni Łuszczkiewicz. The committee of the charity society decided, however, to build another, two-floor pavilion as an addition to the existing one. The preparation of a detailed design was entrusted to Tomasz Pryliński. The building was erected in the years 1881–1882 and consumed the amount of 35,610 guildens [9, p.11; 10, p. 29; 26].

Since the very beginning of the existence of the hospital, its director was its co-founder and the then head of the Chair of Paediatrics, Professor Jakubowski. Wishing to secure the



stable and independent future of the Clinic of Paediatrics, as early as in 1876, he drew up a draft agreement between the Imperial and Royal High Government and the Children's Hospital Care Society in Cracow. According to the contract two patients' rooms (each with 10 beds), one observational room with 4 beds, one lecture hall with benches arranged like in an amphitheatre for ca. 40–50 students, and an office for the professor and accommodation for the assistant were separated in the edifice of St. Louis Hospital for the purposes of the clinic [7, p. 172–173; 26]. The contract stipulating the aforementioned provisions, supported by the university authorities, as well as by the then Ministry of Education, was formally approved only in 1896 [7, p. 172–173; 26; 20, p. 68].

During the first 25 years of its existence, as funds were flowing, St. Louis Hospital was being extended and modernised according to the current possibilities. The most difficult times were brought by the World War II, which caused the biggest losses in people and assets. Within its premises, there were structures and appliances requiring urgent renovation and modernisation. Nevertheless, the general renovation of the hospital was carried out only in late 1950. At the time, besides the 1st Paediatric Clinic in St. Louis Hospital, Cracow had paediatric wards in the Żeromski Hospital in Nowa Huta and in the Narutowicz Hospital, but the demand for paediatric beds was constantly growing [24, p. 23]. During the renovation of St. Louis Hospital held in the years 1957–1959, the 1st Paediatric Clinic, which had been located there until then, was transferred to the premises of a former municipal creche and pre-school at Powstańców Warszawy avenue. In this location, the 2nd Paediatric Clinic started to emerge, with Professor Włodzimierz Mikułowski as its first head [24, p. 23]. After the renovation, both clinics returned to the buildings of St. Lazarus Hospital, where a hospital school for sick children was established, as well.

2.2. Current seat of the Children's University Hospital in Cracow

There was a growing need for Cracow to administer one large, professional clinical centre for sick children. This case was supported by the Polish community in America, especially by our countryman, American architect, Władysław Poray – Biernacki, who in 1958 convinced a Member of the House of Representatives, Congressman Clement J. Zablocki, and two democratic Senators, Hubert H. Humphrey from Minnesota and Jacob K. Javits from New York, to the idea of building a state-of-the-art paediatric hospital in Cracow [24, p. 23].

The first step was to amend the act that was in force at the time, the so-called *Battle Act of 1950*, which forbade the Government of the United States to spend any funds on investments in the communist part of Europe, including Poland. Simultaneously, an organisational committee was established - the American Research Hospital in Cracow in Poland [17], which got engaged in the collection of funds amongst Polish communities in the United States and Canada. At that time, the Congress of the United States of America also passed a resolution on granting a donation for the purpose of building a hospital in Cracow [17].

The architect of Polish descent referred to above, Władysław Poray – Biernacki, worked on the design of the hospital. At first, he was employed by a design studio operating on the American market, Howard T. Fisher & Associates [The company was set up by Howard T. Fisher

in 1943 r. and operated until 1965]. With time, since 1955, he started to run his own designing activities, also within the scheme of his own business called Biernacki-Poray & Associates [3].

The construction works were commenced in September 1961. Before they were completed, on 1 April 1965, the Institute of Paediatrics of the Medical Academy in Cracow was established, which was to be the beneficiary for the hospital complex in Prokocim [25, p. 48]. The construction process lasted 4 years (1962–1965) and it brought about 6 interlinked buildings, which as of the opening day had the floor area of ca. 100 thousand square metres, administered 312 beds, and an operating wing with 5 operating theatres. One three-floor building with a circular floor plan housed wards of the 1st and 2nd Paediatric Clinics, the second three-floor central building on a rectangular floor plan housed the Ward of Paediatric Surgery, the Radiology Department, and classrooms, with a lecture hall for ca. 400 people. Another-third – building housed surgeries and consulting rooms, the fourth one – laboratories, the fifth one – animal quarters, and the sixth one – a hospital chapel [24, p. 23]. The opening ceremony of the Institute of Paediatrics in the new facility was held on 11 December 1965, and the first young patient was admitted on 1 February 1966.

The Institute was characterised by modern and original architecture. It referred to the American functionalism – a popular style in the then architecture, urban planning, and applied arts, the main assumption of which was to highlight the role of the function, which dictated technical elements, such as the structure, materials, as well as the aesthetic values, i.e. the form and detail. The design was based on the concept of the so-called pure architecture – free from excessive subjectivism and traditionalism, the goal of which was most of all to satisfy the needs of a sick child and to blur the line between the art of architecture and engineering, which gave positive results in building hospitals. State-of-the-art materials and technical solutions were applied. We find here large glazing almost freed from the structure of the building. The architect freely designed simple elevations and floor plans of the hospital and created interiors which were neutral in colours and modestly furnished. The author paid special attention to the utilitarian problems of the children's hospital and both external and internal architecture was determined by its function. He avoided the use of many colours that were expressed more by the tinge of the materials that were used.

The shape of the building was based primarily on the grid of straight lines (beyond throw so-called rotunda), which was the result of the parameterization in force and technology that was used. Poray-Biernacki created this way a calm and logical architecture, which had to be subordinated to a functional program adapted to a hospital facility.

The building was created in difficult times, which today is usually associated with “PRL's mediocrity” [22, p. 7–8]. At the time, attempts were made to implement modern architecture, which would, however, correspond to the ideals associated with the imagery of a new society propagated by socialist systems. Efforts to create a new architecture realizing the objectives of functional building and at the same time opening to the future, often encountered major systemic and economic difficulties. A customs description of architectural forms that dominated at that time was carried out by Jerzy Hryniewiecki, tying them with the cycles of 5-year economic plans. The years 1961–1965 was “the dominance of urban design projects that take over the theme of the composition and bring the individual buildings to the



simplest form and purely technical functions, limited by norms and realization capacities” [23, p. 5–34]. Despite the atmosphere of that time, Poray-Biernacki realized the modern hospital facility at world level [22, p. 41–46, 164–170]. However, it must be remembered that his extraordinary ally was constant - substantive and financial – support of the foreign organizational committee – the American Research Hospital in Cracow.

The architect met with recognition both abroad and in Poland. For the design of the Cracow-based hospital, Poray-Biernacki was awarded with the Order of Polonia Restituta, one of the highest orders granted to civilians in Poland [6].

Initially, the location of the hospital was a certain drawback. Situated far from the city centre, it lacked any favourable connections. The tram line terminated in the area of the railway station in Płaszów, and there was no public transport from there whatsoever. With time, the line was extended and it reached Prokocim Stary, but the distance to the hospital was still at least one kilometre. The area was dominated by fields and grasslands, as well as dispersed low-rise residential development. Therefore, at the very beginnings of the existence of the hospital, there was no public transport and no illuminated pavements for pedestrians [30, p. 26–27; 24, p. 24].

Since its establishment, the Institute of Paediatrics functioned not only as a specialised children’s hospital, but also as a state-of-the-art teaching and scientific centre. It developed very dynamically, and as early as in 1971, the construction of another building was launched – it was to house administration and hospital laboratories [25, p. 49]. The new pavilion was put into use in 1974. In the same year, having disbanded the organisational committee, which had played such an important role in obtaining funds for the construction of the hospital, the American government established a foundation: Project HOPE, which ever since then was to manage the funds allocated to its further development and the exchange of medical specialists between the Polish hospital and American centres. One of its first decision was a plan to establish the Children’s Medical Rehabilitation Centre, the construction of which was commenced in 1975. For many subsequent years, thanks to this foundation the entire hospital was provided with e.g. medical equipment and relevant spare parts, as well as medical journals and handbooks for its personnel. The foundation also played a significant role in organising numerous courses and conferences addressed to physicians, nurses, and other healthcare professionals, not only from the Institute of Paediatrics in Cracow, but from the entire territory of Poland [24, p. 23]. With time, individual units of the hospital became more and more specialised. Next clinics and medical teams were gradually separated from the existing structures, such as e.g. the Paediatric Surgery Clinic established in 1967, or the Neonatology Clinic with 38 beds, opened in 1968. The diagnostic possibilities of the hospital were likewise growing constantly. In 1974 a new Electron Microscopy Laboratory came into being, and in 1981 a Lipid Biochemistry and Haematology Laboratory was established [24, p. 23–45]. Simultaneously, efforts were made to improve the quality of the medical care offered in the hospital.

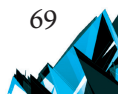
The next stage in the development of the hospital in Prokocim was an agreement signed in 1984 between the Project HOPE Foundation and the Ministry of Health and Social Welfare on the construction of a new Klemens J. Zabłocki Children’s Outpatient Treatment Centre (CALD). The previous one, planned to service 25 thousand appointments per year, was not able to cater for the needs of nearly 100 thousand young patients who came to the existing

outpatient clinic every year. Three years later the construction of a new facility, planned to have the floor area of ca. 19 thousand square metres, was commenced by laying a foundation act on 13 April 1987. The design had been prepared by a Cracow-based architect, Jerzy Urbanik, in collaboration with – among others – Krystyna Łyczakowska, Marta Cierniak, and Krzysztof Dobrowolski [The basic design team included also architects: Adam Mróz, Artur Tatka, Andrzej Zajac]. Przedsiębiorstwo Eksportu Usług Technicznych EXBUD-Kielce was the Main Contractor in the project [25, p. 51].

The postmodern architecture of the new building allowed to observe the way in which theories and solutions developed in the West were confronted with technologies available in the twilight of communism and the dawn of the political transformation, as well as complicated state standards applied in Polish hospital building [12, p. 10–30]. The centre, erected on a rectangular floor plan, had 5 internal patios, adding light to the rooms situated around them. The style was based on concrete blocks, glass slats and structural roof coverings. Staircases were separated from the overall body of the building, in the form of rounded avant-corps crowned with conical roofs and equipped with window openings divided to numerous small, square fields, adding light to the interiors. The main entrance zone was emphasised with a structure the form of which was a reference made to an open portico, crowned with numerous triangular openwork roofs. The largest of them was located above the door opening. Two floors of the building housed the admission room, 35 specialist consulting rooms, 6 diagnostic labs, and a Day Hospital Ward, providing top-quality medical care [12, p. 45–46].

The ceremony opening the first CALD unit, which was the Medical Emergency Treatment Ward, was held on 11 December 1993. It was Poland's first unit which provided 24h-assistance in life-threatening conditions, with full operating and resuscitation facilities, as well as a blood centre and a pharmacy [24, p. 30]. In the same year, the institute along with all the other units of the Medical Academy returned under the wings of the Jagiellonian University, and ever since then was called the Polish-American Institute of Paediatrics, Collegium Medicum. Over subsequent years, it developed further, e.g. in 1995, the construction of a three-floor Transplant Centre was launched.

With time, in 1997, under a decision of the Minister of Health, another change of its organisational structure was introduced. The teaching scientific activities got separated from the medical ones. The Institute became a unit of the Faculty of Medicine, Collegium Medicum, Jagiellonian University, and the medical function was taken over by the Children's Clinical Hospital of Collegium Medicum, Jagiellonian University [24, p. 35; 25, p. 53]. The end of the 1990s also saw the end of the financial support from the Project HOPE Foundation, and the funds for the operation and development of the hospital started to be obtained most of all from the state budget and from independent charities. New wards and medical units, as well as multi-speciality surgical teams, still managed to be established. In 1999, the Radiotherapy Centre was launched, and in 2002 – the Transplant Centre. The Paediatric Burn Centre was established in 2003, the state-of-the-art Neonatal Pathology and Intensive Care Ward started to operate in 2007, and eventually, the new Nutritional Treatment Ward opened in 2008 [25, p. 54–56]. Simultaneously, renovation and modernisation works were undertaken in individual existing units.



In the early 21st century, the university hospital had operated already for 35 years. The constant and growingly fast development of medical sciences brought about the establishment of new, highly-specialised hospital wards and laboratories; nevertheless, the existing complex of buildings, despite all the renovation works undertaken, did not provide a proper location for them and its thorough and comprehensive modernisation was becoming more and more imperative.

3. Issue of protection and modernisation of the Children's University Hospital in Cracow

The Children's University Hospital in Cracow was founded at the time when the standards of treatment, and most of all opportunities offered by medicine, were completely different. In the 21st century, a properly operating children's hospital stands not only for well-equipped operating theatres and rich diagnostic facilities, but also for safe and functional space, creating favourable working conditions for the medical staff and good healing climate for young patients. Today, the hospital should be adjusted to the contemporary technical and sanitary requirements, but also to the EU regulations. Failure to undertake modernisation works or further delay in this respect may lead to gradual degradation of the existing hospital buildings. For this reason, in 2000, the then management of the hospital, led by Maciej Kowalczyk, brought into being the Hospital Revival Action, which got engaged in – without limitations – diversified marketing and information activities, thanks to which the need to modernise the hospital started to be supported by the then authorities and more and more members of the society.

The stock-taking of the unit carried out at the time revealed that the entire building, apart from the newest part of the Children's Outpatient Treatment Centre (CALD), first of all, requires an urgent renovation, and secondly a thorough modernisation [25, p. 2]. Most of all, the entire hospital infrastructure had suffered natural wear and tear: floors, ceilings, walls, water supply, sewerage, electrical and heating installations. The need of changes was dictated not only by their condition, but also by the more and more restrictive standards in force, to which the facility had to be adjusted so as to make it satisfy the requirements of the contemporary academic hospital management. Simultaneously, the demand for next medical units furnished with the most advanced equipment was constantly growing.

Modernisation of the Operating Block and the Intensive Care Ward was the most urgent. Regrettably, there was always the problem of lack of funds, due to which many new concepts and designs were developed, but none of them was ever implemented. Eventually, an opportunity to obtain a financial support from the European Union and the Ministry of Health appeared. The total amount of the subsidy which was obtained at the time was nearly PLN 50M. The first works conducted by Mirbud were launched on 1 September 2009 [25, p. 5]. In two years' time, in September 2011, the unit was put into use. In 2010 funds were obtained (also from the European Union) for the renovation of the Surgery Ward.

Despite the unquestionable success in the gradual modernisation of the hospital, the pace of the works was so slow that all the necessary investments would take an incredibly long time. Hence the management was making efforts so as to draw up a comprehensive schedule

of modernisation of the entire hospital and to secure the relevant funds. A preliminary design of its reconstruction was developed as early as in 2009 by architect Katarzyna Grychowska, employed at the time by Healthcare Design and Investment Office in Katowice. The programme and spatial concept covered the floor area of ca. 75 thousand square metres and the space of ca. 35 thousand cubic metres [4].

After a period of great efforts on the part of the management of the hospital, with considerable support from the municipal and regional authorities and from the Jagiellonian University, as well as Deputies and Senators of the Republic of Poland, on 24 May 2011 the Council of Ministers approved the Programme of Reconstruction of the Children's University Hospital in Cracow for the years 2011-2016 and allocated PLN 201M to this purpose [25, p. 7]. The investment was planned to be implemented in stages, and its most important assumption was to adjust the facility to contemporary standards and technical requirements, as well as to create better conditions for hospitalisation and treatment of sick children.

It was necessary to reorganise the functions of the hospital so as to make sure that the passageways for hospitalised patients would not intersect with those intended for visitors and for patients of numerous consulting centres of the hospital. It was also planned to separate the access road for ambulances from that for vehicles of the staff and carers, as well to build a helipad on the roof of one of the extended buildings, for patients to be transported to emergency departments. Electrical, water supply and sewerage, heating, and air-conditioning installations were reconstructed in compliance with the contemporary requirements. The hospital wards were provided with a new interior design so as to create favourable conditions for sick children and their guardians.

The order of the modernisation works was dictated by the need to maintain the continuity of the hospital's operation. Such a situation always causes a lot of problems in the activities of such a big medical facility. Therefore, the first stage comprised covering the terraces of the main building in order to provide premises for the administration staff. The previous office space was allocated to the extension of the paediatric wards of the hospital. Next, it was planned to create a new Emergency Department along with a helipad and to reorganise the Children's Outpatient Treatment Centre, which after the completion of the modernisation was to be a place occupied by subsequent wards renovated at the time, located in the 'round building', as well as the 'rehabilitation building'. The last works to be performed were the works relating to the so-called street architecture – the hospital garden, pavements for pedestrians, car parks for the staff and patients, and fencing around the entire area.

Already in 2014, the reconstruction of the building housing a central server room, a technical control room, and a transformer station, was completed, thanks to which the hospital got equipped with a modern electrical network. In 2015 the thoroughly modernised Surgical Ward with 69 beds was open. It was equipped with new furniture and specialist medical equipment, and its internal functional organisation was changed, as well. At the same time, the reconstruction and renovation of the lecture hall, classrooms and offices was completed in the oldest part of the hospital. The renovation covered also the Cardiac Surgery Ward and the Dissecting Room. An additional lift shaft was built in the building of the Children's Outpatient Treatment Centre, which in turn was linked with the renovated



Admission Room. The extension and the renovation of the building housing the new Emergency Department with state-of-the-art medical treatment rooms, a waiting room for the patients, and physicians' rooms along with a locked driveway for ambulances and a helipad, was completed. The hospital pharmacy obtained a completely new look – it got equipped with a state-of-the-art automatic drug preparation system, and so did the kitchen, furnished with the most advanced cooking and food distribution system.

The next in line was the reconstruction and extension of one of the oldest hospital buildings, the so-called 'round building', housing paediatric wards of cardiology, gastroenterology, rheumatology and environmental diseases, oncology and haematology, and nutritional treatment, comprising the total of 200 beds for sick children. The facility was also equipped with a new wing, housing auxiliary and social premises. The investment was crowned with a gala opening of the building held on 23 January 2017. One of the most characteristic and valuable elements of the architecture of the hospital received a new life and became more functional and friendlier for the patients and the staff.

Since ca. 2011 to 2017, paediatric wards with nearly 300 beds for patients were renovated, and additional floor surface of ca. 1000 square metres was obtained and allocated to new medical treatment rooms, physicians' rooms, offices, utility rooms, and – most importantly for the patients – play and education rooms for children, as well as rest and refreshment rooms for their guardians, and a new Emergency Department with a helipad. All this was the result of the first stage of the modernisation of the Children's University Hospital in Cracow-Prokocim completed in the beginning of 2017. The entire modernisation is (re)planned to be completed by 2018 [Initially the modernisation of the hospital was to be completed by 2016].

The modernisation of the children's hospital is still in progress. The management succeeded in obtaining another PLN 9M from the EU funds for the launch of a new Traumatology Centre, which is to be one of twelve centres of the kind to be opened in Poland. First, six such centres will be launched, and the one in Cracow will be one of them. For the centre to be opened, it is necessary to have a helipad, the most advanced diagnostic equipment, and trauma teams on duty 24h a day – groups consisting of e.g. an emergency medicine specialist, an anaesthetist, a general surgeon, an orthopaedist, and a neurosurgeon. The hospital in Cracow satisfies these conditions, and the money from the Ministry of Health and the European Union will be used to organise a seat for the new unit and to purchase state-of-the-art medical equipment necessary for it to operate [5].

Thanks to the funds obtained both from the state budget and from the EU, and thanks to the support of many foundations, so far more than a half of the hospital has been modernised, which in fact stands for a new life for this facility. The former hospital buildings no longer satisfied the requirements which have to be fulfilled by such facilities, and due to the more and more rapid development in medical sciences, this specialist university unit, which combines therapeutic functions with those relating to the education of future physicians and medical staff, faced new challenges.

In case of former valuable complexes, which the Children's University Hospital in Cracow undoubtedly is, the revitalisation process is particularly important. The hospital had to be revived by adjusting its oldest buildings to the requirements posed by the contemporary

hospital management and introducing to their interiors modern functional systems and the most advanced equipment. Both these measures called for considerable changes in the spatial structure of the entire hospital. The reconstruction and extension of its hospital buildings could be regarded as successful modernisation projects or simple renovation measures, depending on their scale. However, one question remains: how far can we go in our interference with the structure of the facility so as not to deprive it of its initial value?

Considering the fact that there is no valid precise list of assets of the contemporary culture, comprising all important examples of hospital architecture from the second half of the 20th century, only preservation maintenance seems to be an effective form of protecting these facilities against too far-fetched changes and transformations. On the other hand, the experience of recent years shows that valuable hospital buildings erected after the war – if they want to maintain their therapeutic activities – more and more often are subjected to intensive transformations, which result from necessary comprehensive modernisations. Due to their specific characteristics, hospitals are particularly sensitive to all sorts of ‘renewals’ which are to adjust them to the needs of the contemporary hospital management.

Summing up, valuable hospitals erected in the difficult period of communism in Poland, and the Children’s University Hospital in Cracow among them, as assets of the contemporary culture should be covered with relevant protection the areas of which must be, however, specifically determined. In case of some facilities, such protection could cover only their most valuable elements. In case of complexes of buildings or urban and landscape projects, it is necessary to demarcate limits of the area which is to be protected, although it may turn out that the entire area should be protected instead. In order to make sure that such measures will be effective, it is necessary not only to precisely determine the limits of this protection, but also to support it with a detailed analysis of the values to be protected. The criteria for assessment of these values should be based on the definition of - first - artistic values representing original artistic solutions, creatively addressing contemporary stylistic tendencies in architecture - second - uniqueness and innovativeness in creating architectonic and spatial, or technological solutions - thirdly - novelty in realizing projects - fourthly - historical significance through e.g. an eminent artist or an outstanding document of the epoch [29, p. 7–14; 15, p. 38–50].

Under no circumstances, however, should the existing valuable tissue be destroyed, and all new additions should be combined with it in a collision-free way. It is very important to protect the authenticity of an object when over time it undergoes changes unforeseen by its creator. The Children’s University Hospital in Cracow has specific function and form, was created with specific materials, technologies and construction methods. Preservation of the authenticity is equivalent to the conservation of characteristics comprising the object’s existence. The individual features differ in their significance, and the loss of some of them doesn’t endanger the integrity of the entire object [14, p. 79–90].

Skilfully conducted modernisation measures in Prokocim’s hospital additionally provide its users with an opportunity to get closer to the history of hospitals. They provide a chance to protect a valuable building, to develop a common urban space, and to transform it so as to make it satisfy the needs of the community that uses it. Pragmatism, but also the creativity of such decisions, foster the creation of a balance between the application of the most



advanced material-related and technological solutions, and the need to protect the identity of the building, which is a part of our cultural heritage.

Nowadays hospital architecture should draw inspiration from its long traditions, as well as from the latest achievements of science and technology. On the one hand, it should search for new patterns of shaping space – its functions and forms, and on the other hand, it should adapt and modernise former concepts based on recognised and proven canons. It should constantly follow the transformations in medicine, but at the same time, it should become architecture which is more and more friendly for a sick person, whose sense of well-being fosters the healing process.

Undoubtedly, the atmosphere of the hospital space has its effect on its users. It should be a healing space at all levels: mental, spiritual, and physical, and architecture and interior design influence each of them [16]. Reduction of the patient's sense of alienation and stress, boosting their satisfaction with professional treatment, and promotion of the healthy lifestyle should constitute a priority in designing the new and modernising the existing health architecture, the best example of which are the aforementioned measures undertaken in the Children's University Hospital in Cracow.

References

- [1] Barucki T., *Architektura Polski*, Warszawa 1985.
- [2] *Encyklopedia Krakowa*, Wydawnictwo Naukowe PWN, Warszawa 2000.
- [3] Grimes W., *Władysław O. Biernacki-Poray, 71, Architect* [in:] <http://www.nytimes.com/1995/10/07> (access: 14.04.2017).
- [4] <http://www.atelierkato.pl/referencje.html> (access: 14.04.2017).
- [5] <http://www.dziennikpolski24.pl> (access: 14.04.2017).
- [6] <http://www.nytimes.com/1995/10/07> (access: 14.04.2017).
- [7] Jakubowski M., *Klinika Pediatria*, [in:] *Zbiory i Zakłady Przyrodnicze i Lekarskie Krakowa*, J. Rostański (ed.), Kraków 1881.
- [8] Jakubowski M., *Rzut oka na dzieje, piśmiennictwo i zakłady pediatryczne*, Kraków 1865.
- [9] Jakubowski M., *Szpital Św. Ludwika dla Dzieci w Krakowie. Jego powstanie, rozwój i działalność dziesięciolecie od 1876-1885*, Kraków 1886.
- [10] Jakubowski M., *Kronika Szpitala Św. Ludwika dla Dzieci w Krakowie. Okres 25 lat, od roku 1876 do roku 1900*, Kraków 1900.
- [11] Jakubowski M., *Szpital Św. Ludwika w Krakowie (na łóżek 72)*, Kraków 1877.
- [12] Jencks Ch., *Architektura późnego modernizmu*, Warszawa 1989.
- [13] Fabiański M., Purchla J., *Przewodnik po architekturze Krakowa*, Kraków 2001.
- [14] Kulig A., *Autentyzm dzieł architektury XX wieku – chroniony czy zagrożony?*, *Wiadomości Konserwatorskie*, 49/2017.
- [15] Lewicki J., *Ocena wartości zabytków epoki modernizmu. Przeszłość, teraźniejszość i przyszłość*, *Wiadomości Konserwatorskie*, 49/2017.
- [16] Miłunski F., *Architektura w służbie zdrowia*, *Miejsca*, 2006, (8).

- [17] Morgan T., *The American Research Hospital in Krakow in Poland*, U.S. Govt. Print. Off, 1966.
- [18] Popiela T., Skulimowski M., *Historia wydziału lekarskiego w okresie od repolonizacji Uniwersytetu do odzyskania niepodległości (1864–1918)*, [in:] *Sześćsetlecie Medycyny Krakowskiej*, L. Tochowicza (ed.) T. II, Kraków 1964.
- [19] Porębowicz, S., *Rozrost i elastyczność wewnętrzna rozplanowania jako kryteria nowoczesności projektów szpitali*, Szpitalnictwo Polskie, 1970, No. 14(6).
- [20] Rostański J., *Kraków pod względem lekarsko-przyrodniczym*, Kraków 1900.
- [21] Seyda B., *Dzieje medycyny w zarysie*, Warszawa 1973.
- [22] Springer F., *Żle urodzone. Reportaże o architekturze PRL-u*, Kraków 2012.
- [23] Szafer P.T., *Współczesna architektura polska*, Warszawa 1988.
- [24] *Uniwersytecki Szpital Dziecięcy w Krakowie i Instytut Pediatrii. 50 lat w leczeniu najczęściej chorych dzieci*, M. Oberc, J. Szydłowska, M. Kowalczyk, Kraków 2015.
- [25] *Uniwersytecki Szpital Dziecięcy w Krakowie. Czas odnowy*, publication commemorating the opening of the Central Operating Block with the Central Sterilization Room and the Intensive Care Unit, Kraków 2011.
- [26] Waszczyszyn E., *Zabytkowy szpital w przestrzeni współczesnego miasta. Problemy konserwatorskie*, [in:] *Karta Krakowska 10 lat później*, ed. by Kadłuczka A., Kraków 2011.
- [27] Weeks, J., *Designing for patients care. Educations and research*, World Hospital, 1969, No. 5(4).
- [28] Weeks, J., *Hospital design for growth and change*, World Hospital, 1969, No. 5(1).
- [29] Węclawowicz-Gyurkovich E., *Podsumowanie ogólnopolskiej konferencji „Ochrona dziedzictwa architektury polskiej II połowy XX w.”* Warszawa 16–17 XI, *Wiadomości Konserwatorskie*, 49/2017.
- [30] *Wspomnienia pracowników Instytutu*, [in:] *Kronika XXX-lecia Instytutu*, Z.1, CM UJ, Kraków 1997.





Fig. 1. The Children's University Hospital in Cracow, archival photography of the construction of the hospital, 04.11.1962 (source: Archive of the Children's University Hospital in Cracow)



Fig. 2. The Children's University Hospital in Cracow, archival photograph of the construction of the hospital, 31.05.1965 (source: Archive of the Children's University Hospital in Cracow)



Fig. 3. The Children's University Hospital in Cracow, view of the Children's Outpatient Treatment Centre (CALD) from beginning of 90. and the oldest part of the hospital after renovation and modernization (photo by E. Waszczyszyn, 20.04.2017)



Fig. 4. The Children's University Hospital in Cracow, the entry to the Children's Outpatient Treatment Centre (CALD) (photo by E. Waszczyszyn, 20.04.2017)



Fig. 5. Children's University Hospital in Cracow, view of the oldest part of the hospital after renovation and modernization (photo by E. Waszczyszyn, 20.04.2017)



Fig. 6. Children's University Hospital in Cracow, view of the oldest part of the hospital after renovation and modernization, window detail (photo by E. Waszczyszyn, 20.04.2017)



Fig. 7. Children's University Hospital in Cracow, view of the oldest part of the hospital after renovation and modernization and after covering of the terraces of the main building (photo by E. Waszczyszyn, 20.04.2017)



Fig. 8. Children's University Hospital in Cracow, view of the expanded part of the oldest part of the hospital after renovation and modernization (photo by E. Waszczyszyn, 20.04.2017)



Fig. 9. Children's University Hospital in Cracow, view of the interior of the oldest part of the hospital after renovation and modernization, hospital corridor (photo by E. Waszczyszyn, 20.04.2017)



Fig. 10. Children's University Hospital in Cracow, view of the interior of the oldest part of the hospital after renovation and modernization, rest area for sick children (photo by E. Waszczyszyn, 20.04.2017)

Grzegorz Kurowski

Otmar Vogt (ozvogot@pk.edu.pl)

Jan Ogonowski

Faculty of Chemical Engineering and Technology, Cracow University of Technology

PAINT-DEGRADING MICROORGANISMS

MIKROORGANIZMY ZDOLNE DO DEGRADACJI POWŁOK MALARSKICH

Abstract

The paper presents a literature survey of research on the biodeterioration and biodegradation of paint coatings, as well as study findings on the biodegradation of varnish coatings and agents used for their removal.

Keywords: biodeterioration, biodegradation, paints, microorganisms

Streszczenie

W artykule przedstawiono przegląd literaturowy dotyczący biodeterioracji i biodegradacji powłok malarskich. Przedstawiono również badania dotyczące biodegradacji substancji używanych do usuwania powłok lakierniczych.

Słowa kluczowe: kluczowe: biodeterioracja, biodegradacja, powłoka malarska, mikroorganizmy

1. Introduction

Paint and varnish coatings are used for protective and/or decorative purposes, and their properties largely depend on their ingredients, i.e. on the film-forming substances employed. The latter primarily include macromolecular organic substances or substances that turn into macromolecular compounds by way of reactions that accompany the formation of a varnish layer. These substances also determine the possible uses and applications of paint coatings. Film-forming substances consist of resins and other accessory substances. In order to invest coatings with specific properties a mix of resins is also often used [1].

The literature distinguishes between the processes of biodeterioration and biodegradation. Biodeterioration involves the destruction of an economically important substance by microorganisms. The term is often used to refer to a deterioration in the properties of substances that are normally immune to microbiological attack, such as metals, plastics, medicines, cosmetics, paints and varnishes, fuels, oils, and other objects. Biodeterioration is thus an adverse phenomenon. In contrast, biodegradation is viewed as beneficial. The biodeterioration of varnish coatings caused by microorganisms is often accompanied by the deterioration of their protective and decorative properties. The proliferation of microorganisms on and inside paint may cause it to peel. The growth of fungi, in turn, may lead to discolorations [2, 3].

The microbiological resistance of resins used in varnishes can be measured on the Wasserbauer scale, which assesses the proliferation of microorganisms on a scale from 0 to 3, where 3 indicates robust growth. Varnishes with the index of 2 or more are described as non-resistant to microorganisms. The values of microbiological resistance for selected synthetic resins are as follows: acrylic resin 0.51; phenol formaldehyde resin 0.73; urea-formaldehyde resin 0.17; polyester resin 1.20; epoxy resin 0.19; polyvinyl resin 0.51 [4].

2. Biodeterioration and biodegradation of paint films

2.1. Biodeterioration

Only water-based paints are susceptible to biodeterioration. The cellulose ethers they contain as thickening agents are often attacked by enzymes produced by fungi and bacteria, which enter the paint at the manufacturing stage together with contaminated ingredients [3]. Orehek et al. studied the biodeterioration of carboxymethyl cellulose by *Bacillus subtilis* subsp. *subtilis* NCIB 3610. The bacteria turned the original pseudo-plastic non-Newtonian fluid that contained carboxymethyl cellulose into a Newtonian fluid. In addition, the fluid's viscosity dropped from 10 to 1.4 mPas and the concentration of reducing sugars increased eighteen-fold. *Bacillus subtilis* subsp. *subtilis* NCIB 3610 was also shown to synthesize cellulases that effectively break down chemical bonds in carboxymethyl cellulose; however, it is not able to reutilize the end products of hydrolysis [5].

Obidi et al. [2009] studied the microbiological contamination of painting products. Fresh samples were taken to isolate the following bacterial strains: *Bacillus brevis*, *Bacillus laterosporus*,

Bacillus polymyxa, *Lactobacillus gasseri*, *Lactobacillus brevis*, *Proteus mirabilis*, *Escherichia coli* and fungi *Aspergillus niger*, *Aspergillus flavus*, and *Penicillium citrinum*. Monitored over a period of 10 months, the population of isolated microorganisms increased. For bacteria, the population grew from $1.6 \cdot 10^1$ cfu·cm⁻³ in a fresh sample to $4.7 \cdot 10^5$ cfu·cm⁻³ 10 months later. The corresponding figures for fungi were $1.0 \cdot 10^1$ and $5.5 \cdot 10^3$ cfu·cm⁻³, respectively [6]. In addition, the researchers took samples of water-based paints to isolate a strain of *Pseudomonas aeruginosa* [7].

Shirakawa et al. (2011) studied microorganisms living in four different acrylic paints exposed to atmospheric conditions for seven years in Sao Paulo and the coastal town of Ubatuba in Brazil. They identified bacteria from the cyanobacteria group. The main strains present on the surface of all four types of paint belonged to *Gloeocapsa* and *Scytonema*. In addition, the following genera were detected: *Synechocystis*, *Synechococcus*, *Xenococcus*, *Chroococciopsis*, *Myxosarcina*, *Oscillatoria*, *Leptolyngbya*, *Chlorella*, *Stichococcus*, *Coccioid chlorophyte*, *Klebsormidium*. The researchers also observed certain differences depending on the location from which the samples were obtained. Paint degradation related to the presence of biofilm was shown to be very low, but progressed faster in Ubatuba [8]. Other researchers also isolated cyanobacteria from painted buildings in Latin America [9-11]. *Stenotrophomonas maltophilia* were also detected in water-based acrylic paints [12].

De Souza and Gaylarde (2002) isolated microorganisms from varnished wood panels exposed to atmospheric conditions for two years in South Brazil. The following bacteria were isolated from the samples: *Corynebacterium* sp., *Bacillus* sp., as well as bacteria from the *Enterobacteriaceae* family. Yeasts were also detected: *Rhodotorula lactosa*, *Exophiala jeanselmei*, as well as fungi: *Aureobasidium* sp., *Penicillium* sp., *Paecilomyces* sp. [13].

Pangallo et al. (2015) investigated the biodeterioration of epoxy resins. The most frequently isolated bacteria present on surfaces coated with epoxy resins belonged to the Firmicutes phylum, especially the genus of *Bacillus*. γ -Proteobacteria, and particularly *Pseudomonas*, were also frequently detected [14]. Similar results were obtained by Cappitelli et al. [15], who studied surfaces coated with epoxy resins and isolated Gram-positive bacteria of the *Bacillus*, *Brevibacillus*, *Micrococcus*, *Staphylococcus* and *Kocuria* genus, as well as Gram-negative bacteria from the genus of *Pseudomonas*, *Agrobacterium* and *Ochrobacter* [15]. In addition, researchers [14] identified several species of black yeast: *Friedmanniomyces endolithicus*, *Pseudotaeniolina globosa*, *Phaeococcomyces catenatus* and *Catenulostroma germanicum*.

The authors of paper [16] listed the following species of microorganisms: *Aspergillus versicolor*, *Aspergillus niger*, *Cladosporium herbarum*, *Cladosporium sphaerospermum*, *Engyodontium album*, *Penicillium aurantiogriseum*, *Penicillium chrysogenum*, *Trichoderma longibrachiatum*, *Debaryomyces hansenii*, *Rhodotorula mucilaginoso*, *Bacillus subtilis*, *Bacillus licheniformis* and *Rhodococcus fascians*, all isolated from paints [16].

Bacteria present in paint on stone walls were studied by Gurtner et al. [17]. The researchers analysed mural paintings in the chapel of an Austrian castle and a German church. The following bacteria were isolated: bacteria from the phylum of Actinobacteria, such as *Arthrobacter* sp., *Actinobispora* sp., *Amycolata* sp., *Asiosporangium* sp., *Frankia* sp., *Geodermatophilus* sp., *Nocardioides* sp., *Promicromonospora* sp., *Pseudonocardia* sp., *Rubrobacter* sp., *Streptomonospora* sp., *Saccharopolyspora* sp., *Sphaerobacter* sp. and *Thermocrisium* sp., bacteria from the *Proteobacteria*

phylum, such as *Aquaspirillum* sp., *Chromohalobacter* sp., *Deleya* sp., *Erythrobacter* sp., *Halomonas* sp., *Porphyrobacter* sp., *Pseudomonas* sp., *Rhizobium* sp., *Salmonella* sp. In addition, the researchers isolated bacteria from the genera of *Bacillus*, *Paenibacillus*, *Micrococcus*, *Staphylococcus*, *Methylobacterium*, and *Halomonas* [17]. Similar results were obtained by the authors of articles [18, 19], who studied mural paintings in Spain and Italy. Apart from the phyla mentioned above, researchers who wrote papers [20, 21] also isolated *Cyanobacteria* (*Chlorogloea microcystoides*, *Chroococcus lithophilus*, *Gloeocapsa* spp., *Gloeotheca rupestris*, *Pseudocapsa dubia*). Bacteria from the *Actinobacteria* phylum of the *Pseudonocardia*, *Streptomyces*, *Nocardia*, *Rhodococcus*, *Nocardioides*, *Amycolatopsis*, *Saccharothrix*, *Brevibacterium*, *Microbacterium* genus and *Bacillus*, *B. megaterium* were also isolated in Paleolithic cave paintings in southern Spain [22–24].

Apart from bacteria, wall paintings also contained fungi from the genus of *Cladosporium*, *Penicillium*, *Nectria* and yeast from the genus of *Rhodotorula* [25].

Okunye et al. [26] studied species of fungi capable of causing the degradation of paint. Thirty six samples obtained from various locations in Southeastern Nigeria helped isolate *Rhizopus* and *Aspergillus* spp., which were the most frequently present. Other genera of fungi as: *Absidia*, *Monilia*, *Alternaria*, *Fusarium* and *Penicillium* were also detected. [26]. *Aspergillus*, *Penicillium*, *Botrytis*, *Acrodictys*, *Mucor*, *Absidia*, *Alternaria*, *Cladosporium*, *Rhizopus*, *Cephalosporium*, *Fusarium*, *Helminthosporium*, *Trichoderma*, *Acremonium*, *Pullularia*, *Monilia*, *Epicoccum* and *Udeniomyces* were likewise isolated genera from the painted surfaces investigated in studies [27–29]. Likewise, species of fungi *Aspergillus niger* and *Cladosporium* sp. were isolated from paint coatings [30, 31]. Fungi such as *Aureobasidium* sp. and *Cladosporium* sp. [9, 11, 14, 15, 32–36] were also frequently found. In addition, Lugauskas et al. (2003) identified fungi from the genera of *Aureobasidium pullulans*, *Aspergillus niger*, *Aspergillus versicolor*, *Cladosporium cladosporioides*, *Paecilomyces sulphurellus*, *Trichoderma viride*, *Ulocladium atrum* as capable of degrading polymer materials [37].

The following fungi species: *Acremonium charticola*, *A. strictum*, *A. kiliense*, *Acremonium* spp., *Aspergillus sydowii*, *Aureobasidium pullulans*, *Beauveria* sp., *Cladosporium* sp., *Cladosporium sphaerospermum*, *Chrysosporium* sp., *Engyodontium album*, *Mycelia sterilia*, *Scopulariopsis brevicaulis*, *Verticillium lecanii*, *Verticillium suchlasporium*, *Verticillium* sp. were isolated from walls painted in antiquity [38]. The same species were isolated and described in study [39], which also isolated species from the genera of *Arthrobacter* and *Bacillus* [39]. Researchers studying similar subjects isolated the following genera of fungi and bacteria: *Pseudomonas alcaligenes*, *Nocardia asteroides*, *Arthrobacter* spp., *Bacillus subtilis*, *Pseudomonas aeruginosa*, *Streptococcus* spp., *Pseudomonas fluorescens*, *Micrococcus roseus*, *Bacillus pumilus*, *Streptomyces* spp., *Micrococcus luteus*, *Alternaria alternata*, *Aspergillus versicolor*, *Chaetomium globosum*, *Cladosporium cladosporioides*, *Aspergillus terreus*, *Aspergillus oryzae*, *Aureobasidium* spp., *Penicillium* spp. and *Penicillium stoloniferum* [40].

Studies carried out by Romero-Noguera et al [41, 42] looked into the biodeterioration of turpentine varnishes. Based on their findings, researchers identified the following microorganisms capable of degrading Venetian turpentine coatings: *Chrysonilia sitophila*, *Streptomyces celluloflavus*, *Bacillus amyloliquefaciens*, and *Arthrobacter oxydans* [41]. Moreover they identified microorganisms capable of degrading Sandarac coatings (*Chrysonilia sitophila*, *Penicillium chrysogenum*, *Rhizopus oryzae*, *Mucor indicus*) and Manila Copal coatings

(*Aspergillus niger*, *Aureobasidium pullulans*, *Chrysonilia sitophila*, *Penicillium chrysogenum*, *Pleospora herbarum*, *Rhizopus oryzae*) [42].

Polish studies also looked into the presence of microorganisms in paint coatings. Scientists isolated fungi present in paint inside residential buildings and public utility buildings. The following genera of fungi were isolated in residential buildings: *Acremonium stictum*, *Alternaria alternata*, *Alternaria tenuissima*, *Aspergillus niger*, *Aspergillus versicolor*, *Penicillium cyclopium*, *Penicillium janthinellum*, *Penicillium thomii*, *Penicillium viridicatum*, *Rhizopus nigricans*, *Trichoderma viride*. Researchers also studied fungi in the Registry Office, an office building, a university, a polytechnic, a primary school, a hospital, and a hotel in the Silesian Voivodship. 20 species of fungi were isolated, including: *Acremonium stictum*, *Alternaria alternata*, *Aspergillus niger*, *Aspergillus ustus*, *Aspergillus versicolor*, *Cladosporium cladosporides*, *Chateomium globosum*, *Gliocladium catenulatum*, *Gliocladium roseum*, *Memnoniella echinata*, *Paecilomyces variotii*, *Penicillium albidum*, *Penicillium chrysogenum*, *Penicillium citrinum*, *Penicillium rugulosum*, *Penicillium terrestre*, *Penicillium variabile*, *Scopulariopsis brevicaulis*, *Stachybotrys chartarum*, *Trichoderma viride* [4].

Research was also conducted on fungi present in paint coatings in food industry plants, breweries, dairies, and fruit preserve factories.

Alternaria sp., *Cladosporium herbarum*, *Leptographium* sp., *Phoma violacea*, *Trichothecium* sp., *Trichoderma koningii* were isolated in food industry plants, *Aspergillus flavus*, *Botrytis cinerea*, *Acremonium* sp., *Chalaropsis* sp. in breweries, and *Geotrichum* sp. and *Fusarium* sp. in dairies. Paint in fruit preserve factories was shown to contain the following species of fungi: *Aspergillus nidulans*, *Aspergillus flavus*, *Aspergillus oryzae*, *Aspergillus versicolor*, *Penicillium glabrum*, *Penicillium italicum*, *Penicillium spinulosum*, *Botrytis cinerea*, *Aureobasidium pullulans*, *Cladosporium elatum*, *Cladosporium herbarum* and *Alternaria tenuis*, as well as various species of *Fusarium* sp. In chocolate-processing factories, the following species of fungi were detected: *Mucor hiemalis*, *M. racemosus*, *M. spinosus*, *Rhizopus nigricans*, *Aspergillus ochraceus*, *Aspergillus versicolor*, *Penicillium cyclopium*, *Penicillium expansum*, *Penicillium globrum* and *Penicillium brevicompactum*. The following fungi were isolated in a juice factory: *Cladosporium sphaerospermum*, *Chaetophoma* sp. and *Acremonium strictum* [4].

Table 1. Microorganisms isolated from paint exposed to atmospheric conditions [2, 43]

Microorganism	Oil paint	Emulsion paint	Microorganism	Oil paint	Emulsion paint
Fungi			Bacteria		
<i>Alternaria dianthicola</i>	+	+	<i>Alcaligenes recti</i>	+	
<i>Aureobasidium pullulans</i>	+	+	<i>Alcaligenes</i> sp.	+	
<i>Aspergillus flavus</i>	+	+	<i>Bacillus cereus</i>	+	
<i>Botryodiplodia malorum</i>	+		<i>Bacillus mycoides</i>	+	+
<i>Botrytis cinerea</i>	+		<i>Bacillus sphericus</i>	+	

Microorganism	Oil paint	Emulsion paint	Microorganism	Oil paint	Emulsion paint
<i>Cephalosporium carpogenum</i>	+	+	<i>Bacillus</i> sp.	+	
<i>Cladosorium sphaerospermum</i>	+	+	<i>Flavobacterium invisibile</i>	+	
<i>Cladosporium</i> sp.	+	+	<i>Flavobacterium marinum</i>	+	+
<i>Fusarium flocciferum</i>	+		<i>Micrococcus albus</i>	+	
<i>Helminthosporium spiciferum</i>	+	+	<i>Micrococcus candidus</i>	+	
<i>Paecilomyces variotii</i>	+		<i>Micrococcus ureae</i>	+	
<i>Penicillium oxalicum</i>	+	+	<i>Sarcina flava</i>	+	+
<i>Phoma glomerata</i>	+	+	<i>Pseudomonas cepacia</i>	+	
<i>Stemphylium consortiale</i>	+	+	<i>Pseudomonas maltophilia</i>	+	
<i>Torula nigra</i>		+			

2.2. Biodegradation

Giacomucci et al. [44] studied the biodegradation of nitrocellulose paint by the *Desulfovibrio desulfuricans* bacterium. The study measured changes in the concentration of nitrates, nitrites, and the concentration of ammonia in a culture medium. In order to determine the chemical changes in the paint, infrared spectroscopy was performed, as well as colorimetric assays designed to identify the changes in colour caused by microorganisms. Microscopic measurements showed that bacteria were able to adhere to the surface of the paint and alter its adhesive properties. The study determined that *Desulfovibrio desulfuricans* was capable of degrading nitrocellulose used as a binder in the products mentioned [44].

Ishfaq et al. [45], who authored a study, researched the biodegradation of paint coatings by fungi and bacteria. The researchers detected fungi such as *Aspergillus* sp., *Phanerochaete chrysosporium* sp. and *Rhizopus* sp. on an agar culture medium that contained paint. The study also used bacteria from the *Bacillus* sp. and *Pseudomonas* sp. group. Results showed that, under stress conditions, the selected species of fungi and bacteria were capable of adhering to the surface of the paint and causing its degradation. It was also observed that the selected fungi showed a greater ability to degrade paint coatings than bacteria [45].

Study [46] identified bacteria present in waste produced by the varnish industry. The following bacteria were identified as able to cause the biodegradation of the sludge: *Bacillus subtilis*, *Bacillus licheneformis*, *Bacillus cereus*, *Bacillus megaterium*, and *Pseudomonas fluorescens*.

In addition, the authors of the study determined bacterial generation time under growth conditions. These were, respectively, 44.67 minutes for *Pseudomonas fluorescens*; 45.04 minutes for *Bacillus subtilis*; 35 minutes for *Bacillus licheniformis*, 18 minutes for *Bacillus cereus*, 19 minutes for *Bacillus megaterium*, and 53 minutes for mixed bacterial cultures.

2.3. Biodegradation of polyurethane coatings

The study conducted by Wojturska [47] concerned the susceptibility of polyurethane coatings to enzymatic degradation. The coatings were obtained as the result of polyaddition of polycarbonate diol or polyesterether polyols of different degrees of branching, 1,4-butanediol to 2,2,4-trimethyl-1,6-hexamethylene diisocyanate (TMPD). Polyurethane coatings were subjected to enzymatic degradation by lipases isolated from *Candida antarctica* type A (Novozym 735), type B (CALB L), *Aspergillus* sp. (Nowozym 51032), lipase isolated from *Rhizomucor miehei* (Pallatase 20000 L) or lipase from *Thermomyces lanuginosus* (Lipozyme TL 100 L). Samples were subjected to enzymatic activity for 6 weeks, each at its optimum temperature, i.e. T = 50°C for Novozym CALB L and 36°C for all the rest. Studies showed that lipase type B isolated from *Candida antarctica* (Novozym CALB L) has the greatest affinity to the synthesized PUR coatings; the smallest change was caused by the lipase isolated from *Thermomyces lanuginosus* (Lipozyme TL 100 L).

The authors of [48] studied the growth of bacteria in a culture medium, using polyurethane coatings as the sole source of carbon. They used the following bacteria: *Acinetobacter calcoaceticus*, *Pseudomonas* sp. and *Arthrobacter globiformis*. All the microorganisms in the study showed an ability to grow using polyurethane coatings as the only source of carbon [48]. In addition, the following bacteria were also shown to degrade polyurethane coatings: *Pseudomonas chlororaphis*, *Pseudomonas putida*, *Pseudomonas cepacia*, *Pseudomonas aeruginosa*, *Corynebacterium* sp., *Acinetobacter calcoaceticus*, *Arthrobacter globiformis* [49–51].

3. Biodegradation of varnish removal agents

Studies were also conducted to investigate the biodegradation of waste produced in the process of varnish removal. Arquiga et al. [52] studied the decomposition of paint stripping wastewater, which contained methylene chloride at the concentration of 5000 mg·dm⁻³, phenol at the concentration of 1800 mg/dm³, and other organic ingredients, such as paraffin waxes, cellulose derivatives, petroleum sulfonates, and naphthalene at the total concentration of 2200 mg·dm⁻³. Active sludge was used. The study identified bacteria present in active sludge and their ability to grow in the presence of phenol and varnish stripping wastewater. In addition, researchers investigated whether these substances can serve as a source of carbon for the bacteria [52]. They isolated such bacteria as *Enterobacter* sp., *Klebsiella* sp., *Corynebacterium* sp., *Pseudomonas* sp., *Bacillus* sp., *Acinetobacter* sp.. The bacteria most frequently isolated from the wastewater were Gram-negative, with the majority belonging to the *Pseudomonas* genus. Gram-positive bacteria, represented by the genus *Bacillus*, were less

numerous. Studies showed that most of the isolated bacteria can biodegrade phenol and paint stripping wastewater. However, the most important role in the process is played by bacteria from the *Pseudomonas* and *Bacillus* genus [52].

Vanderberg-Twary et al. [53] researched the biodegradation of solvents used in varnish removal agents, focusing on the biodegradation of typical agents of this kind. They prepared two kinds of agents: one contained toluene, acetone, and dichloromethane, the other – dichloromethane, isopropyl alcohol, and methanol. The study used two types of bacteria: *Hyphomicrobium* sp. and *Rhodococcus rhodochrous* sp. [53].

The study showed that the selected strains of bacteria are able to degrade solvents present in paint stripping agents [53].

In patent US8202424, Almadidy et al. [54] described the biodegradation process of paint stripping waste containing high levels of organic substances. Their study used two kinds of bacteria: *Pseudomonas fluorescens* and *Bacillus subtilis* and one type of fungus, *Cunningham elegans*. They tested the changes in the COD of waste samples subjected to microbial activity. The study showed that it is possible to achieve as much as 47% decrease in COD with the use of the selected bacterial strains [54].

4. Conclusions

As evidenced by the above literature survey, there is a great diversity of microorganisms capable of growing on and in paint coatings. The most frequently observed bacteria belong to the genus *Bacillus* and *Pseudomonas*. The most frequently detected of fungi were *Penicillium* sp., *Aspergillus* sp. and *Cladosporium* sp. Studies on the biodegradation of varnishes and their stripping agents show that some of these microorganisms can be used for the deliberate degradation of old varnish coatings.

References

- [1] Joint publication, *Powłoki malarsko-lakiernicze*, WNT, Warszawa 1983.
- [2] Zyska B., *Mikrobiologiczna korozja materiałów*, WNT, Warszawa 1977.
- [3] Ravikumar H.R., Rao S.S., Karigar C.S., *Biodegradation of paints: a current status*, Indian Journal of Science and Technology, vol. 5(1), 2012, 1977–1987.
- [4] Zyska B., Żakowska Z., *Mikrobiologia Materiałów*, Wydawnictwo Politechniki Łódzkiej, Łódź 2005.
- [5] Orehek J., Dogsa I., Tomšič M., Jamnik A., Kočar D., Stopar D., *Structural investigation of carboxymethyl cellulose biodeterioration by Bacillus subtilis subsp. subtilis NCIB 3610*, International Biodeterioration & Biodegradation, vol. 77, 2013, 10–17.
- [6] Obidi O.F., Aboaba O.O., Makanjuola M.S., Nwachukwu S.C.U., *Microbial evaluation and deterioration of paints and paint-products*, Journal of Environmental Biology, vol. 30(5), 2009, 835–840.

- [7] Obidi O., Nwachukwu S., Aboaba O., *Investigation on the Biodegradative Potential of Pseudomonas aeruginosa on Water-based Paints*, Researcher, vol. 2(1), 2010, 57–67.
- [8] Shirakawa M.A., Loh K., John V.M., Silva M.E.S, Gaylarde C.C., *Biodeterioration of painted mortar surfaces in tropical urban and coastal situations: Comparison of four paint formulations*, International Biodeterioration & Biodegradation, vol. 65, 2011, 669–674.
- [9] Gaylarde P.M., Gaylarde C.C., *Algae and cyanobacteria on painted buildings in Latin America*, International Biodeterioration & Biodegradation, vol. 46, 2000, 93–97.
- [10] Crispim C.A., Gaylarde P.M., Gaylarde C.C., *Algal and Cyanobacterial Biofilms on Calcareous Historic Buildings*, Current Microbiology, vol. 46, 2003, 79–82.
- [11] Crispim C.A., Gaylarde C.C., Gaylarde P.M., *Biofilms on church walls in Porto Alegre, RS, Brazil, with special attention to cyanobacteria*, International Biodeterioration & Biodegradation, vol. 54, 2004, 121–124.
- [12] La Rosa F.R., Giese E.C., Dekker R.F.H., Sánchez Pelayo J., de Melo Barbosa A., *Microbiological contamination of water-based paints from an industry in the state of Paraná, Brazil*, Semina: Ciências Exatas e da Terra, Londrina, vol. 29, 2008, 85–92.
- [13] De Souza A., Gaylarde C.C., *Biodeterioration of varnished wood with and without biocide: implications for standard test methods*, International Biodeterioration & Biodegradation, vol. 49, 2002, 21–25.
- [14] Pangallo D., Bučková M., Kraková L., Puškárová A., Šaková N., Grivalský T., Chovanová K., Zemánková M., *Biodeterioration of epoxy resin: a microbial survey through culture-independent and culture-dependent approaches*, Environmental Microbiology, vol. 17, 2015, 462–479.
- [15] Cappitelli F., Principi P., Pedrazzani R., Toniolo L., Sorlini C., *Bacterial and fungal deterioration of the Milan Cathedral marble treated with protective synthetic resins*. Science of the Total Environment, vol. 385, 2007, 172–181.
- [16] Cappitelli F., Zanardini E., Sorlini C., *The Biodeterioration of Synthetic Resins Used in Conservation*, Macromolecular Bioscience, vol. 4, 2004, 399–406.
- [17] Gurtner C., Heyrman J., Pinar G., Lubitz W., Swings J., Rölleke S., *Comparative analyses of the bacterial diversity on two different biodeteriorated wall paintings by DGGE and 16S rDNA sequence analysis*, International Biodeterioration & Biodegradation, vol. 46, 2000, 229–239.
- [18] Heyrman J., Swings J., *16SrDNA Sequence Analysis of Bacterial Isolates from Biodeteriorated Mural Paintings in the Servilia Tomb (Necropolis of Carmona, Seville, Spain)*, Systematic and Applied Microbiology, vol. 24, 2001, 417–422.
- [19] Imperi F., Caneva G., Cancellieri L., Ricci M.A., Sodo A., Visca P., *The bacterial aetiology of rosy discoloration of ancient wall paintings*, Environmental Microbiology, vol. 9, 2007, 2894–2902.
- [20] Nugari M.P., Pietrini A.M., Caneva G., Imperi F., Visca P., *Biodeterioration of mural paintings in a rocky habitat: The Crypt of the Original Sin (Matera, Italy)*, International Biodeterioration & Biodegradation, vol. 63, 2009, 705–711.
- [21] Zucconi L., Gagliardi M., Isola D., Onofri S., Andaloro M.C., Pelosi C., Pogliani P., Selbmann L., *Biodeterioration agents dwelling in or on the wall paintings of the Holy*

- Saviour's cave (Vallerano, Italy)*, International Biodeterioration & Biodegradation vol. 70, 2012, 40–46.
- [22] Groth I., Vettermann R., Schuetze B., Schumann P., Saiz-Jimenez C., *Actinomycetes in Karstic caves of northern Spain (Altamira and Tito Bustillo)*, Journal of Microbiological Methods, vol. 36, 1999, 115–122.
- [23] Gonzalez I., Laiz L., Hermosin B., Caballero B., Incerti C., Saiz-Jimenez C., *Bacteria isolated from rock art paintings: the case of Atlanterra shelter (south Spain)*, Journal of Microbiological Methods vol. 36, 1999, 123–127.
- [24] Stomeo F., Portillo M.C., Gonzalez J.M., Laiz L., Saiz-Jimenez C., *Pseudonocardia in white colonizations in two caves with Paleolithic paintings*, International Biodeterioration & Biodegradation, vol. 62, 2008, 483–486.
- [25] Rosado T., Gil M., Mirão J., Candeias A., Caldeira A.T., *Oxalate biofilm formation in mural paintings due to microorganisms – A comprehensive study*, International Biodeterioration & Biodegradation, vol. 85, 2013, 1–7.
- [26] Okunye O.L., Morakinyo K.O., Ayedun J.S., *Isolation and Characterization of Fungi Associated with in-Can Degradation of Paint*, Journal of Environment and Earth Science, vol. 3(12), 2013, 142–145.
- [27] Saad D.S., Kinsey G.C., Kim S., Gaylarde C.C., *Extraction of genomic DNA from filamentous fungi in biofilms on water-based paint coatings*, International Biodeterioration & Biodegradation, vol. 54, 2004, 99–103.
- [28] Giannantonio D.J., Kurth J.C., Kurtis K.E., Sobeck P.A., *Molecular characterizations of microbial communities fouling painted and unpainted concrete structures*, International Biodeterioration & Biodegradation, vol. 63, 2009, 30–40.
- [29] Aina V.O., Adewumi A.A.J., Haruna H., Zakari A., *Isolation and Identification of Fungi Associated with the Deterioration of Painted Wall surfaces within Kaduna Polytechnic*, Asian Journal of Medical Sciences, vol. 3(6), 2011, 250–253.
- [30] Gaylarde C.C., Gaylarde P.M., *A comparative study of the major microbial biomass of biofilms on exteriors of buildings in Europe and Latin America*, International Biodeterioration & Biodegradation, vol. 55, 2005, 131–139.
- [31] Shirakawa M.A., Tavares R.G., Gaylarde C.C., Taqueda M.E.S., Loh K., John V.M., *Climate as the most important factor determining anti-fungal biocide performance in paint films*, Science of the Total Environment, vol. 408, 2010, 5878–5886.
- [32] Garg K.L., Jain K.K., Mishra A.K., *Role of fungi in the deterioration of wall paintings*. Science of the Total Environment, vol. 167, 1995, 255–271.
- [33] Ciferri O., *Microbial Degradation of Paintings*, Applied and Environmental Microbiology, vol. 65, 1999, 879–885.
- [34] Shirakawa M.A., Gaylarde C.C., Gaylarde P.M., John V., Gambale W., *Fungal colonization and succession on newly painted buildings and the effect of biocide*, FEMS Microbiology Ecology, vol. 39, 2002, 165–173.
- [35] Saad D.S., Saad D.S., Kinsey G.C., Paterson R., Gaylarde C., *Ergosterol analysis for the quantification of fungal growth on paint films. Proposal for a standard method*, Surface Coatings International Part B: Coatings Transactions, vol. 86, 2003, 131–134.

- [36] Gaylarde C.C., Morton L.H.G., Loh K., Shirakawa M.A., *Biodeterioration of external architectural paint films— A review*. International Biodeterioration & Biodegradation, vol. 65, 2011, 1189–1198.
- [37] Lugauskas A., Levinskaitė L., Pečiulytė D., *Micromycetes as deterioration agents of polymeric materials*, International Biodeterioration & Biodegradation, vol. 52, 2003, 233–242.
- [38] Gorbushina A.A., Petersen K., *Distribution of microorganisms on ancient wall paintings as related to associated faunal elements*, International Biodeterioration & Biodegradation, vol. 46, 2000, 277–284.
- [39] Gorbushina A.A., Heyrman J., Dornieden T., Gonzalez-Delvalle M., Krumbein W.E., Laiz L., Petersen K., Saiz-Jimenez C., Swings J., *Bacterial and fungal diversity and biodeterioration problems in mural painting environments of St. Martins church (Greene–Kreienzen, Germany)*, International Biodeterioration & Biodegradation, vol. 53, 2004, 13–24.
- [40] Santos A., Cerrada A., García S., San Andrés M., Abrusci C., Marquina D., *Application of Molecular Techniques to the Elucidation of the Microbial Community Structure of Antique Paintings*, Microbial Ecology, 58, 2009, 692–702.
- [41] Romero-Noguera J., Bolívar-Galiano F.C., Ramos-López J.M., Fernández-Vivas M.A., Martín-Sánchez I., *Study of biodeterioration of diterpenic varnishes used in art painting: Colophony and Venetian turpentine*, International Biodeterioration & Biodegradation, vol. 62, 2008, 427–433.
- [42] Romero-Noguera J., Martín-Sánchez I., Doménech-Carbó M.T., Osete-Cortina L., López-Miras M.M., Bolívar-Galiano F., *Analytical characterisation of the biodeterioration of diterpenoid labdanic varnishes used in pictorial techniques: Sandarac and Manila copal*, International Biodeterioration & Biodegradation, vol. 90, 2014, 99–105.
- [43] Seves A.M., Sora S., Ciferri O., *The Microbial Colonization of Oil Paintings. A Laboratory Investigation*, International Biodeterioration & Biodegradation, vol. 37, 1996, 215–224.
- [44] Giacomucci L., Toja F., Sanmartín P., Toniolo L., Prieto B., Villa F., Cappitelli F., *Degradation of nitrocellulose-based paint by Desulfovibrio desulfuricans ATCC 1354*, Biodegradation vol. 23, 2012, 705–716.
- [45] Ishfaq S., Ali N., Tauseef I., Khattak M.N.K., Shinwari Z.K., Ali M.I., *Analysis of paint degradation by fungal and bacterial species*, Pakistan Journal of Botany vol. 47, 2015, 753–760.
- [46] Dwipayana, Ariesyady H.D., (2010) Identification of bacterial diversity in waste recycling paint sludge by conventional microbiological technique. http://www.ftsl.itb.ac.id/kk/rekayasa_air_dan_limbah_cair/wp-content/uploads/2010/11/pe-ww7-dwipayana-15305020.pdf (access: 10.03.2016).
- [47] Wojturska J., *Degradacja enzymatyczna powłok poliuretanowych*, Ochrona Przed Korozją, vol. 52(12), 2009, 595–600.
- [48] Halim A., El-Sayed M.M., Mahmoud W.M., Davis E.M., Coughlin R.W., *Biodegradation of Polyurethane Coatings by Hydrocarbon-Degrading Bacteria*, International Biodeterioration & Biodegradation vol. 37, 1996, 69–79.
- [49] Gautam R., Bassi A.S., Yanful E.K., Cullen E., *Biodegradation of automotive waste polyester polyurethane foam using Pseudomonas chlororaphis ATCC55729*, International Biodeterioration & Biodegradation vol. 60, 2007, 245–249.

- [50] Howard G.T., *Biodegradation of polyurethane: a review*, International Biodeterioration & Biodegradation, vol. 49, 2002, 245–252.
- [51] Howard G.T., *Microbial biodegradation of polyurethane*. Recent Developments in Polymer Recycling, 2011.
- [52] Arquiaga M.C., Canter L.W., Robertson J.M., *Microbiological characterization of the biological treatment of aircraft paint stripping wastewater*, Environmental Pollution vol. 89, 1995, 189–195.
- [53] Vanderberg-Twary L., Steenhoudt K., Travis B.J., Hanners J.L., Foreman T.M., Brainard J.R., *Biodegradation of paint stripper solvents in a modified gas lift loop bioreactor*, Biotechnology and Bioengineering, vol. 55, 1997, 163–169.
- [54] Almadidy A., Lavayssiere N., *Microbial degradation of water-borne paint containing high levels of organic solvent*, patent No. US8202424, 2012.

Tomasz M. Majka (tomasz.majka@pk.edu.pl)

Agnieszka Leszczyńska

Krzysztof Pielichowski

Department of Chemistry and Technology of Polymers, Faculty of Chemical Engineering and Technology, Cracow University of Technology.

MECHANICAL PROPERTIES VS BIOTIC DEGRADATION OF POLYAMIDE/LAYERED SILICATES NANOCOMPOSITES

WŁAŚCIWOŚCI MECHANICZNE VS DEGRADACJA BIOTYCZNA NANOKOMPOZYTÓW POLIAMID/KRZEMIANY WARSTWOWE

Abstract

The purpose of this study was to obtain polyamide 6 nanocomposites with different organically modified clays and to study the biotic degradation behaviour vs mechanical properties of the obtained materials. Thermal stability of pure organoclays was investigated using thermogravimetric analysis (TGA). The prepared nanocomposites were characterized by X-ray diffraction (XRD) and infrared spectroscopy (IR). The evolution of mechanical properties was also studied. The obtained results confirm good interactions of nanofillers with the polymer, showing the formation of intercalated and/or partially exfoliated structures. The nanocomposites showed higher thermal stability compared to pure polymer, and advantageous mechanical properties. Finally, a discussion related to the effects of biotic degradation on mechanical properties of PA6/MMT nanocomposites is presented.

Keywords: polymer nanocomposites, polyamide 6, layered silicates, montmorillonite, biotic degradation

Streszczenie

Celem przeprowadzonych badań było otrzymanie nanokompozytów poliamidowych z różnym typem organicznie modyfikowanej glinki. Stabilność termiczna czystych glinek mierzono za pomocą analizy termogravimetrycznej (TGA). Otrzymane nanokompozyty zostały scharakteryzowane przy wykorzystaniu dyfrakcji promieni rentgenowskich (XRD) oraz spektroskopii w podczerwieni (IR). Zbadano również zmiany własności mechanicznych. Uzyskane wyniki potwierdzają dobrą interkalację nanonapełniaczy z matrycą polimerową, pokazując tworzenie struktur częściowo eksfoliowanych. Nanokompozyty wykazały wyższą stabilność termiczną w porównaniu do czystego poliamidu. Artykuł zamyka dyskusja związana ze skutkami degradacji biotycznej nanokompozytów PA6/MMT na ich właściwości mechaniczne.

Słowa kluczowe: nanokompozyty polimerowe, krzemiany warstwowe, poliamid 6, montmorylonit

1. Introduction

The search for the new hybrid polymer materials with improved processability and better thermal properties than commercially available polymers has received significant attention from both academia and industry. In the last 20 years, natural clay–polymer nanocomposites have become an active field of research because of their unique microstructures and enhanced properties. However, in some areas of industry, there is a great demand for materials with improved functional properties relative to those on the market. The growing needs are related primarily to the military sector, automotive and packaging, where high demands are made for polymer products [1–7]. New applications involve a need to develop innovative solutions or modification of polymer materials produced. The growing interest is focused on polymeric nanocomposites, i.e. materials, consisting of two separable phases: a continuous phase (matrix) and a dispersed phase comprised of nanoparticles having at least one dimension less than 100 nm [1-3]. The improvement of certain properties allows functional or structural nanocomposites to be obtained. Among different nanofillers, special attention was paid to layered silicates, particularly montmorillonite (MMT). Extensive studies have shown that proper dispersion of particles of montmorillonite in a polymer matrix can result in improved mechanical, thermal, barrier and flame retardant properties [2–7]. Polyamide 6 nanocomposites with montmorillonite, although well-known, suffer from thermo-oxidative degradation and flammability [8–11].

It is well known that MMT can substantially improve thermoplastic polymer matrix properties even at a low load (3-5%) of the nanofiller [12]. Wu et al. [13] reported that modulus, bending strength and heat distortion temperature of these nanocomposites increased with increasing clay content, and tensile strength increased up to 4.3 wt% and then decreased with further increases of montmorillonite content in PA6/organoclay nanocomposites prepared by melt compounding. Another study [14] showed that both nanocomposites prepared with 1% and 3% of MMT were exfoliated, but at 5% load MMT was intercalated among two basal planes of MMT, leading to an expansion of the interlayer spacing. However, the thermogravimetric results indicated that the nanocomposites showed largely higher thermal stability compared to the pure polymer.

Current studies are focused on e.g. biopolymer/clay nanocomposites [15] or polyamide /polyolefin/clay blend nanocomposites [16]. Although numerous studies have been devoted to polyamide/MMT nanocomposites, the problems associated with biotic degradation behaviour vs mechanical properties changes have not been sufficiently addressed.

The aim of this work is to analyse the structure, morphology, thermal stability and mechanical properties of polyamide/montmorillonite (PA6/MMT) nanocomposites subjected to biotic degradation. The nanocomposites were prepared by the melt intercalation method, where three different clays were used. Each of the clays was characterized by thermogravimetric analysis (TGA) and only two of them were recommended for processing with PA6. The prepared nanocomposites were then characterized by X-ray diffraction (XRD) and infrared spectroscopy (IR). The mechanical properties – flexural strength – of PA6/MMT nanocomposites after long-term biotic degradation have been determined.

2. Experimental

2.1. Materials

The nanocomposites were produced using a PA6 melt-compounded with three types of organo-modified montmorillonites (OMMT), developed for extrusion compounding with PA6. Dellite® 43B (D43B), supplied by Laviosa, was a dimethyl-benzyl-hydrogenated tallow ammonium salt modified clay; Dellite® 72T (D72T) was a dimethyl-di(hydrogenated tallow) ammonium salt modified clay with low modifier content. Nanomer® I30P (NI30P), supplied by Nanocor, was a clay organophilized with high content of dimethyl-di(hydrogenated tallow) ammonium salt. The PA6 (Tarnamid® T30), supplied by Grupa Azoty SA, was a high-viscosity extrusion grade PA designed for technical applications.

2.2. Sample preparation

Before the preparation of nanocomposites, the materials were dried in a laboratory vacuum oven. Polyamide was dried at 80°C for 3 hours. Dried matrix and 3 wt% of nanofiller were mechanically premixed then compounded using a twin co-rotating screw extruder Thermo Scientific Rheomex PTW 16/25 XL with a screw speed of 150 rpm, a temperature profile in the range of 245°C (feed throat) to 260°C (die). The compounds were extruded as single fibre of 4 mm diameter, which was hauled into a quenching water trough prior to being pelletized (Table 1).

Table 1. Processing conditions of PA6/MMT nanocomposites

Twin co-rotating screw extruder								
Feeder performance (%)	Rotational speed (rpm)	Heating zones						
		1	2	3	4	5	6	Die
0.3	150							
Temperature (°C)		245	245	245	250	255	250	260
Atmospheric venting		-----	-----	-----	-----	YES	-----	-----
Length of the zones (mm)		80	60	60	64	60	76	23
L/D		5.00	3.75	3.75	4.00	3.75	4.75	-----
Cooling tank								
Length of cooling surface (mm)		1500						
Tank volume (dm ³)		27						
Height of bath (mm)		1081						
Water temperature (°C)		18						
Pelletizer								
Size of pellets (mm)		1						
Rotational speed (1/s)		12						

The dried nanocomposites were moulded to form flexural bar specimens using a Zamak WT12 injection moulding machine. The cylinder temperature profile was 232°C and the mould temperature was maintained at 90°C. To minimize moisture uptake moulded specimens were immediately wrapped in aluminium foil and sealed in polyethylene bags. Prior to mechanical testing, all specimens were dried for 3 h at 80°C and then conditioned in a dessicator for a minimum of 3 days.

2.3. X-Ray Diffraction (XDR)

X-ray diffraction (XRD) measurements were performed using an X-ray powder diffractometer Phaser D2 with a Cu K α 2 ($\lambda=1.54 \text{ \AA}$). Measurements were made in two systems:

- ▶ in the range of 5–40° angle 2θ – a gap of 1 mm and an aperture of 3 mm in the trading system with step 0.5 at 180°;
- ▶ in the range of 0.5–3° angle 2θ – a gap of 0.2 mm and an aperture of 1 mm in the stationary system.

2.4. Infrared spectroscopy (IR)

Studies using Fourier transform infrared spectroscopy (FTIR) were performed in reflectance mode using a Perkin Elmer Spectrum 65 spectrometer with ATR accessory (diamond crystal / ZnSe – spectral range 4000 - 650 cm^{-1} , penetration depth ~ 2 microns, number of scans 32, resolution 4 cm^{-1});

2.5. Thermal analysis

Thermogravimetric measurements were made using a thermogravimetric analyser Netzsch TG 209. Measurements were carried out using the following measurement conditions:

- ▶ Measurement temperature range of 20°C-605°C;
- ▶ Heating rate of 10°C/min;
- ▶ Oxidizing environment, air flow of 15 cm^3 /min.

In each case, the measuring corundum ($\alpha\text{-Al}_2\text{O}_3$) cell was open and the mass of the sample ~ 5 mg.

Thermal conversion of PA6/OMMT nanocomposites was investigated by Differential Scanning Calorimetry using a Mettler Toledo DSC823° calorimeter. Calorimetric measurements were carried out in an oxidizing atmosphere. The calorimeter was calibrated using indium and zinc standards. The sample (ca. 10 mg) was placed in an aluminium pan, and sealed in the press. The measurements were performed according to the two programs that allowed the evaluation of thermal properties:

- ▶ Heating from -30 to 300°C at 20°/min;
- ▶ Cooling from 300 to -30°C at 5°/min.

2.6. Mechanical tests

The static flexural strength of the nanocomposites was recorded using a universal testing machine Zwick Z005 TH in accordance with DIN EN ISO 178, and with a deformation speed of 2 mm / min. Tests were carried out at three temperatures: 0, 25 and 80°C. The reported values of strength were the arithmetic average of at least eight measurements. As a result of the test, $\sigma = f(\epsilon)$ curves and the bending modulus values, flexural bending stress, flexural strength, elongation area at the bending strength and the bending stress at failure were obtained.

2.7. Biotic degradation

Biotic degradation was carried out for 475 days. The PA6/OMMT composite panels were obtained by compression press (Zamak P-200) with heated punches. Press conditions were as follows: press temperature – 230°C; time of process – 5 minutes; compression pressure – 25 bar. Collected plates with dimensions 100x100x10 mm were screwed together by screws to form a cube. The composite cube was buried underground at a depth of 1 metre in the ground with pH of ca. 6.8.

3. Results and discussion

3.1. TGA of fillers

Fig. 1 shows thermogravimetric curves of commercial organoclays. For all tested clays, content of absorbed moisture oscillates around 2–3%. The course of TG curves reveals two decomposition stages. The first stage of degradation in the range up to approx. 200°C is related to the evaporation of water and the emission of low molecular weight modifier compound present in the system.

Depending on the type of quaternary salt, the onset and end temperatures of the first stage differ; to compare the results of thermogravimetric analysis, representative indices of 5, 10, 20%, and the maximum weight loss, have been identified – Table 2.

Table 2. Weight loss parameters of organoclays

Sample	T _{5%} [°C]	T _{10%} [°C]	T _{20%} [°C]	T _{max} [°C]	T _{onset} [°C]
Dellite® 43B	211	232	288	227	197
Dellite® 72T	253	283	337	298	245
Nanomer® I30P	306	340	502	340	292

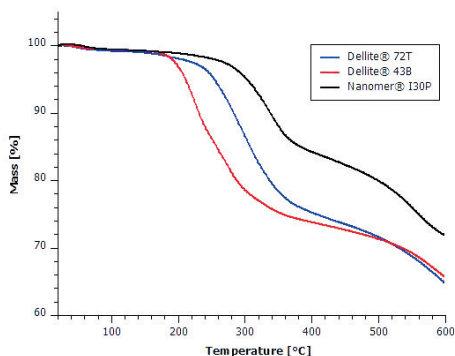


Fig. 1. Thermogravimetric curves of commercial organoclays

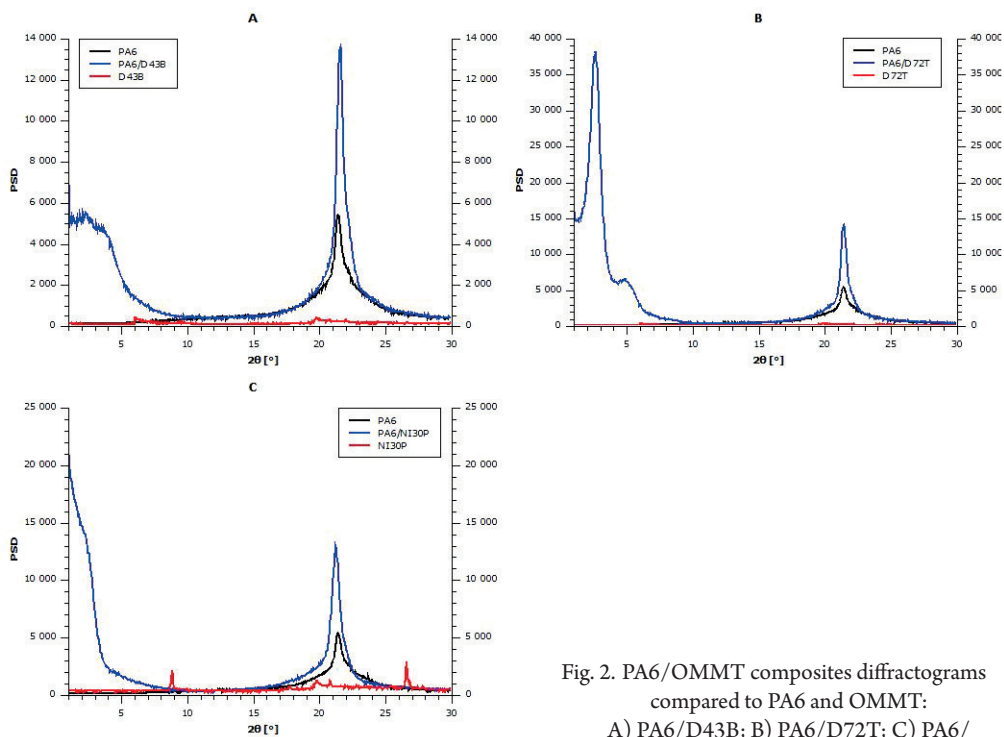


Fig. 2. PA6/OMMT composites diffractograms compared to PA6 and OMMT:

A) PA6/D43B; B) PA6/D72T; C) PA6/NI30P

Among the clays, the lowest value of weight loss reached montmorillonite modified with dimethyl-benzyl-hydrogenated tallow ammonium salt. On the other hand, the highest values of the indicators were found for NI30P. The second stage of degradation in the range of 400-600°C may be linked to decomposition reactions resulting in the formation of low molecular weight compounds.

3.2. X-Ray Diffraction (XDR)

The analysis of the XDR results allowed information to be acquired on the structure of the final compositions. Polyamide 6 may crystallize in three crystal forms: a stable phase α , a metastable phase β , and an unstable phase γ . Phase α is formed by arrangement of chains in regular conformations (all trans), while phase γ is a quasi-hexagonal structure form, in which the chains are helically twisted due to an irregular arrangement of intermolecular hydrogen bonds between amide groups. The meta-stable phase β is a rare intermediate phase from phase α to γ , or vice versa [17–19]. In Fig. 2 diffractograms of polyamide 6 and three PA6/OMMT composites are presented. In a 2θ angle range of $20\text{--}24^\circ$, one may observe reflections that come from the crystal phase of polyamide 6.

The most intense maximum around the angle $2\theta = 21^\circ$ stands for the created phase α_{100} . The intensity of this reflection increases together with the addition of layered silicate to the matrix, which would suggest that the presence of a filler helps in creating a stable phase α . The conducted deconvolution of curves (Fig. 3) reveals the presence of other reflections. For polyamide 6 there is an overlap of five reflections, while the maximum angle 2θ at 13.6 and 22.0° occurs due to the presence of phases γ_{020} and γ_{001} , accordingly. Reflection around the angle $2\theta 23^\circ$ indicates the presence of phase $\alpha_{002}/202$. The distribution of phases α and γ analysed

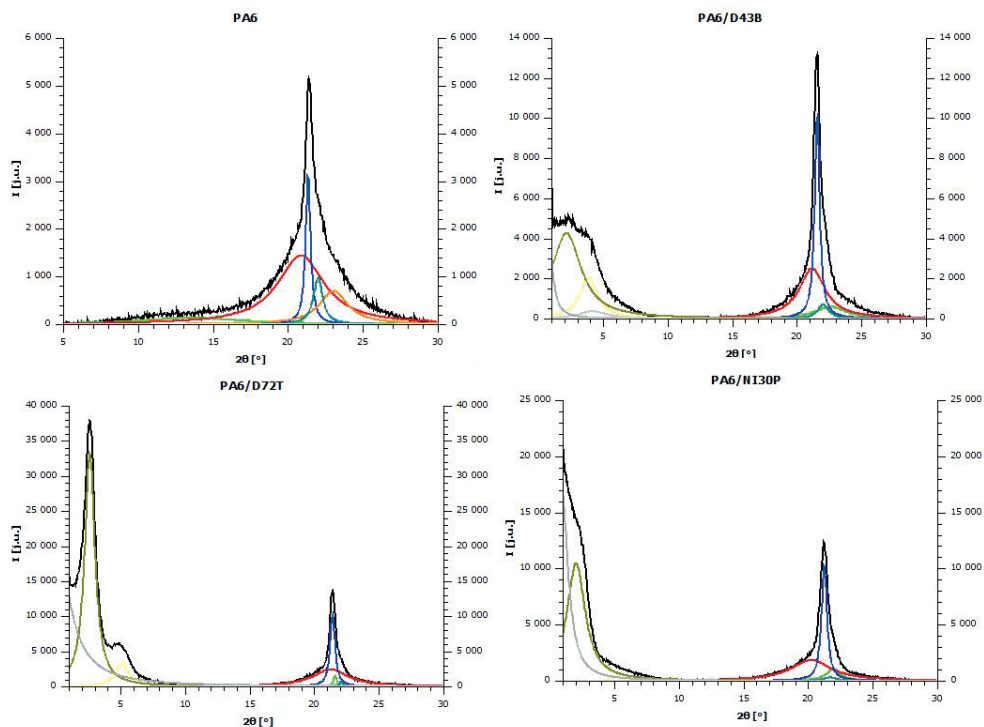


Fig. 3. PA6 and PA6/OMMT diffractograms with deconvolution results

Table 3. Crystallinity degree and interplanar distances determined for PA 6 and PA6/OMMT nanocomposites

Sample	Reflex No	2 Theta [°]	d [nm]	X _c [%]
PA6	1	13.6	0.65	39.8
	2	20.8	0.43	
	3	21.3	0.42	
	4	22.0	0.40	
	5	23.0	0.39	
PA6/D43B	1	0.7	12.86	53.7
	2	2.1	4.13	
	3	3.9	2.29	
	4	4.2	2.11	
	5	21.2	0.42	
	6	21.6	0.41	
	7	22.0	0.40	
	8	22.6	0.39	
PA6/D72T	1	2.6	3.38	40.8
	2	5.1	1.73	
	3	21.2	0.42	
	4	21.4	0.42	
	5	21.6	0.41	
	6	22.1	0.40	
PA6/NI30p	1	0.9	9.77	49.9
	2	2.0	4.37	
	3	5.1	1.74	
	4	20.2	0.44	
	5	21.2	0.42	
	6	21.7	0.41	
	7	21.9	0.41	

in this case is slightly different from that described in Ref. [18, 19]. In the quoted publications, the contribution of phase γ was more visible, while the addition of clay had a significant influence on the initiation of creating a phase that would be less thermodynamically stable. In Ref. [20] it has been suggested that the addition of layered silicate forces amide groups to be placed beyond the layer created by the chains. Thus, conformation changes of chains occur, which limits the creation of hydrogen bonds and the appearance of phase γ . In the diffractograms, elaborated as part of our research, the contribution of phase γ is minimal and visible only after performing the deconvolution. In the case of all of the compared material (regardless the clay used), the contribution of phase α was predominant, while the addition of layered silicates caused a decreased intensity of reflections originating from phase γ .

The interplanar distance of the organic clays was calculated as:

- ▶ For Dellite® 43B – reflection at 6.0° corresponds to interplanar distance of 1.47 nm;
- ▶ Dellite® 72T – reflection at 5.8° corresponds to interplanar distance of 1.52 nm;
- ▶ Nanomer® I30P - reflection at 8.9° corresponds to interplanar distance of 0.99 nm.

For PA6/OMMT systems (Dellite® 43B), the presence of a wide band of reflections at 0.7, 2.1, 3.9 and 4.2 of the 2θ angle suggests an efficient separation of montmorillonite layers in the range of 2.1–12.9 nm.

Implementation of Dellite® 43B and Nanomer® I30P organoclay helps in the creation of a polyamide nanocomposite with a large degree of separation of the clay layers, which suggests the presence of partial exfoliation of OMMT layers in the matrix. The presence of the preferred exfoliated structure results in an enhancement of physical and mechanical properties of a polymer nanocomposite [17–20].

The performed deconvolution of diffractograms has also enabled separation of signals from the amorphous and crystal phase of polyamide 6. Thus, according to the procedure described in Refs. [17–20], it was possible to calculate the crystallinity of polyamide 6 in the compositions analysed. In Table 3, the maximum reflections and their corresponding interplanar distances have been presented on the basis of the Bragg equation, as well as the crystallinity degree. The crystallinity degree of a pure matrix was 39.8%. The addition of 3% (m/m) of Dellite® 43B and Nanomer® I30P montmorillonite has increased the degree of crystallinity by almost 10%.

3.3. Infrared spectroscopy (IR)

The IR spectra of polyamide composites (Fig. 4) reveal strong bands due to vibration of N-H and C=O bonds, which are characteristic of secondary amides:

- ▶ $\nu = 3330\text{--}3050\text{ cm}^{-1}$ band of N-H bonds vibrations,
- ▶ $\nu = 2950\text{--}2800\text{ cm}^{-1}$ band of C-H bonds vibrations,
- ▶ $\nu = 1680\text{--}1630\text{ cm}^{-1}$ band of C=O bonds vibrations,
- ▶ $\nu = 1560\text{--}1510\text{ cm}^{-1}$ band of N-H bond deformations,



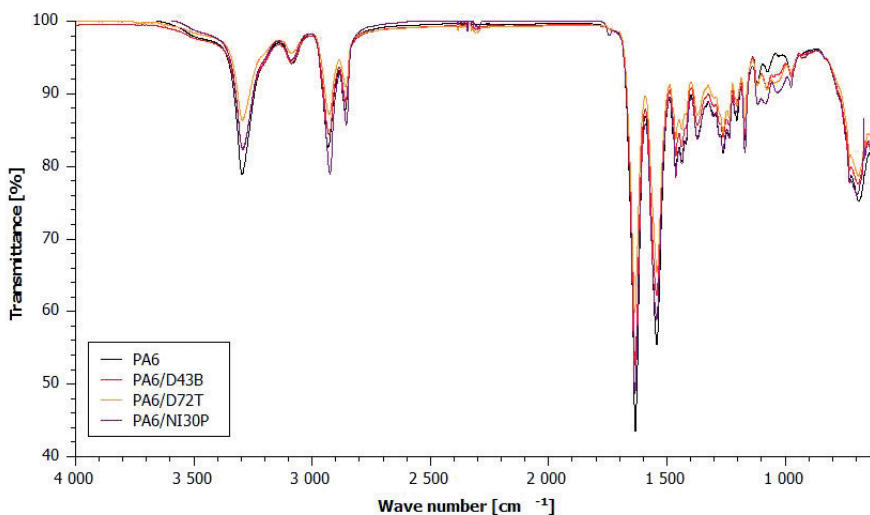


Fig. 4. IR spectra of PA6 – black curve; PA6/D43B – red curve; PA6/D72T – orange curve; PA6/NI30P – blue curve

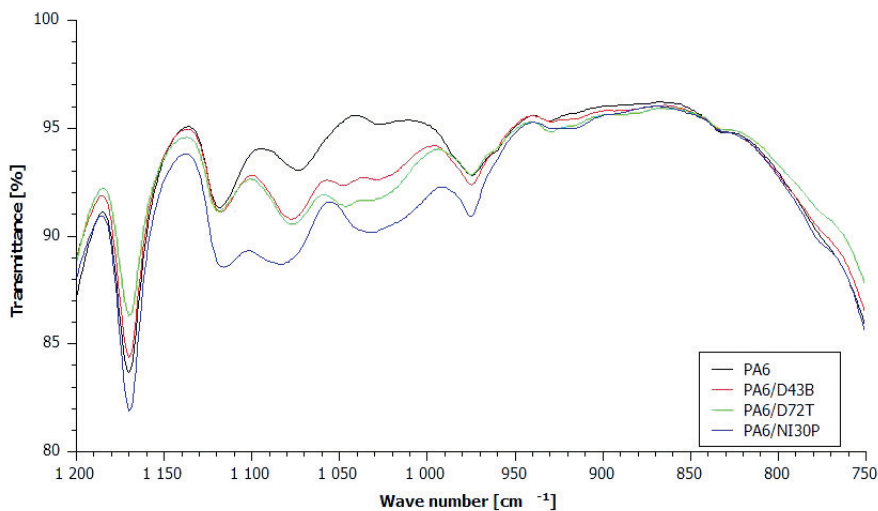


Fig. 5. IR spectra of PA6 – black curve; PA6/D43B – red curve; PA6/D72T – orange curve; PA6/NI30P – blue curve in the range of 1200-750 cm⁻¹

On the IR spectra of polyamide composites in the range of wave numbers of 1200-930 cm⁻¹ (Fig. 4), there are bands related to vibrations of the CONH group in the plane that is applied for determination of the crystallinity degree of a polyamide. Fig. 5 shows the IR spectra in the range of wave numbers of 1200-750 cm⁻¹ of polyamide 6 and PA6/OMMT nanocomposites. Bands at $\nu = 1029$ and 928 cm⁻¹ are connected with the contribution of the

crystal phase α and γ . The weak band at $\nu = 960 \text{ cm}^{-1}$ is linked with the contribution of the crystal phase α . The band at $\nu = 975 \text{ cm}^{-1}$ is connected with the contribution of crystal phase γ . The band at $\nu = 1120 \text{ cm}^{-1}$, as amorphous band, occurs in all of the composites analysed and may be used as a reference band during the determination of the crystallinity degree [21].

The following values of absorbancy were acquired in order to calculate the crystallinity degree:

- ▶ $A\alpha$ - band of phase α and γ at $\nu = 929 \text{ cm}^{-1}$
- ▶ $A\gamma$ - band of phase α and γ at $\nu = 975 \text{ cm}^{-1}$

The crystallinity degree (X_c) was calculated from the following equation [21]:

$$X_c = 58.82 \cdot K + 32.06 \quad (1)$$

where: $K = A\alpha + 2/3 \cdot A\gamma$

Then, the contribution of phases α and γ has been calculated from the following equations [21]:

$$X_\alpha = 10.76 \cdot \ln(A\alpha) + 57.37 \quad (2)$$

$$X_\gamma = X_c - X_\alpha \quad (3)$$

The acquired values for the total crystallinity degree and the crystallinity degree of phases α and γ for respective polyamide materials have been presented in Table 4.

The characteristic absorption bands from phase α are 1463 cm^{-1} (scissoring vibrations of CH_2 group), 1419 cm^{-1} (scissoring vibrations of CH_2 group), 1373 cm^{-1} (wagging vibrations from amide III and CH_2 group), 1204 cm^{-1} (twisting vibrations of CH_2 group), 959 cm^{-1} (deformation vibrations in plane CO-NH), 928 cm^{-1} (deformation vibrations in plane CO-NH). The characteristic absorption bands for phase γ have been noted at 1437 cm^{-1} (scissoring vibrations of CH_2 group), 1373 cm^{-1} (wagging vibrations of CH_2 group), 1237 cm^{-1} (twisting vibrations of CH_2 group), 975 cm^{-1} (deformation vibrations in plane CO-NH) and 725 cm^{-1} (vibrations from amide V group) [21].

3.4. Differential Scanning Calorimetry

For the DSC analysis PA6/OMMT composites with three types of nanofiller (Dellite 43B, Dellite 72T and Nanomer I30Pm) and a reference PA6 were applied. In Fig. 6 DSC the curves obtained at a heating rate of $20^\circ/\text{min}$ and cooling rate of $5^\circ/\text{min}$ are shown. The glass transition temperature, the temperatures for melting of crystal phases, changes of melting enthalpy, the specific heat, the temperature of crystallization and the crystallinity degree have been determined in accordance with the procedure described in [22, 23] and presented in Table 4.

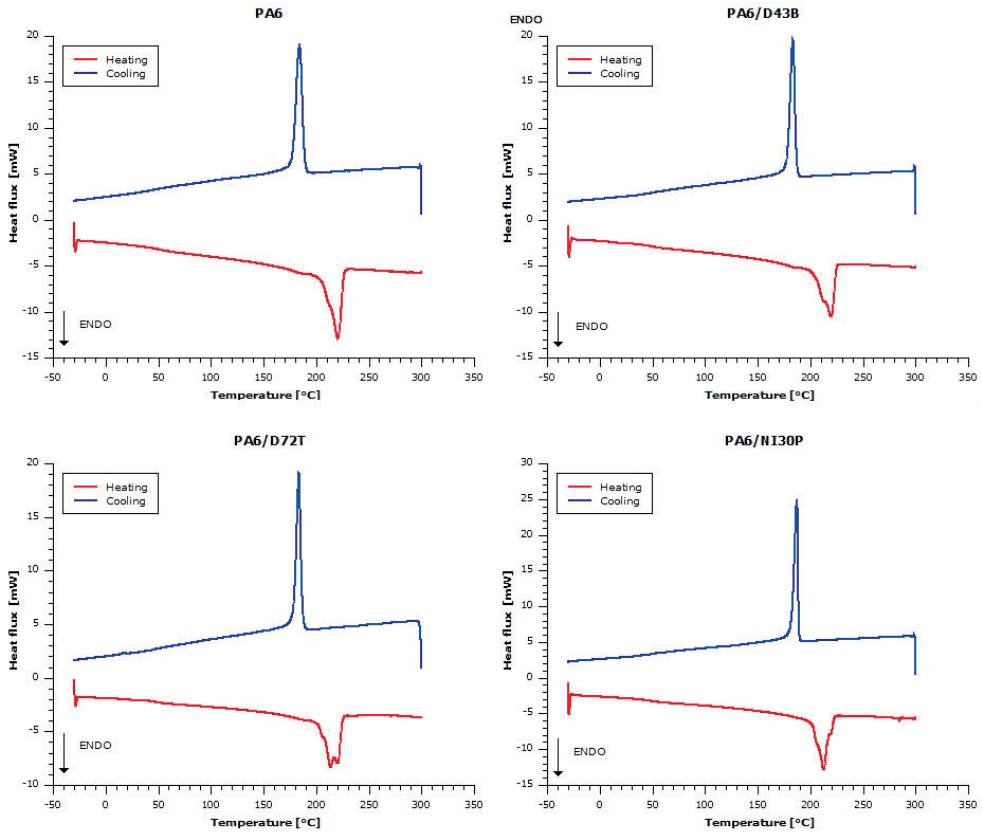


Fig. 6. DSC thermograms of polyamide 6 and PA6/OMMT composites

Table 4. Glass transition temperatures (T_g); melting temperature T_m ; change of melting enthalpy ΔH_m ; change of specific heat ΔC_p ; crystallinity degree X_c and crystallization temperature T_c and change of crystallization enthalpy ΔH_c for PA6 and PA6/OMMT composites

Sample	T_g [°C]	T_m [°C]	ΔH_m [J/g]	ΔC_p [J/gK]	X_c [%]	T_c [°C]	ΔH_c [J/g]
PA6	49	220	80.38	0.260	33.49	184	69.92
PA6/D43B	49	220	72.26	0.224	31.04	184	65.14
PA6/72T	48	214	61.24	0.196	26.31	184	63.30
PA6/NI30P	48	214	67.46	0.244	28.98	187	63.47

On the basis of the DSC analyses conducted it can be found that the implementation of OMMT had almost no influence on the glass transition temperature (T_g) of the polyamide matrix, neither did it change the temperature at which the crystallization occurred at its maximum speed. Moreover, ΔC_p and ΔH_m decreased. These observations suggest that the

addition of aluminosilicate helps in the transition from the glass state into the elastic state, since significantly less heat is required for the system in order to achieve the transition point between those two states. The presence of 3% weight of organoclay resulted also in lowering the X_c value. Depending on the nanofiller used, various values of crystallinity degree have been noted. The rather slow cooling of samples, however, caused a small decrease of the contents of the crystal phase of the nanocomposite, despite its crystallization capacity. This may suggest that the formation of the amorphous phase of polyamide 6 depends on the type of organic modifier present in the composite system.

3.5. Mechanical properties (Flexural strength)

The maximum value of bending stress acquired from charts $\sigma = f(\epsilon)$ (Fig. 7) for static bending resistance tests of polyamide 6 and two types of PA6/OMMT nanocomposites that were performed at temperatures of 0, 25 and 80°C have been described in Fig. 8 and Table 5.

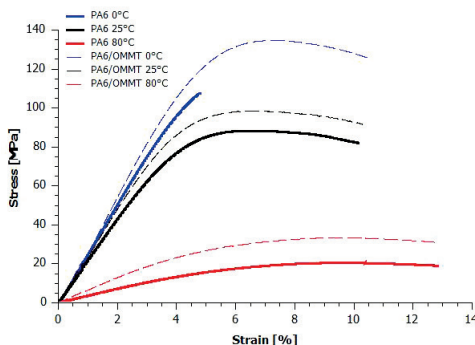


Fig. 7. Chart of dependencies in the deformation function for PA6, PA6/D43B and PA6/D72T performed at 0, 25 and 80°C

In the glass and brittle state at temperature of 0°C, due to a strong firmness, the bending stress increases rapidly along with the increase of deformation. The dependency described for polyamide 6 indicates a linear manner until the time of destruction. At a temperature of 0°C, the mobility of PA6 chain segments is rather minor and the applied bending strength causes the occurrence of small elastic stress. The organization of chains towards the stress is, in these conditions, more difficult and, after exceeding the critical point, there is a direct breaking of bonds in the macrochain. The addition of nanofiller causes not only an increase in its resistance against the strength applied, but is the factor which protects the structure against appearance of microfractures that would start the destruction process. While analysing the curve $\sigma = f(\epsilon)$, it may be assumed that the material enforced with montmorillonite resembles a tendency to change its state from glass and brittle into a pseudo-elastic one. This may be caused by the fact that dispersed organoclay absorbs part of the bending strength applied on the surface of a particular sample and, due to this fact, the material shows a higher resistance to static bending. Together with the increase of the used strength, there is a certain maximum capacity of energy

that may be stored by the nanoparticles stored in the matrix. After this capacity is exceeded, the resistance becomes successively lower, together with an increase of deformation.

In the area of the glass state with forced flexibility at the temperature of 25°C, the dependency $\sigma = f(\epsilon)$ also shows a linear character and, then, the curve is marked due to the occurrence of gaps and stress cracks in the micro-deformed area. Then, the effect of flow towards the cold is present and caused by the orientation processes of polymer chains. In the case of the tests for static bending resistance, the percentage of deformation is lower than in the case of static tensile resistance testing. Tests are ended after exceeding 7% deformation from the maximum point.

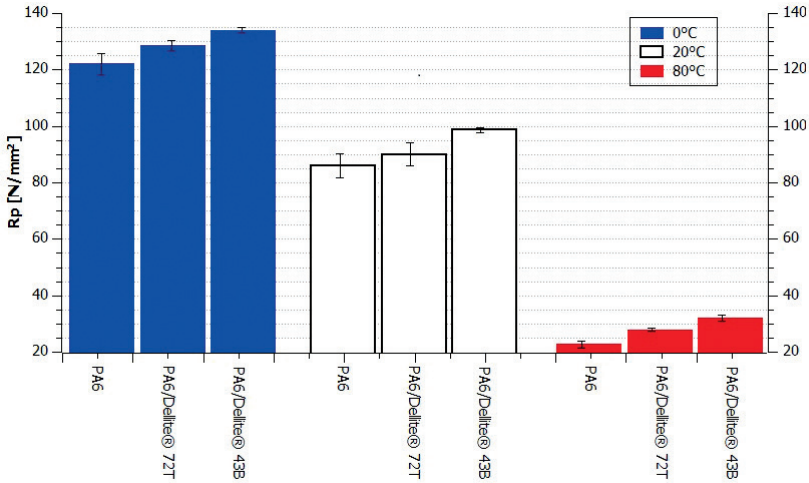


Fig. 8. Resistance against static bending (R_p) of PA6, PA6/D43B and PA6/D72T at various measuring temperatures

Table 5. Mechanical factors of PA6/MMT nanocomposites at various measuring temperatures: ϵ_f - bending modulus, σ_{jc} - flexural bending stress, ϵ_{fM} - elongation area at the bending strength, σ_{fB} - the bending stress at failure ϵ_{fB} - elongation area at the at failure

Sample	Temperature [°C]	ϵ_f [MPa]	σ_{jc} [MPa]	ϵ_{fM} [%]	σ_{fB} [MPa]	ϵ_{fB} [%]
PA6	0	2464	90.66	7.20	113.4	10.2
PA6/D43B		2832	101.18	6.98	124.6	10.0
PA6/D72T		2720	98.78	7.12	119.6	10.0
PA6	20	2022	69.76	7.08	80.0	10.4
PA6/D43B		2528	82.60	6.36	91.8	10.2
PA6/D72T		2304	75.68	6.68	83.8	10.4
PA6	80	384	13.74	10.14	21.1	13.0
PA6/D43B		682	21.78	9.10	29.8	12.4
PA6/D72T		558	18.12	9.56	25.9	13.0

When the temperature increases above the temperature of glass transition (test temperature equal 80°C), the polymer enters into a highly elastic range, becoming soft and flexible. In the range of small deformations, the stress-deformation curve is a linear process type, with a smaller module of E Young than in the state of forced elasticity in temperature of 25°C. The polymer becomes deformed when there is a fixed stress strength. Increase of temperature above T_g is advantageous for the mobility of segments of polymer chains and, different from the curve determined in ambient temperature, in a highly elastic area the stress increases monotonically, but non-linearly, along with the deformation. The elastic effects connected with entropic elasticity and the adhesive and elastic effects perform a predominant role.

3.6. Biotic degradation

Mechanical strength tests of composite materials, after their biotic degradation caused by biological factors – mainly enzymes produced by microbes, allows evaluation of the resistance of the material against biotic degradation. Tests have been conducted for 475 days in fertile soil of pH=6.8. Polyamide composites with layered silicate undergo a partial biological degradation that leads to worsening of mechanical properties. Results of the maximum stress of samples exposed to biotic degradation compared with the stress values of reference samples have been presented in Fig. 9 and Table 6.

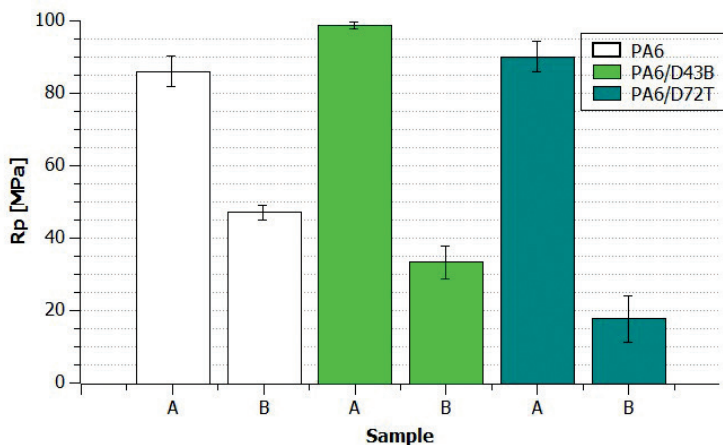


Fig. 9. Maximum stress (Rp) of reference samples (A) compared with the stress values of samples exposed to biotic degradation (B)

These results suggest that the maximum stress of PA6/OMMT composites after biotic degradation is significantly smaller than in the case of the reference material. It may be assumed that hydrolysis processes in the presence of organoclay play a key role. Adsorption and desorption processes in conditions of cyclic soaking and drying of the material are more intense than in the case of unmodified PA6. They lead to the creation of internal mechanical stress. Water helps microbes to develop and is advantageous for the enzyme reactions which

Table 6. Mechanical factors of PA6/MMT nanocomposites exposed to biotic degradation: ϵ_i – bending modulus, σ_{fc} – flexural bending stress, ϵ_{fM} – elongation area at the bending strength, σ_{fB} – the bending stress at failure, ϵ_{fB} – elongation area at the at failure

Sample	ϵ_i [MPa]	σ_{fc} [MPa]	ϵ_{fM} [%]	σ_{fB} [MPa]	ϵ_{fB} [%]
PA6	698	---	6.9	40	6.9
PA6/D43B	488	---	5.8	32	6.1
PA6/D72T	246	---	3.2	18	3.2

cause the microbiological degradation. Then, the influence of biological factors causes degradation of organic compounds into simple molecules in an enzymatic manner [24–28].

The variable results of static bending resistance may be caused by different chemical activity in which the components of the composite analysed serve as the feed for microbes or in which dissimilation processes occur, as a result of which the composite material is destroyed chemically by the substances produced and precipitated by microbes. Among the materials analysed, the sample with 3% content of clay organophilized with dimethyl benzyl ammonium salt had the highest Rp (flexural strength) value. The results clearly indicate that biological degradation of the polyamide sample was the smallest. The lowest resistance was noticed in the PA nanocomposite with montmorillonite organophilized with dimethyl ammonium salt with two long hydrocarbonic chains. The samples with organoclay were much more biologically active than the samples of pure polyamide 6. An increase in biological activity could be caused not only by the presence of a layered silicate, but also quaternary amphiphilic salt and a small amount of chloride ions. Long exposure of nanocomposite samples in a strongly biological environment has created an opportunity for organoclay to react/interact with organic and non-organic compounds present in soil. Apart from strictly biological factors, such as the presence of microbes, synergistic factors co-responsible for lowering of mechanical properties of those materials one may include rapid cyclic temperature changes, absorptivity of water and other compounds in soil.

From technological point of view, knowledge about biotic degradation is important as it may limit further applications in e.g. the building industry and infrastructure sector. After 475 days the mechanical properties of PA6/MMT nanocomposites decreased to values similar to those presented in Ref. [29], where PA6/clay nanocomposites were flushed in deionized water. For future developments, polyamide 6/clay nanocomposites could be considered as useful tools for the design of environmental or medical markers, capable of detection of specific interactions with biological species.

4. Conclusions

The influence of the three types of organoclays modified with dimethyl-benzyl-hydrogenated tallow ammonium and dimethyl-di(hydrogenated tallow) ammonium salt on the physical properties of polyamide 6 nanocomposites was studied. X-ray diffraction results showed that in the presence of each type of nanofiller formation of a stable α phase of PA6 is

facilitated, whereby the amount of γ phase is minimal and visible only after deconvolution. For all the materials compared, the presence of α phase was dominant, and the addition of phyllosilicates caused a reduction in the intensity of reflections derived from γ phase. The incorporation of any of the fillers did not affect glass transition temperature. Regardless of the testing temperature, composites reinforced with dimethyl-benzyl-hydrogenated tallow ammonium-modified montmorillonite were characterized by the highest mechanical strength. This can be explained by the presence of the aromatic ring in the structure of the quaternary salt that affects the stiffness of the composite structure and resulting flexural strength.

During the biotic degradation as a result of absorption and desorption of water under cyclic soaking and drying of the material, swelling and contraction leads to the formation of internal mechanical stresses. Pristine polyamide was the most stable material against biological degradation and the PA composite with 3% content of clay organophilized with dimethyl benzyl ammonium salt had the highest Rp value; one can point to the role of the organomodifier when analysing this effect.

The research (work) was supported by the European Union through the European Social Fund within "Cracow University of Technology development programme – top quality teaching for prospective Polish engineers; University of the 21st century" project (contract no.UDA-POKL.04.01.01-00-029/10-00).

References

- [1] Scala E.P., *A Brief History of Composites in the U.S.—The Dream and the Success*, Journal of the Minerals, Metals and Materials Society, 48(2), 1996, 49–48.
- [2] Pavlidou S., Papaspyrides C.D., *A review on polymer-layered silicate nanocomposites*, Progress in Polymer Science, 33, 2008, 1119–1198.
- [3] Ray S.S., Okamoto M., *Polymer/layered silicate nanocomposites: a review from preparation to processing*, Progress in Polymer Science, 28, 2003, 1539–1641.
- [4] Leszczyńska A., Njuguna J., Pielichowski K., Banerjee J.R., *Polymer/ montmorillonite nanocomposites with improved thermal properties Part II*, Thermochemica Acta, 454, 2007, 1-22.
- [5] Leszczyńska A., Njuguna J., Pielichowski K., Banerjee J.R., *Polymer/ montmorillonite nanocomposites with improved thermal properties Part I*, Thermochemica Acta, 453, 2007, 75.
- [6] Leszczyńska A., Pielichowski K., *Application of thermal analysis methods for characterization of polymer/montmorillonite nanocomposites*, Journal of Thermal Analysis and Calorimetry, 93(3), 2008, 677–687.
- [7] Pielichowski K., Leszczyńska A., *Polyoxymethylene-based nanocomposites with montmorillonite: oan introductory study*, Polimery, 2, 2006, 60–66.
- [8] Vaccari A., *Clays and catalysis: a promising future*, Applied Clay Science, 14, 1999, 161–244.
- [9] Kim N.H., Malhotra S.V., Xanthos M., *Modification of cationic nanoclays with ionic liquids*, Microporous Mesoporous Materials, 96, 2006, 29.



- [10] Morfis S., Philippoulos C., Papayannakos N., *Application of Al-pillared clay minerals as catalytic carriers for the reaction of NO with CO*, Applied Clay Science, 13, 1998, 203.
- [11] Christidis G.E., Scott P.W., Dunham A.C., *Acid activation and bleaching capacity of bentonites from the islands of Milos and Chios, Aegean, Greece*, Applied Clay Science, 12, 1997, 329.
- [12] Duleba B., Spišák E., Greškovič F., *Optimization of injection molding process by DOE*, Procedia Engineering, 96, 2014, 75–80.
- [13] Wu S., Jiang D., Ouyang X., Wu F., Shen J., *The structure and properties of PA6/MMT nanocomposites prepared by melt compounding*, Polymer Engineering & Science, 44, 2004, 2070–2074.
- [14] Kherrou D.E., Belbachir M., Lamouri S., Bouhadjar L., Chikh K., *Synthesis of Polyamide-6/ Montmorillonite Nanocomposites by Direct In-situ Polymerization Catalysed by Exchanged Clay*, Oriental Journal of Chemistry, 29(4), 2013, 1429–1436.
- [15] Stoeffler K., Utracki L.A., Simard Y., Labonte S., *Polyamide 12 (PA12)/clay nanocomposites fabricated by conventional extrusion and water-assisted extrusion processes*, Journal of Applied Polymer science, 130(3), 2013, 1959–1974.
- [16] Dintcheva Z.T., Filippone G., Arrigo R., La Mantia F.P., *Low-Density Polyethylene/ Polyamide/Clay Blend Nanocomposites: Effect of Morphology of Clay on Their Photooxidation Resistance*, Journal of Nanomaterials, vol. 2017, Article ID 3549475, 2017, 9, doi:10.1155/2017/3549475.
- [17] Rotter G., Ishida H., *FTIR separation of nylon-6 chain conformations: Clarification of the mesomorphous and γ -crystalline phases*, Journal of Polymer Science Part B, 30, 1992, 489.
- [18] Goderis B., Klein P.G., Hill S.P., Koning C.E., *A comparative DSC, X-Ray and NMR study on the crystallinity of isomeric aliphatic polyamides*, Progress in Colloid and Polymer Science, 130, 2005, 40–50.
- [19] Liu T.X., Liu Z.H., Ma K.X., Shen K.Y., He C.B., *Morphology, thermal and mechanical behavior of polyamide 6/layered-silicate nanocomposites*, Composites Science and Technology 63, 2003, 331–337.
- [20] Lincoln D.M., Vaia R.A., Wang Z.G., Hsiao B.S., *Secondary structure and elevated temperature crystallite morphology of nylon-6/layered silicate nanocomposites*, Polymer, 42, 2001, 1621.
- [21] Hanna A.A., *Thermal and dielectric properties of nylon 6*, Thermochimica Acta, 76, 1984, 97–103.
- [22] Wu Q., Liu X., Berglund L.A., *FT-IR spectroscopic study of hydrogen bonding in PA6/clay nanocomposites*, Polymer 43(8), 2002, 2445–2449.
- [23] Katoh Y., Okamoto M., *Crystallization controlled by layered silicates in nylon 6–clay nanocomposite*, Polymer, 50, 2009, 4718–4720.
- [24] Kaczmarek H., Bajer K., *Metody badania biodegradacji materiałów polimerowych*, Polimery, 51(10), 2006, 13–18.
- [25] Foltynowicz Z., Jakubiak P., *Poli(kwas mlekowy) – biodegradowalny polimer otrzymany z surowców roślinnych*, Polimery, 47, 2002, 11–12.

- [26] Pielichowski K., Majka T.M., Leszczyńska A., Giacomelli M., *Optimization and Scaling up of the Fabrication Process of Polymer Nanocomposites: Polyamide 6/Montmorillonite Case Study*; Springer-Verlag Berlin Heidelberg, in: Niuguna J. (Ed) *Structural Nanocomposites*, 2013, 75–103.
- [27] Majka T.M., Bartyzel O., Raftopoulos K.N., Pagacz J., Leszczyńska A., Pielichowski K., *Recycling of polypropylene/montmorillonite nanocomposites by pyrolysis*, *Journal of Analytical and Applied Pyrolysis*, 119, 2016, 1–7.
- [28] Majka T.M., Leszczyńska A., Kandola B.K., Pornwannachai W., Pielichowski K., *Modification of organo-montmorillonite with disodium H-phosphonate to develop flame retarded polyamide 6 nanocomposites*, *Applied Clay Science*, 139, 2017, 28–39.
- [29] Majka T.M., Pielichowski K., Leszczyńska A., *Wpływ chłonności płynów eksploatacyjnych stosowanych w motoryzacji przez kompozyty PA-6/MMT na ich właściwości mechaniczne*, *Czasopismo Techniczne*, 9-M/2012, 147–154.



Dmytro Hladyshv

Lviv Polytechnic, Department of Architectural Constructions

Roman Kinasz (rkinash@agh.edu.pl)

Faculty of Mining and Geoengineering, Department of Geomechanics, Civil Engineering,
AGH University of Science and Technology

LOCAL ACTION OF LOADING ON CONCRETE CUBES

MIEJSCOWE DZIAŁANIE OBCIĄŻENIA NA KOSTKI BETONOWE

Abstract

In the work, different approaches to projecting concrete elements have been considered, taking into account concrete deformation under tension in a casing around the crushing area at different levels of loading. The purpose of the investigations was the detection of the principal factors, which discern the taken empirical approaches to the computing models from the experimental character of tension strain distribution and concrete deterioration around the loading area. There have been determined, through an experiment, the lateral strains of the concrete tension along the perimeter of the side cube faces with the size of cube edges that equal to 300 mm with the formation of the cracks and up to the deterioration. The approach to calculations of the effect made by the samples of the above-mentioned type on local vertical loading with the reducing coefficient of their working conditions should be considered in the regulatory documents.

Keywords: local loading, stress, strain, deformation

Streszczenie

W artykule przedstawiono i porównano różne podejścia do projektowania elementów betonowych z uwzględnieniem odkształcenia betonu na rozciąganie w objętości wokół pola zgniecenia przy różnych poziomach obciążenia. Celem badań było określenie głównych czynników, które odróżniają podejścia empiryczne do schematów obliczeniowych dla częściowo obciążonych obszarów od eksperymentalnego charakteru rozkładu deformacji rozciągania oraz zniszczenia betonu objętości wokół powierzchni obciążenia. Doświadczalnie określono poprzeczne naprężenia rozciągające betonu na obwodzie bocznych powierzchni kostki o wymiarach żeber 300 mm przy powstawaniu pęknięć i do zniszczenia.

Słowa kluczowe: obciążenie miejscowe, naprężenie, odkształcenie, deformacja

1. Introduction

The local surface compression of concrete in structural elements of a square section is quite common in cases of the transmission of loading through metal centring spacers of different cross-sections from the above-placed elements onto the ones that are placed lower.

The probability of bearing failure under local loading of concrete elements has been practically proven, but in different normative documents, this nature of destruction is described by way of using various calculation designs. When calculating local loading on purely concrete elements according to the old Construction Norms and Regulations, SNIP 3], the following condition should be met:

$$F_{loc} \leq \psi R_{b,loc} A_{loc,1} \quad (1)$$

Norms [2–4, 10] use the same approach to the calculation, but they do not allow to assess the level of loading during the initiation and formation of vertical micro-cracks and macro-cracks.

The formation of vertical cracks in concrete should be considered as its transition typical condition, for practically this loading is the onset of the process of destruction of concrete in time and, in addition, subject to the conditions of its use. Therefore, the assessment of deformation characteristics of the early stage micro-crack formation in concrete and its boundary deformation during concrete macro-crack formation, with its transverse tension through compression by longitudinal forces, is a very topical task.

Of course, one should also pay attention to the nature of the kinematics of the crack formation and destruction of concrete structures during compression by longitudinal forces in the event of the transfer of loading onto them through the metal centred spacers of different sections.

2. Analysis of regulatory approaches and studies

The calculation of concrete elements for local bearing according to the normative literature is performed under the effect of compression force on the limited area taking account of the increased resistance of concrete to the compression within the area of bearing due to the formation of the volumetric stress state of the concrete that is under it. The issue of actual concrete tension deformation in a holder around the area of bearing at different levels of loading has not been sufficiently dealt with.

In the old norms [3], the calculation of the elements for bearing was performed using dependence (1). According to it, the main factor of influence adopted for heavy concrete of grade above B7,5 is the formula

$$\phi_b = \sqrt[3]{A_{loc2} / A_{loc1}} \leq 2,5 \quad (2)$$

which deals with the limited value of the root of the cubic relationship of the calculated reference area of bearing A_{loc2} to the area of direct bearing A_{loc1} . The same approach is offered by the valid norms [2, 4].

The height of concrete elements of a square section, which are included into the behaviour in transverse tension under the effect of centred local loading, are neither registered, standardised, nor limited, but they can vary within a significant range. Norms [2–4, 10] do not regulate the height or thickness of the concrete elements for local compression calculations.

While calculating the effect of local loading, evenly distributed over the area A_{c0} (Figure 1), the concentrated force of resistance $F_{Rdu,p}$ in pursuance of the valid norms [5], is established using the formula:

$$F_{Rdu} = A_{c0} f_{cd} \sqrt{\frac{A_{c1}}{A_{c0}}} \leq 3 f_{cd} A_{c0} \quad (3)$$

where: A_{c0} – is the area of loading; A_{c1} – is the maximal calculated area of distribution; f_{cd} – is the calculated value of concrete compression strength.

As we see, norms [5] in formula (2) continue to apply the approach to power F_{Rdu} through the calculated value of concrete compression strength, as the case is with norms [2–4], but they take into account the height h of the distribution of loading, and other additional new parameters and their limitations are introduced.

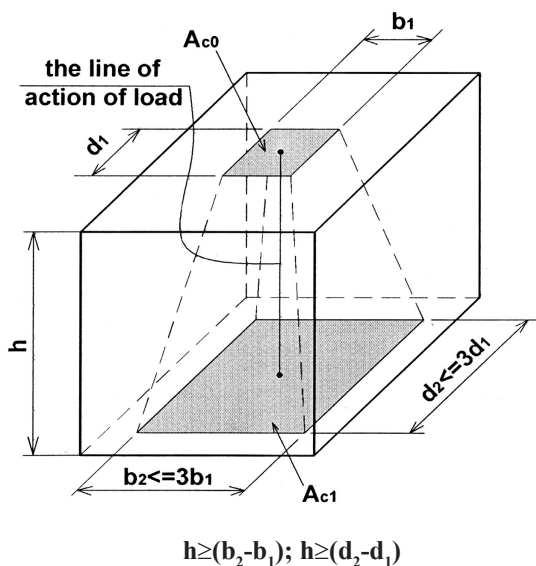


Fig. 1. Calculated distribution for partially loaded areas

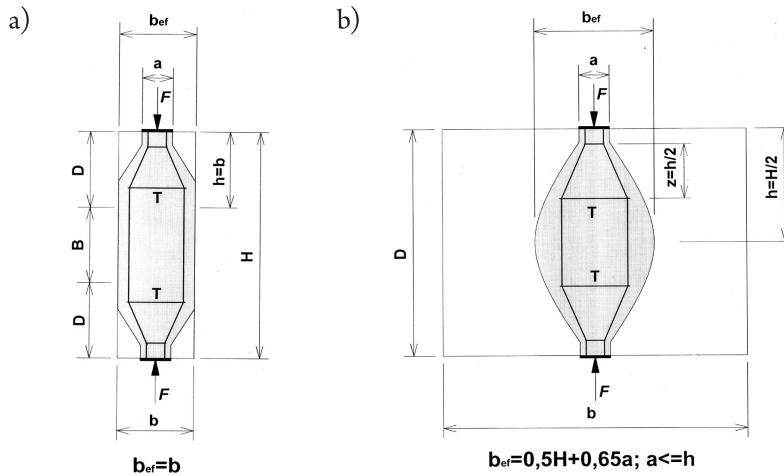


Fig. 2. Parameters for determining the transverse tension force T in compressed fields of concrete elements for designing the distributed reinforcement:
 B – area without discontinuity (homogeneous); D – area with discontinuity (heterogeneous)
 a – partial discontinuity; b – total discontinuity

The principle of taking crack formation into account is quite empirical. The tension strength T , which should provide for the appropriate transverse reinforcement, is taken from the value of the longitudinal force F according to the geometric parameters $a=b_1$ and $b \geq b_2$ or d_2 (same type designation in Fig. 1 and Fig. 2).

For areas with partial distribution $b \leq H/2$ (Fig. 2, a), use the formula:

$$T = 0,25 \times \left(\frac{b-a}{b} \right) \times F; \quad (4)$$

For areas with partial distribution $b > H/2$ (Fig. 2, b), use the formula:

$$T = 0,25 \times \left(1 - 0,7 \frac{a}{h} \right) \times F. \quad (5)$$

3. Purpose of the study and research tasks

As we see from the similar calculation designs of bearing listed in normative documents [2–4] and the design for partially loaded areas of bearing under the valid norms [5], they have relation to their calculated approach, when designing concrete elements. During the compression exerted by longitudinal forces, this especially refers to the assessment of characteristics of stage-wise and boundary deformation, crack formation and kinematics of crack formation, as well as to the destruction of concrete due to transverse tension in the holder around the area of loading.

The purpose of the present study is to identify the main factors that distinguish between traditional empirical approaches to calculation designs for partially loaded areas and the experimental nature of the distribution of tensile deformation, micro-crack formation and destruction of concrete of the holder around the area of loading.

The carried-out study aimed at identifying appropriate transverse tension deformation of concrete along the perimeter of side cube faces during the formation of micro- and macro-cracks to the destruction of various levels of loading on the partially loaded area in the volumetric concrete elements in the form of cubes with edges measuring 300 mm.

4. The results of the study

According to the set task, let us check the features of the behaviour for boundary conditions adopted in design [5] of the calculated distribution for partially loaded areas and the thereto corresponding formula (2).

It is necessary to observe the calculated boundary conditions:

- ▶ the relationship of the area of base A_{c1} to the upper area A_{c0} of the partially loaded area should be $A_{c1}/A_{c0} \geq 9$;
- ▶ the minimum height h should be no more than the double maximum amount of the area the loading area A_{c0} - $h \geq 2 \times b_2$ or $h \geq 2 \times d_2$.

Such conditions will accommodate the testing samples in the form of a cube (Fig. 3).

The manufactured concrete cubes (K1-3×3, K2-3×3) with edge dimensions of 300 mm (Fig. 3) are used to establish the influence of the nature of the tension stress distribution along the lateral faces of the volumetric concrete elements on the nature of their micro and macro-crack formation and destruction. The cubes are made of heavy concrete of class C18/22,5 (compression strength: the medium strength is $f_{cm,cube} = 28,5 \text{ MPa}$; the characteristic strength is $f_{ck,cube} = 22,5 \text{ MPa}$; calculated strength is $f_{cd} = 13 \text{ MPa}$).

The cubes were loaded from above applying uniformly distributed loading q_{loc} within a limited area $A_{c0} = 0,1 \times 0,1 = 0,01 \text{ m}^2$. The bottom face of the cube $A_{c1} = 0,3 \times 0,3 = 0,09 \text{ m}^2$ was lying tightly on a hard press plate (Fig. 4, 5).

While testing the cubes, the tensile strain was measured along their perimeter at three levels of 0,05 m, 0,15 m, and 0,25 m from the bottom face. All samples were tested under the application of centred loading. The loading was applied stage-wise up to the expected destructive force.

The characteristics and the analysis of the tested concrete cubes, applying methods [1] and [5] respectively, are shown in Table 1 and 2.

According to calculation methods [5], it is not quite clear how to assume the assured calculated value, which is less than the boundary value, of the component of formula (3). The boundary value is $3/2,08 = 1,442$ times bigger, than the average experimental value.

The fact that the destruction of the tested samples resulted not only due to bearing, but due to the combination of compression stresses (deformations) and bearing and splitting tension is evidenced not only by concrete horizontal tensile deformations measured in the course of testing, but also by the nature of crack formations during the destruction of tested samples (Fig. 4, 5).

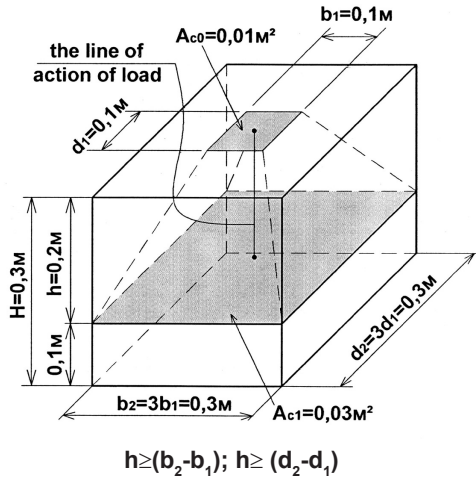


Fig. 3. Calculated distribution along the height h for the partially loaded area A_{s0} that corresponds to the boundary conditions of the calculated design shown in Fig. 1 and to boundary limit of formula (2)

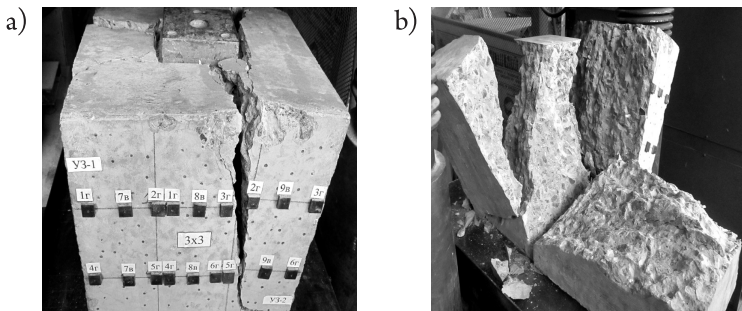


Fig. 4. The nature of crack formation (a) and destruction (b) under locally applied loading on the concrete cube K1-3x3 ($A_{s0}=0,01 \text{ m}^2$ $A_{c1}=0,0425 \text{ m}^2$)

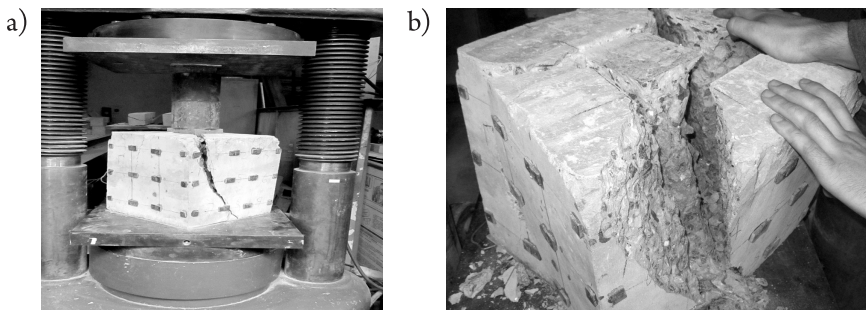


Fig. 5. The nature of crack formation (a) and destruction (b) under locally applied loading on the concrete cube K2-3x3 ($A_{s0}=0,01 \text{ m}^2$ $A_{c1}=0,0441 \text{ m}^2$)

Table 1. Typical data for the concrete cubes when using methods [1]

Mark sample	$l = b = h, \text{ m}$	$A_{loc1}, \text{ m}^2$	$A_{loc2}, \text{ m}^2$	$\sqrt[3]{A_{loc2} / A_{loc1}}$		$F_u, \text{ kN}$	$R_{m.cube}, \text{ MPa}$	The calculated value		$N_{m.cub}, \text{ kN}$	$N_{kubum}, \text{ kN}$	N / F_u
				$\varphi_{b,f}$	φ_{kubum}			$N_{b,f}, \text{ kN}$	$N_{kubum}, \text{ kN}$			
K1-3x3	0,3	0,01	0,09	2,08<	2,5	800	28,5	270,4	325	592	713	0,338
K2-3x3	0,3	0,01	0,09	2,08<	2,5	900	28,5	270,4	325	592	713	0,338

Heavy concrete class: B 22,5 [5] (compression strength: the medium value is $R_{m.cube} = 28,5$ MPa; the normative prism value is $R_{b,ser} = 16,75$ MPa; the calculated value is $R_b = 13$ MPa; the value of tension is $R_{bt,ser} = 1,5$ MPa; $R_{bt} = 0,975$ MPa).

Table 2. Typical data for the concrete cubes when using methods [5]

Mark sample	$l = b = h, \text{ m}$	$A_{c0}, \text{ m}^2$	$A_{c1}, \text{ m}^2$	$\sqrt{A_{c1} / A_{c0}}$		$F_{pr}, \text{ kN}$	$f_{m.cub}, \text{ MPa}$	The behaviour value		$F_{R,ducf}, \text{ kN}$	$F_{R,duit}, \text{ kN}$	$F_{R,dui}, / F_u$
				fact	ultum			$F_{R,ducf}, \text{ fact}, \text{ kN}$	$F_{R,ducf(3)ultum}, \text{ kN}$			
K1-3x3	0,3	0,01	0,0425	2,06<	3	800	28,5	268	390	588	855	0,335
K2-3x3	0,3	0,01	0,0441	2,10<	3	900	28,5	273	390	598	855	0,303

Heavy concrete class: C18/22,5 [4] (compression strength: the medium value is $R_{m.cube} = 28,5$ MPa; the characteristic value $- f_{ck.cube} = 22,5$ MPa; the calculated value is $f_{cd} = 13$ MPa).

The actual nature of the distribution of tension strain at three levels of height h and along the perimeter of the lateral faces of cubes of the partially loaded area is shown in Figures 6, 7, and 8.

It should be noted that while testing the samples for a central application of force onto a fixed area, uneven distribution of horizontal tension strain $\epsilon_{bt} \times 10^5$ according to the perimeters at three levels along the height h (Fig. 6, 7, and 8) is observed.

Tension strain values at the top - first level of the lateral faces in all tested samples were bigger, than at the second and third levels closer to their lower carrying face (Fig. 10, 11).

The diagrams of unit strains of concrete under the application of the die block according to the stress under it (Fig. 11) show that the curve has three areas of stress change σ under the die block from the corresponding forces $F_{R,dui} (N_p)$: the first one – from 0 MPa to 30 MPa, the second one – from 30 MPa to 80 MPa and the third one – from 80 MPa to F_u MPa, which describe the transitional stage-wise phases of the concrete stress-deformed state.

The comparison of the two curves of dependence (Fig. 11) of the stress under the side block on the unit strain of the pushing pyramids with the equal height (the first dependence

at $h=H=0,3$ m and the second dependence at the unit height $h=2b_1=0,2$ m) showed that, compared with the full diagram of concrete compression according to the characteristic parameters, the first curve corresponds better to the full concrete compression diagram, graphed using methods [1].

The stress values $\sigma \gg 30$ MPa in the typical transition point between the first and the second areas describes the beginning of the micro-crack formation in the concrete of cubes and is characterised by unit strains of $\epsilon_{bt}=10 \times 10^{-5}$ (Fig. 6). The stress value of $\sigma \gg 30$ MPa, thereby, coincides practically with the average stress calculated value according to data [1] $\sigma_1 = N/A_{loc.1} = 270,4/0,01 = 27,04$ MPa (Table. 1) and the processing of experimental data according to formula (2) of norms [5] $\sigma_2 = \sigma_1 = F_{Rdt}/A_{c0} = (268+273)/2/0,01 = 27,05$ MPa (Table 2).

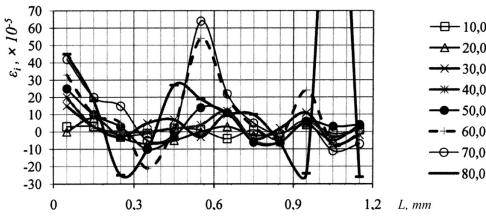


Fig. 6. Unit strain along the perimeter of a cube at the level of 0,25 m from its base

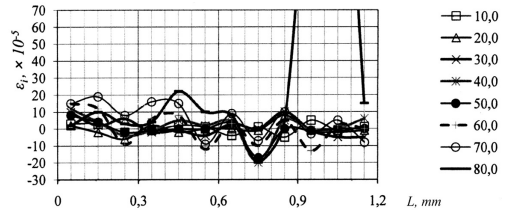


Fig. 7. Unit strain along the perimeter of the cube at the level of 0,15 m from its base

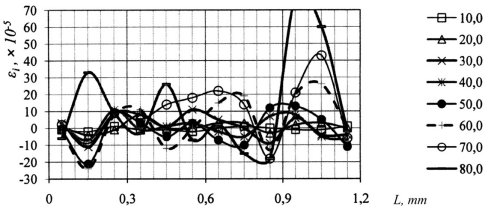


Fig. 8. Unit strain the perimeter of the cube at the level of 0,05 m from its base

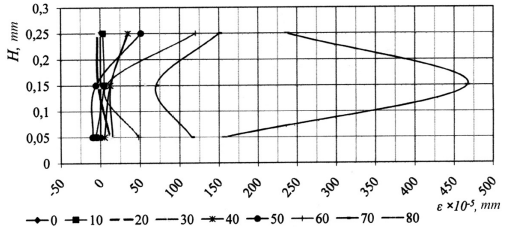


Fig. 9. The total strain along the perimeter of the cube at three levels from its base (0,25 m, 0,15 m, and 0,05 m)

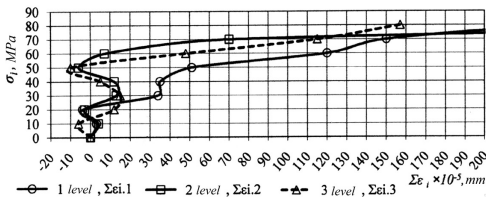


Fig. 10. Dependence of total strain along the perimeter of the cube at three levels from its base (0,25 m, 0,15 m, and 0,05 m) from the stresses under the die

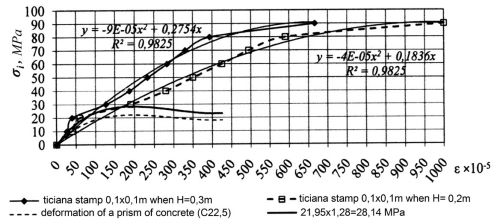


Fig. 11. Dependence of stress under the on the relative unit strain of the pushing pyramid with the cube height of $H=0,3$ m and the relative height of $h=2b_1=0,2$ m compared with the full concrete compression diagram

The stress value $\sigma \gg 80$ MPa in the typical transition point between the second and the third areas describes the beginning of the macro-crack formation in concrete cubes and is characterised by unit strains of $\varepsilon_{bt} = 400 \times 10^{-5}$ (Fig. 6) [1]. The stress value of $\sigma \gg 80$ MPa, thereby, is practically approaching the boundary values, both according to norms [3] $\sigma_{1,u} = N_{b,ultum} / A_{loc.1} = 713 / 0,01 = 71,3$ MPa (Table 1) and according to norms [5] $\sigma_2 = \sigma_{1,u} = F_{Rdu,ult} / A_{c0} = 855 / 0,01 = 85,5$ MPa (Table 2).

The uniform transmission of bearing loading on the top face of the concrete samples increases, without doubt, their actual carrying capacity, but the values of destructive loading, obtained during the tests, can be considered somewhat bigger compared with those, which may be in natural structures. This is due primarily to the fact that during the study, a rather high rigidity base was used (bottom steel press plate) under the bottom face of the cubes, which is difficult to achieve under actual conditions of construction.

5. Conclusions

With reference to the nature of the destruction of the tested concrete cubes under the effect of local vertical loading along their geometrical axis, the following was revealed:

- ▶ the processed experimental data on concrete cube samples, tested for the effect of local loading, revealed, according to the assessment of the concrete stress-strain state through the fixed levels of tensile strain, a stable value of the calculated concentrated forces of the levels of crack formation and destruction of the cube samples;
- ▶ destruction is due to bearing-cracking of concrete, starting from the top face under the fixed area of the partially loaded area, where, due the vertical and transverse strain, the volume of the incomplete pushing pyramid increases, from which pyramid the longitudinal crack formation and splitting of the cube body are running;
- ▶ the bottom face of the incomplete pushing pyramid, which bottom face is running along the full height of the cube, covers an area on the lower bearing face of the cube, which is less than the area of the cube;
- ▶ the analysis of the stress curves under the die block depending on the unit strains of the pushing pyramids equal to the height with the full diagram of the concrete compression strain made it possible to determine that the transition of the cube concrete to the micro-crack formation is carried out at the unit strain values of $\varepsilon_{bt} = 10 \times 10^{-5}$, and to the macro-crack formation and gradual physical destruction is carried out at the relative boundary deformation values of $\varepsilon_{bt} = 400 \times 10^{-5}$ according to the characteristic parameters of the full concrete compression diagram;
- ▶ the results of comparing the test data with the data of the calculation using the methods of the old Construction Norms and Regulations, SNIP [3] and the current standard DSTU [4] showed that the derived calculated and behaviour values of centred forces, listed in Tables 1 and 2, are practically similar $N = F_{Rdu} = 270$ kN;
- ▶ as to calculation methods [5], it is not completely clear, how to take the assured, lower value for the boundary the calculated value of the component of formula (2), or

should the boundary value be taken right away. Then, in view of the data obtained during testing the cubes, the relationship of concentrated forces is $N/F_{Rdu} = 270/390 = 0.69$ (calculated forces [3] to behaviour forces [5]). Other data are listed in Tables 1 and 2.

- ▶ the approach to calculating the effect of local vertical loading on the samples of the dealt with type should be taken account of in [5] taking into consideration the nature of their destruction and the hardness of the base under their bottom face using the reduction factor of their behaviour conditions.

The nature of the crack formation, while changing the stress-strain state in flat concrete elements under symmetrical loading in the conditions of changing their height with other parameters remaining constant, is shown in [6].

The thesis was conducted within the framework of the regulation researches of AGH University of Science and Technology, Poland No 11.11.100.197.

References

- [1] Bambura A.N., Bachinskiy V.Ya., *Guidelines for the refined calculations of reinforced concrete elements, taking into account the full diagram of concrete compression*, NIISK of the USSR State Committee for Construction, Gosstroy, Kyiv 1987.
- [2] Code of Regulations, SP 52-101-2003. Concrete and reinforced concrete structures without reinforcement pre-strain.: Russian State Committee for Construction, Moscow 2004.
- [3] Construction Norms and Regulations, SNIP 2.03.01-84. Concrete and reinforced concrete structures. Design standards. TSITP of the USSR State Committee for Construction, Gosstroy, Moscow 1986.
- [4] Construction Standards of the Republic of Belarus, SNB 5.03.01-02. Concrete and reinforced concrete structures, Ministry of Architecture and Construction of Belarus, Minsk 2003.
- [5] DSTU B V.2.6-156: 2010. Concrete and reinforced concrete structures made of heavy concrete. Design Rules, Ministry of Regional Development, Kyiv 2011.
- [6] Hladyshev D.N., *The local effect of stress on concrete elements // Resource-saving materials, structures, buildings and edifices*, Collection of scientific papers National University of Water Management, Issue 29, Rivne 2014, 94–101.
- [7] Knauff M., *Obliczanie konstrukcji żelbetowych według Eurokodu 2: zasady ogólne i zasady dotyczące budynków*, Wydawnictwo Naukowe PWN, Warszawa 2015.
- [8] Łapko A., *Projektowanie konstrukcji żelbetowych: wg Eurokodu 2 i PN-B-03264:1999*, Arkady, Warszawa 2000.
- [9] Manual for designing foundations on a natural basis for columns of buildings and structures (Addendum to the SNiP 2.03.01-84 and SNiP 2.02.01-83), Lenpromstroyproekt of the USSR State Committee for Construction, Gosstroy – TSITP of the USSR State Committee for Construction, Gosstroy, Moscow 1989.
- [10] PN-EN 1992-1-1:2008/NA:2010 Eurokod 2: Projektowanie konstrukcji z betonu.

Remigiusz Wojtal (rwojtal@pk.edu.pl)

Institute of Road & Railroad Engineering and Transport, Faculty of Civil Engineering,
Cracow University of Technology

Laurence R. Rilett

Keith W. Klaasmeyer Chair in Engineering and Distinguished Professor, Department
of Civil Engineering, University of Nebraska-Lincoln, Lincoln, NE, USA

CALIBRATION OF MICROSIMULATION MODELS OF ADVANCE DETECTION AND WARNING SYSTEMS AT SIGNALIZED INTERSECTIONS

KALIBRACJA MODELI MIKROSYMULACYJNYCH SYSTEMÓW WCZESNEJ DETEKCJI I OSTRZEGANIA NA SKRZYŻOWANIACH Z SYGNALIZACJĄ ŚWIETLĄ

Abstract

This paper presents a calibration and validation procedure for microsimulation models, which used metrics (mean, variance, and mean absolute percentage error) with statistical tests (t-test, Shapiro-Wilk, Kolmogorov-Smirnov and Wilcoxon), to compare empirical and simulated data. A genetic algorithm was used to identify calibration parameters set. The paper justifies the approach using VISSIM microsimulations to analyze two safety countermeasures: Advance Detection System and Advance Warning System, which can be applied at signalized intersection. The end result was a calibrated and validated model, which could be used to compare different safety countermeasures at rural signalized intersections in the state of Nebraska (USA). The proposed approach could be utilized in similar studies.

Keywords: signalized intersection, safety countermeasures, calibration, validation, simulation model

Streszczenie

W artykule przedstawiono procedurę kalibracji i walidacji modeli mikrosymulacji z użyciem mierników (średnia, wariancja i średni bezwzględny błąd procentowy), oraz testów statystycznych (studenta, Shapiro-Wilka, Kolmogorow-Smirnowa oraz Wilcoxon) dla porównania danych empirycznych i z symulacji. Algorytm genetyczny wykorzystano do doboru parametrów kalibracyjnych. Metodologia została sprawdzona w programie VISSIM do analizy dwóch systemów poprawy BRD: Systemu Wczesnej Detekcji i Wczesnego Ostrzegania, które stosuje się na skrzyżowaniach z sygnalizacją świetlną. W wyniku uzyskano skalibrowany i zweryfikowany model, który można użyć do porównania środków poprawy BRD na zamiejskich skrzyżowaniach z sygnalizacją w stanie Nebraska (USA). To podejście można wykorzystać w podobnych analizach.

Słowa kluczowe: skrzyżowanie z sygnalizacją, środki poprawy brd, kalibracja, walidacja, model symulacyjny

1. Introduction

The main advantage of using microsimulation tools to examine potential traffic safety improvements is their ability to handle the dynamic and stochastic nature of traffic. However, it is critical that the user calibrates and validates the models for the particular site and safety countermeasure being studied. Traffic micro-simulation models are very useful tools for simulating the effects of proposed network improvements, and have the potential to model a subset of safety countermeasures. Typically, these countermeasures are operation-based, e.g. an Advance Detection System (ADS) and an Advance Warning System (AWS).

The key features of VISSIM include driver and vehicle behavior management, a comprehensive toolbox for signal control that allows the user to define signal control logic, and flexibility in collecting disaggregated data. The basic VISSIM model contains a set of default parameter values related to the driver, the vehicle, and the system. However, it is crucial that these default values are adjusted for the specific problem being studied to ensure that the simulation output reflects reality. As these parameters directly affect modeled vehicle interactions, failure to properly calibrate them can result in erroneous conclusions. The calibration process for safety-related studies should include car-following, lane-changing, and signal control parameters.

This paper is focused on the calibration of a VISSIM microsimulation model (PTV, version 5.30, 2011) that will be utilized to examine ADS and AWS safety countermeasures. The model was calibrated for four high-speed, isolated, rural intersections in Nebraska (USA). Due to spatial constraints of the current paper, the process will be described with respect to a selected single location, the intersection of US-77 and Pioneers Boulevard outside of Lincoln, Nebraska. The calibration procedure utilized a genetic algorithm (GA). The GA procedure was selected because, by definition, it examined the entire solution space, and was less likely than other optimization techniques to identify a local minimum. Due to the stochastic nature of the model, nonparametric statistical techniques were also incorporated within the GA to ensure that empirical and simulated data were statistically identical. Because the study was safety-related, it was imperative that the resulting model adequately replicates the distribution of speeds, rather than the average speed only. Therefore, speed distribution was used as the primary calibration measure of effectiveness. Subsequently, the calibrated model was validated to ensure that the simulated speed distributions at key points matched the empirical distributions, that the simulated queues matched the observed queues, and that the simulated delays matched the observed delays. These comparisons were made using appropriate statistical tests. The end result was a calibrated and validated model for Nebraska that replicated driver behavior and could be utilized to compare various safety countermeasures at isolated, rural, high-speed signalized intersections.

2. Methodology

In recent years, a number of studies have been conducted in relation to the development of systematic calibration methodologies. Many of these studies were based on statistical comparisons and utilized GAs to identify potential parameter sets [3, 7, 8, 9, 11]. In addition, many made use of ITS data, which are becoming more widely available.

Mathew et al. [9] proposed heuristics and the GA based optimization for model calibration. The calibration parameters were identified through sensitivity analysis. The optimum values of these parameters were obtained by minimizing the error between the simulation and field delays using the GA technique. The authors found that the time and effort required to accurately tune a large number of potential simulation parameters could be reduced through the use of optimization methods. Park et al. [11] proposed an innovative calibration and validation procedure, successfully applying the approach to several case studies. The first step of the procedure required the identification of key calibration parameters and their acceptable ranges. The generation of a reasonable number of parameter sets using a statistical experimental design was then performed. Each parameter set was run five times to test the statistical feasibility of each set. The GA optimization program obtained an optimal calibration parameter set from the potential parameter ranges that were accepted during the feasibility test step. Since VISSIM is a microscopic and stochastic simulation model, a small number of runs was conducted for each feasible parameter set to reduce variability. An objective function of the GA was obtained through the comparison of field data to simulation output. A recent study indicated that an automatic calibration procedure could be more effective [7] than a manual approach. The authors used the GA procedure to determine ideal parameters. Five steps were proposed for the approach, and the procedure was iterative. The first step was the initialization of the GA. Next, the microscopic simulation model was run with the input file (generated parameters were translated into the appropriate VISSIM format). The model output and selection of the potential parameter set was then evaluated. This was the most important component in the calibration procedure. The model output was evaluated using statistical tests (Moses's test, the Wilcoxon test, and the Kolmogorov-Smirnov test). Finally, two descriptive statistics (median and dispersion) and the maximum difference in the cumulative function were tested using nonparametric testing methods. Cunto et al. [3] utilized a calibration procedure consisting of four steps. Their procedure included a heuristic selection of the initial model inputs; statistical screening using a Plackett–Burnman with factorial design; the development of a linear expression relating significant model inputs to safety performance (fractional factorial analysis); and the GA procedure to obtain best estimate model parameters. The next attempt to determine a formal calibration procedure was conducted by Park et al. [10], who implemented an experimental design. They argued that this was appropriate because the number of controllable feasible parameter combinations was so large that the set of possible scenarios could not be evaluated in a reasonable amount of time. This problem was compounded if multiple simulation runs were required to reduce stochastic variability. The authors also used statistical tests, including the t-test and the Kolmogorov–Smirnov test, to determine how well their calibration procedure performed.



Calibration is a process used to determine an appropriate set of model parameters that replicate observed information, such as local driver behavior. In general, the user adjusts the select behavioral parameters until the simulated output matches the empirical data at some predefined statistical level. Validation is then employed to ensure that the calibrated model is appropriate by determining whether the output of the model accurately reproduces the specified behavior. Similar to calibration, validation compares empirical data with simulated data. However, this empirical data need to be different than that used in the calibration.

The objective of the current paper was to demonstrate the calibration and validation of a VISSIM microsimulation model for signalized intersections in Nebraska. The model will be utilized to study ADS and AWS systems, which are engineering countermeasures intended to improve road safety. As such, it was critical that the model accurately reflects driver behavior, as measured by vehicle speeds, when these systems were active. A microsimulation model was developed for four Nebraska test sites. Key inputs included traffic volume, turning movements, and heavy vehicle percentages. The models were then calibrated based on speed distributions at particular points at each location. The calibrated model was then validated by examining speed distributions at other locations and waiting times on the minor approaches.

Fig. 1 displays the calibration procedure adopted in the current study, which is described more thoroughly in subsequent sections.

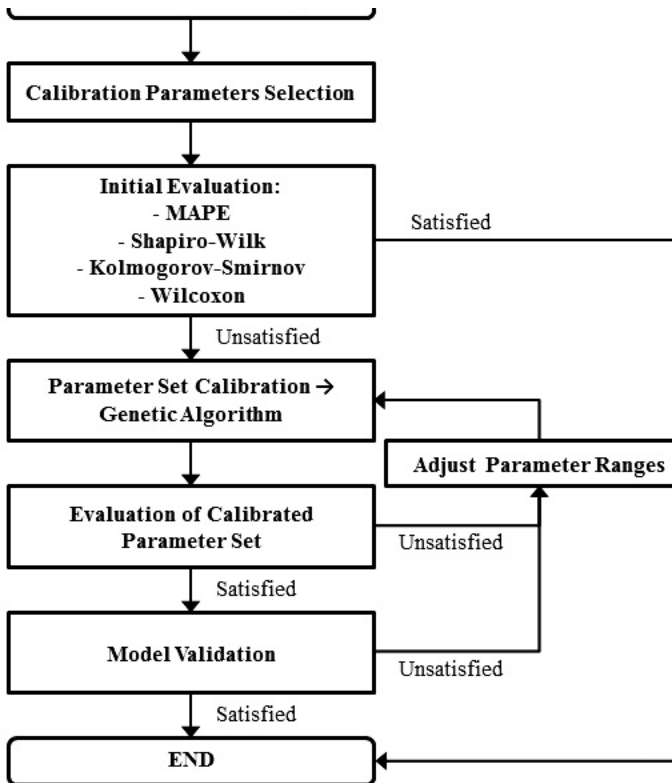


Fig. 1. Procedure for model calibration

3. Calibration procedure

The procedure of the calibration, presented in the current paper, was applied to four intersections in Nebraska: (1) US Highway 77 and Saltillo Rd in Lincoln, (2) US Highway 77 and Pioneers Boulevard in Lincoln, (3) Highway N-133 and Highway N-36 in Omaha, and (4) US Highway 75 and Platteview Road in Bellevue. However, based on space limitations of the current document, the US-77 and Pioneers Blvd test site was treated as the primary focus of the current paper.

3.1. Simulation model setup

The intersection of the US-77 and Pioneers Blvd was located in a rural area approximately five miles south of Lincoln. US Highway 77 is a four-lane divided expressway, as pictured in Fig. 2.

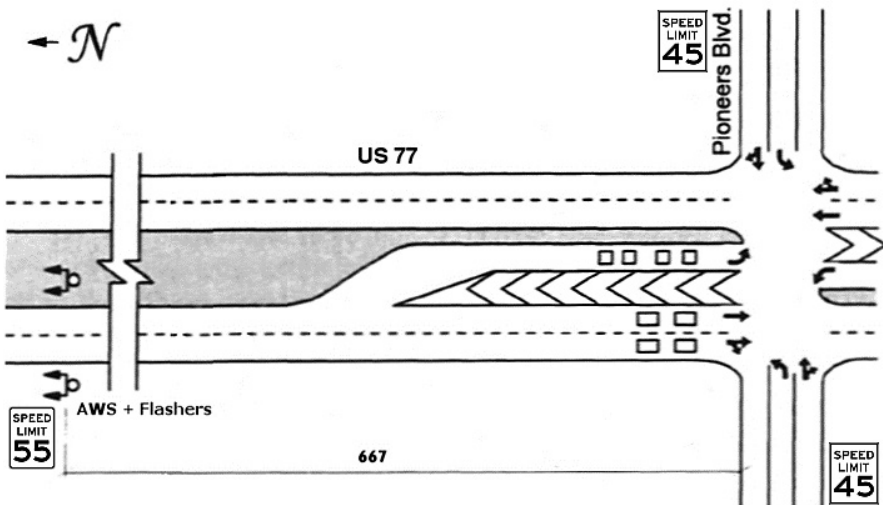


Fig. 2. Study site – US-77 & Pioneers Blvd in Lincoln, NE

The speed limit on US-77 before the intersection was 55 mph, and on Pioneers Blvd was 45 mph in both directions. The SB approach of US-77 was examined because it had been outfitted with an AWS by the Nebraska Department of Roads (NDOR) [14]. This approach had an exclusive left-turn lane, one through lane, and one shared through and right-turn lane. All lanes were 12 ft wide, while the left-turn lane was 180 ft long. The lane configuration for this approach, with detector locations, is pictured in Fig. 2. Additionally, the EB and WB approaches on Pioneers Boulevard had an exclusive left-turn lane and shared through with the right-turn lane. The advance warning sign, which operates in pulse mode, was located 667 ft before the intersection, as shown in Fig. 2. Stop line detection was provided for all approach lanes, and the detectors operated in presence mode. The SB approach on US-77 had one phase, which included all movements. It had a 15.0 sec minimum green, a 2.0 sec passage time, a 50.0 sec maximum green, a 4.5 sec yellow, and a 0.5 sec all-red interval.

This intersection was coded in VISSIM software and the plan and overall geometry of the intersection were based on a scaled map obtained from Google [5]. Lane width and length were confirmed on site. Additional data pertaining to the intersection, including signal timing and the location of the safety countermeasures (AWS and detectors), was obtained from materials provided by NDOR. Empirical traffic data, including volumes, turning movement counts, and heavy vehicles percentages, were collected by the authors. All VISSIM parameters were initially set to the default values. The signal control logic was coded using vehicle actuated programming (VAP) to enable the phase-based, traffic-actuated signal control logic to be implemented in the microsimulation [13].

Due to the nature of the traffic safety countermeasures being analyzed, it was critical that the speed distribution of drivers at critical locations be adequately modeled. Therefore, the empirical data collection effort focused on speed distribution as a function of space and time. The speed and location of every northbound vehicle were collected using wide area detectors (WAD) mounted on a mobile trailer. These devices were connected to programmable controllers. Three WADs were used to collect data: two Wavetronix SmartSensor Advance models and one Wavetronix SmartSensor HD. The two Wavetronix SmartSensor Advance models were used to track approaching vehicles upstream and downstream of the trailer location. The sensors recorded distance and speed at 0.1 sec intervals. The Wavetronix SmartSensor HD was used to record vehicle information equidistant with the pole location [1]. More detailed information on the specific capabilities of these devices can be obtained from the Wavetronix website [6]. The data collection scheme, including the location of each WAD and their coverage (up to 600 ft upstream and downstream), is illustrated in Fig. 3.

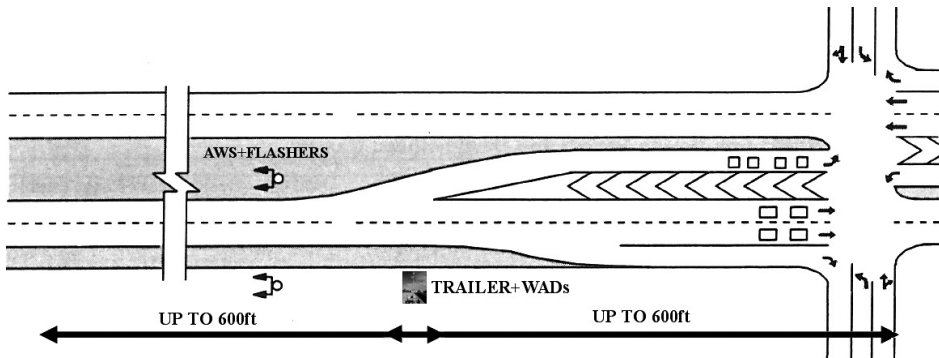


Fig. 3. WADs data coverage [14]

In calibrating a microsimulation model, it is critical that the end application is known so that key metrics are appropriately modeled. For example, for safety applications, the distribution of metrics, such as speed, is often required; however, if the model is calibrated to mean speed, it may fail to perform in the desired manner. The motivation behind the current paper was to analyze two safety countermeasures (ADS and AWS). A complete description of these systems can be found elsewhere [14].

To model the ADS countermeasure, two critical locations were identified, as illustrated in Fig. 4. The first location, denoted as DS1, was 2,000 ft upstream of the stop bar. This location was selected because the status of the traffic signal was deemed to have no effect upon observed speed. The second location, denoted as DS2, was 600 feet from the stop bar. This location was selected in order to model driver behavior in the vicinity of the intersection as a function of traffic signal status.

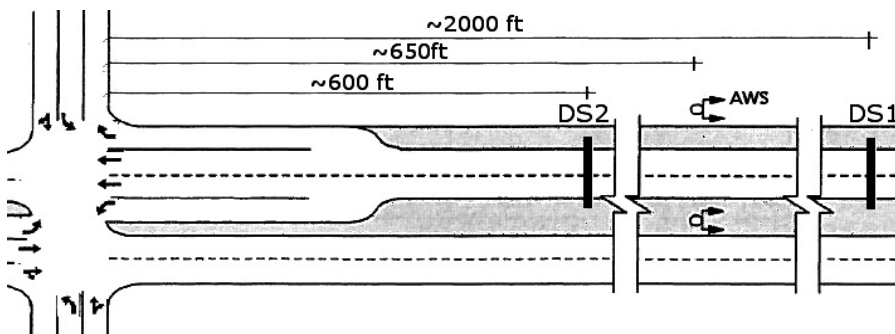


Fig. 4. Location of desired speed decision points [14]

Empirical speed distribution data were collected at both locations, and are presented in Fig. 5. Distribution DS1 corresponded to location DS1, and distribution DS2a corresponded to location DS2. In order to model speed change, a decision point was input in VISSIM. In essence, the desired speed distribution was added as part of the input parameter set. The simulated vehicles were then assigned a speed from this input distribution when they crossed the point. In this way, the behavior of drivers slowing down as they approached the intersection was modeled.

There were three key points to note from Fig. 5. First: because the ADS could not be detected by drivers in real life, it was not necessary to model driver reactions to that system in VISSIM. Second, users are able to add as many decision points as they wish to simulate driver deceleration upon approaching an intersection. While two decision points were adequate for the four test sites analyzed in the current project, other test sites may require additional decision points. Third, the distribution is defined by the user. In the current paper, five points were used to define the curve. However, users may choose as many points as desired in order to model the desired speed distribution.

In essence, the AWS simulation followed ADS logic. The primary difference was that the AWS utilized an AWS device (i.e., a sign with the flashers). Because drivers react to active AWS signs, it was necessary to model the reactions of simulated drivers to the AWS sign in VISSIM. In Nebraska, the recommend sign location was 650 ft from the intersection stop bar. The flasher activated (i.e., began to flash) a few seconds (typically 7 sec, but 8 sec in this case) before the traffic signal transitioned from green to yellow. Any driver upstream of the AWS sign was then in a position to slow and come to a gradual stop. To model this component of driver behavior in VISSIM, two speed distributions were utilized at location DS2: DS2a and DS2b. When AWS flashers were inactive (i.e., from the end of the red phase until 8 sec prior to the transition from

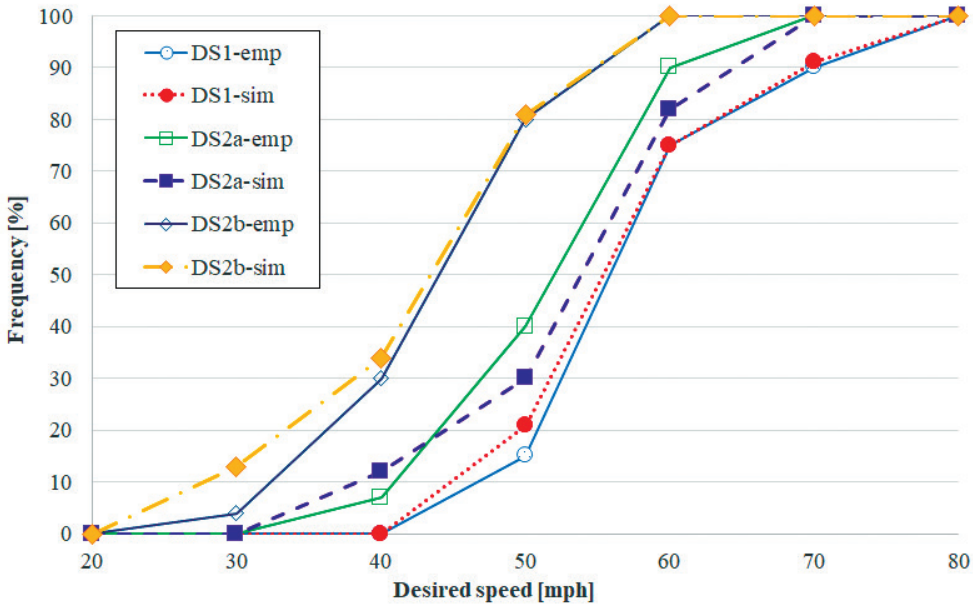


Fig. 5. Empirical and simulated speed distributions at the DS1 and DS2 locations

green to yellow), vehicles crossing DS2 followed the speed distribution DS2a, as detailed in Fig. 5. This was the same speed distribution followed by all vehicles in the ADS simulation. When the AWS flashers were active (i.e., eight 8 sec prior to yellow phase and then during the yellow and red phases) drivers followed the adjusted speed distribution DS2b, as shown in Fig. 5. In this latter situation, each vehicle that crossed location DS2 was assigned a new speed from the speed distribution DS2b, and was traveling, on average, at the slower rate [13]. Therefore, when the AWS was active, simulated drivers traveled, on average, at a slower rate compared to periods when the AWS was off (i.e., during the green signal and until 8 sec prior to the yellow signal). It was further assumed that a change in speed would occur at DS2. Note that the model could be easily adjusted to add additional decision points to reflect a more gradual change in driver speeds. Additionally, when the system was active, change in speed distribution implied that a percentage of vehicles potentially sped up to beat the yellow signal. Note that the percentage of drivers that exhibited this behavior was controlled by the user through the input speed distribution. However, red light running is not modeled in VISSIM, so therefore the number of vehicles entering the intersection during the red signal was precluded from being an MOE in the current analysis.

3.2. Selection of calibration parameters

Calibration was carried out for the parameters of car-following, lane change, desired speed distribution, and signal control [13]. The initial set of VISSIM parameters utilized in the calibration was identified and selected based upon engineering judgment and the review of the salient literature [3, 4, 9–12]. All 19 parameters and their acceptable ranges, identified in the literature review, are presented in Table 1.

Table 1. Set of VISSIM parameters for the initial evaluation and calibration

BEHAVIOR	PARAMETER		DEFAULT VALUE	ACCEPTABLE RANGE	UNIT
CAR FOLLOWING	Number of observed preceding vehicles	NUMB_PRECED	2	1 – 4	-
	Maximum look ahead distance	OBS_DISTANCE MAX	820.21	700 – 900	ft
	CC0 (Standstill Distance)	CC0	4.92	3.0 – 6.0	ft
	CC1 (Headway Time)	CC1	0.9	0.5 – 3.0	s
	CC2 ('Following' Variation)	CC2	13.12	1.0 – 30.0	ft
LANE CHANGE	Waiting time before diffusion	T_DISAPPEAR	60.0	30 – 90	s
	Minimum headway (front/rear)	MIN_LC_GAP	1.64	0.5 – 3.0	ft
DESIRED SPEED 1	Frequency at 50mph	DESIRED_SPEED 1	0.150	0.10 – 0.30	-
	Frequency at 60mph	DESIRED_SPEED 1	0.750	0.60 – 0.80	-
	Frequency at 70mph	DESIRED_SPEED 1	0.900	0.85 – 0.95	-
DESIRED SPEED 2a	Frequency at 40mph	DESIRED_SPEED 2	0.070	0.05 – 0.15	-
	Frequency at 50mph	DESIRED_SPEED 2	0.400	0.30 – 0.50	-
	Frequency at 60mph	DESIRED_SPEED 2	0.900	0.80 – 0.95	-
DESIRED SPEED 2b	Frequency at 30mph	DESIRED_SPEED 3	0.040	0.03 – 0.15	-
	Frequency at 40mph	DESIRED_SPEED 3	0.300	0.25 – 0.45	-
	Frequency at 50mph	DESIRED_SPEED 3	0.800	0.70 – 0.90	-
SIGNAL CONTROL	Reaction to amber signal: α	AMBER_ALPHA	1.59	1.0 – 15.0	-
	Reaction to amber signal: β_1	AMBER_BETA1	-0.26	- 0.40 – -0.20	-
	Reaction to amber signal: β_2	AMBER_BETA2	0.27	0.10 – 0.30	-

The default values of all parameters were utilized in the initial evaluation of the simulation model. The definitions of these parameters and their functions in VISSIM can be obtained from the VISSIM manual [13], as well as from a number of papers describing VISSIM model calibration [3, 4, 9–12].

3.3. Initial evaluation

The primary goal of the initial evaluation was to determine whether the default model (e.g. based on the default values of the parameters) was able to adequately model real traffic conditions at the test intersection. If the output did not match the empirical data, it became necessary to conduct additional steps in the calibration procedure.

A calibration procedure was performed to obtain a set of driving behavior parameters that would result in simulation results similar to the observed empirical values. As discussed previously, the procedure involved the use of a GA and a variety of metrics and associated statistical tests.

Speed—in particular, speed distribution—was selected as a criterion because it effectively characterized the nature of road traffic and could also be used to measure safety. Following each simulation run, the approach speeds of all vehicles were recorded at the cross-section 900 ft from the stop bar. This location was selected for the provision of speed data reflecting driver behavior under the influence of activated flashers. A speed distribution was created and parameters were calculated: mean, median, mode, standard deviation, and kurtosis. Mean absolute percentage error (MAPE) was used as a measure of accuracy to determine the difference between the empirical and simulated speed distributions, as demonstrated in Equation (1).

$$MAPE = \frac{1}{5} \cdot \sum_{i=1}^5 \frac{R_i - S_i}{R_i} \tag{1}$$

where: R_i = empirical speed mean, median, mode, standard deviation, and kurtosis;
 S_i = simulated speed mean, median, mode, standard deviation, and kurtosis.

A MAPE value of less than 5% was targeted to indicate sufficient model fit and merit further analysis.

In addition, the Shapiro-Wilk (S-W) and Kolmogorov-Smirnov (K-S) tests were performed to check the normality of the approach speed. The K-S and Wilcoxon tests were also used to test the hypothesis that the empirical and simulated speed distributions were identical. Descriptions of all statistical tests utilized in the current study can be found elsewhere [14].

After 20 simulation runs, based on the default values of the select VISSIM parameters, the lowest noted MAPE value was 5.4% (see Table 2).

Table 2. Average MAPE in the initial model evaluation

Distribution parameter	Unit	Simulated speed	Empirical speed	MAPE [%]
Mean	[mph]	56.25	54.92	2.4
Median	[mph]	55	55	0.0
Mode	[mph]	58	58	0.0
Standard Deviation	-	7.76	6.30	23.2
Kurtosis	-	0.295	0.290	1.6
avg MAPE =				5.4

The difference between the four distribution parameters was less than 5%, but the difference in standard deviation was higher than 20%. That is, in comparison to the empirical data, simulation speed was distributed more evenly. The MAPE results indicated that a simulation model using the default parameter set could be satisfactory for this type of analysis; however, as discussed previously, simply using the MAPE may not provide adequate results if the user is interested in the distribution of a particular metric, such as speed.

While the functions were similar, it can be seen in Fig. 6 that the greatest differences occurred within the 46-53 mph range, and at speeds above 58 mph. Only for low speeds (< 46 mph), and for speeds in the range of 53-58 mph, were the distributions close. Overall, the simulated speed distribution had higher tails than did the empirical speed distribution. Tests for normality were

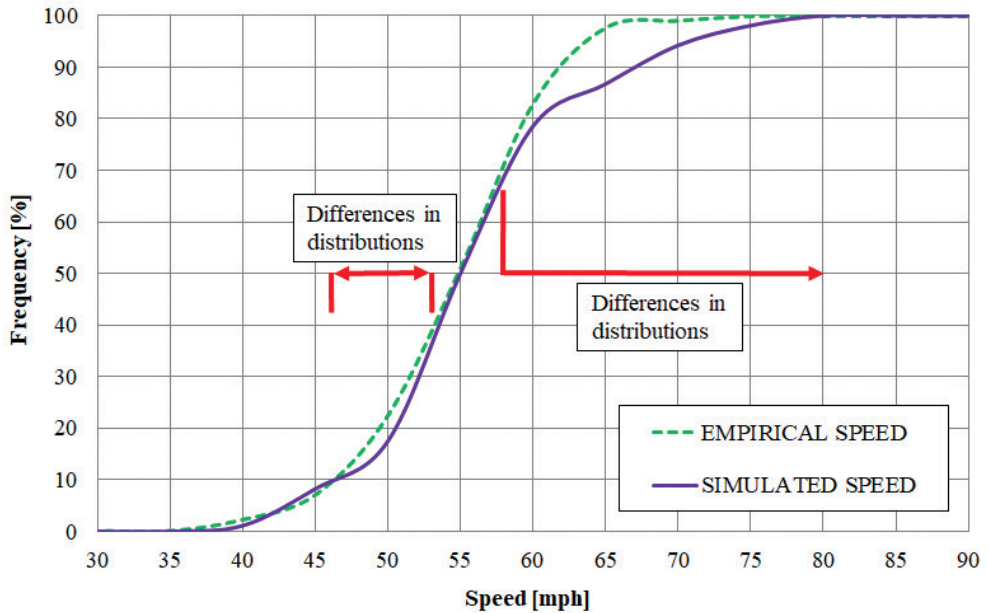


Fig. 6. Graphical comparison of speed distributions

performed to determine whether both speed distributions were normally distributed. The S-W test indicated that, at the 5% significance level, empirical speed ($p = 0.0499$) was normally distributed, but simulated speed ($p < 0.0001$) was not normally distributed. Additionally, the K-S test confirmed the non-normality of both speed distributions with $p < 0.01$. The K-S and Wilcoxon tests were also used to check the equality of the distributions, providing separate results. The K-S test rejected the hypothesis that both speed distributions resulted from the same continuous distribution ($p = 0.0352$ and < 0.05). At the same time, the Wilcoxon test accepted the hypothesis that the analyzed speed distributions were equal ($p = 0.1919$).

In summary, the microsimulation model resulted in a MAPE slightly higher than 5%, and failed the normality test. Additionally, the K-S and Wilcoxon tests provided different results as to the question of whether both speed distributions follow the same continuous distribution. Therefore, it was determined that the model may not have been acceptable and required calibration to facilitate further analysis. A full analysis can be found elsewhere [14].

3.4. Parameter set calibration

The calibration procedure was designed to identify the “best” parameter set for a given problem. In the current study, 19 driving behavior-related parameters were selected for testing. The desired parameter set depended on the MAPE and the results of the statistical tests. Following the calibration, the 10 parameter sets with the lowest MAPE value, having passed all statistical tests, were output.

As discussed previously, many researchers have identified the GA method as an appropriate tool for calibrating traffic microsimulation models. Based on these experiences, a GA was selected for the current research. The theory of the genetic algorithm is based on the Darwinian biological evolutionary processes occurring in nature; as such, it is ideal for solving complex problems. Put simply, a GA sets select parameters as “chromosomes” [8]. The GA for each task must include the following elements: the representation of potential solutions, a method to create an initial population, an evaluation function, a selection procedure, basic operators, and the values of parameters, such as population size, etc. The biggest advantage of using a GA is that, as a product of its approach, it considerably reduces the number of search steps and the amount of time required to determine a solution to a given problem. Fig. 7 illustrates the calibration procedure using GA.

The objective of the calibration process was to minimize MAPE values. First, agents that represented all select parameters needed to be defined. GA uses agent and gene terms, where a gene is represented by the binary digits 0 or 1. One agent is defined as a group of genes

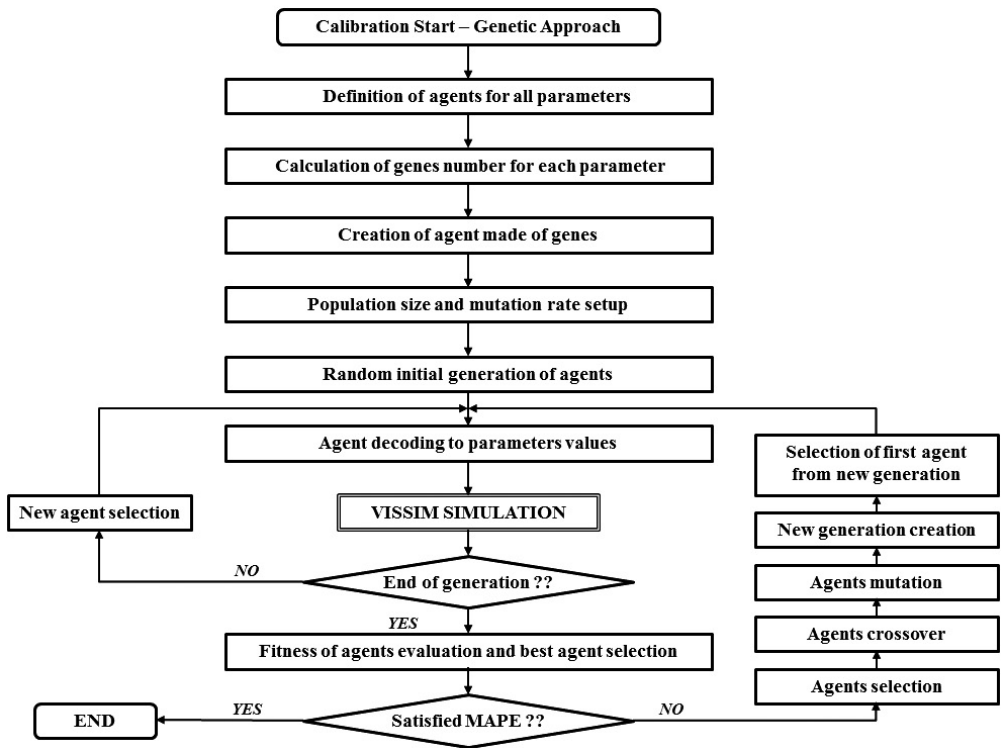


Fig. 7. Procedure based on GA for VISSIM model calibration [14]

used to represent a value of each parameter. One generation is then defined as the specified number of agents. The size of the population is defined as the number of agents included in one generation [15]. Second, for each calibrated parameter, the number of genes needed to

be estimated, which depended on the characteristics of the parameter and the increment of its value. The number of genes (n_i) for each parameter was calculated using the following equation:

$$n_i = \log_2 \left(\frac{\max(x_i) - \min(x_i)}{a_i} + 1 \right) \quad (2)$$

Where: x_i = value of i parameter;
 $\max(x_i)$ = maximum value of x_i ;
 $\min(x_i)$ = minimum value of x_i ;
 a_i = increment value of x_i .

Prior to calculation, the acceptable range of values (i.e., $\max(x_i)$ and $\min(x_i)$ values) for every parameter needed to be identified. The results of the calculations for all 19 VISSIM parameters are presented in Table 3.

Table 3. Number of genes with increments for calibrated parameters

PARAMETER		Min (x_i)	Max (x_i)	Number of genes (n_i)	Increment (a_i)
Number of observed preceding vehicles	NUMB_PRECED	1	4	2	1
Maximum look ahead distance	OBS_DISTANCE MAX	700	900	5	7
CC0 (Standstill Distance)	CC0	3	6	3	0.5
CC1 (Headway Time)	CC1	0.5	3	5	0.1
CC2 ('Following' Variation)	CC2	1	30	5	1
Waiting time before diffusion	T_DISAPPEAR	30	90	6	1
Minimum headway (front/rear)	MIN_LC_GAP	0.5	3	5	0.1
Frequency at 50mph	DESIRED_SPEED 1	0.10	0.30	5	0.007
Frequency at 60mph	DESIRED_SPEED 1	0.60	0.80	5	0.007
Frequency at 70mph	DESIRED_SPEED 1	0.85	0.95	4	0.007
Frequency at 40mph	DESIRED_SPEED 2a	0.05	0.15	4	0.007
Frequency at 50mph	DESIRED_SPEED 2a	0.30	0.50	5	0.007
Frequency at 60mph	DESIRED_SPEED 2a	0.80	0.95	4	0.01
Frequency at 30mph	DESIRED_SPEED 2b	0.03	0.15	4	0.008
Frequency at 40mph	DESIRED_SPEED 2b	0.25	0.45	5	0.007
Frequency at 50mph	DESIRED_SPEED 2b	0.70	0.90	5	0.007
Reaction to amber signal: α	AMBER_ALPHA	1	15	6	0.23
Reaction to amber signal: β_1	AMBER_BETA1	-0.4	-0.2	3	0.03
Reaction to amber signal: β_2	AMBER_BETA2	0.1	0.3	3	0.03
				84	

In the next step of the GA approach, the initial generation, based on agents, was set. These agents were composed of genes. Table 3 illustrates that there were 84 genes representing the 19 calibrated parameters in the current study. A population size of 32 was defined as the number of agents in one generation. In the initial generation, each gene of the agent was randomly assigned a 0 or 1. To decode each agent to the value of the parameter value x_i , the following approach was adopted:

$$x_i = \alpha_i \cdot A \cdot B + \beta \quad (3)$$

Where: $A = (a_1, a_2, a_3, \dots, a_n)$

$$B = \begin{pmatrix} 2^{n_i-1} \\ \cdot \\ \cdot \\ \cdot \\ 2 \\ 1 \end{pmatrix}$$

$i = 1, 2, 3, \dots, 19$

x_i = value of i parameter

a_i = increment value of x_i (0 or 1)

$b_i = \min(x_i)$ as shown in Table 3

A = vector representing agent

B = coefficient vector

n_i = number of genes of agent representing x_i

MAPE, as a measure of accuracy or fitness, was utilized to evaluate agent quality. The given parameter set was deemed to be more ideal if the fitness value was higher (e.g. lower MAPE). The procedure was automated, as shown in Fig. 8. Following each generation, the assessment of each agent was executed based on MAPE, and the best agent was selected.

When one generation was complete and agents were evaluated, steps for agent selection, crossover, and mutate were performed, resulting in the subsequent generation of agents. The selection was based on probability, and agents with lower MAPE values were the most likely to be selected. To crossover, two agents interchanged part of their genes to create two new agents. One agent was mutated to create a new agent by changing one of its genes from 1 to 0, or from 0 to 1 (15). After these steps were completed for the agents of the previous generation, additional agents were presented to create a new generation.

The GA described in the current paper was implemented using the MatLab software. The toolbox developed by the University of Sheffield provided all the necessary functions to implement the GA operators, i.e., selection, crossover, and mutation (2). The complete calibration procedure used in this study combined the MatLab software, Visual Basic, GA toolbox, and VISSIM.

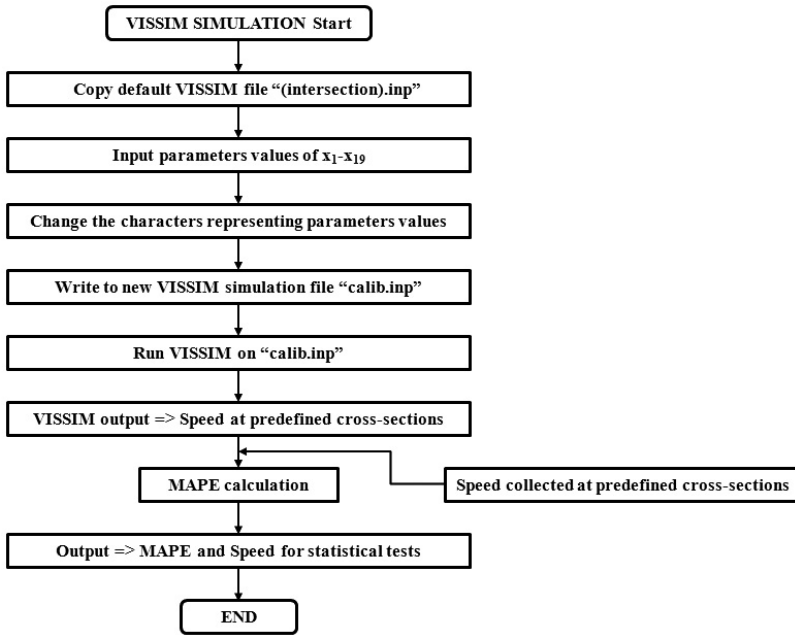


Fig. 8. Procedure for VISSIM model calibration [14]

3.5. Calibration results

After nine generations of the GA, the procedure was complete, and 10 sets of parameters with the lowest MAPE values were recorded, as shown in Table 4.

Table 4. Calibration results for the intersection of US-77 & Pioneers Blvd

		DEFAULT VALUE	1	2	3	4	5	6	7	8	9	10	
PARAMETER	NUMB_PRECED	2	4	4	3	3	3	3	3	4	3	4	
	OBS_DISTANCE_MAX	820.21	889	889	791	791	791	791	791	847	700	889	
	CC0	4.92	3.5	3.5	6.5	6.5	6.5	6.5	6.5	3	5.5	3.5	
	CC1	0.9	1.9	1.3	0.8	1.6	2.9	0.8	0.8	3.6	3.1	1.9	
	CC2	13.12	1	18	12	12	18	12	12	20	8	1	
	T_DISAPPEAR	60.0	70.0	54.0	34.0	34.0	54.0	34.0	34.0	49.0	62.0	70.0	
	MIN_LC_GAP	1.64	3	3.5	3.1	3.1	3.5	3.1	3.1	2.1	3.5	3	
	DESIRED_SPEED 1	0.150	0.210	0.270	0.230	0.230	0.270	0.230	0.230	0.230	0.230	0.210	0.210
		0.750	0.750	0.810	0.810	0.810	0.810	0.810	0.810	0.810	0.750	0.740	0.750
		0.900	0.910	0.950	0.950	0.950	0.950	0.950	0.950	0.950	0.930	0.890	0.910
DESIRED_SPEED 2a	0.070	0.120	0.110	0.110	0.110	0.110	0.110	0.110	0.110	0.060	0.100	0.120	
	0.400	0.300	0.520	0.520	0.520	0.520	0.520	0.520	0.520	0.300	0.330	0.300	
	0.900	0.820	0.900	0.940	0.940	0.900	0.940	0.940	0.940	0.840	0.880	0.820	

PARAMETER	DESIRED_SPEED 2b	0.040	0.130	0.130	0.140	0.140	0.130	0.140	0.140	0.130	0.130	0.130
		0.300	0.340	0.340	0.290	0.290	0.340	0.290	0.290	0.380	0.380	0.340
		0.800	0.810	0.870	0.830	0.890	0.880	0.830	0.830	0.720	0.720	0.810
PARAMETER	AMBER_ALPHA	1.59	12.5	3.53	7.9	7.9	3.53	7.9	7.9	1.23	1	12.04
PARAMETER	AMBER_BETA1	-0.26	-0.25	-0.22	-0.34	-0.34	-0.22	-0.34	-0.34	-0.37	-0.37	-0.37
PARAMETER	AMBER_BETA2	0.27	0.22	0.25	0.22	0.22	0.25	0.22	0.22	0.31	0.13	0.13
TESTS	MAPE [%]	5.45	4.96	5.34	5.8	5.88	6	6.13	6.23	6.23	6.35	6.39
	Kolmogorov-Smirnov (Passed/Not Passed)	N	P	N	P	N	N	P	P	P	P	N
	Wilcoxon (Passed/Not Passed)	P	P	N	P	N	N	P	P	P	P	P

Additionally, each parameter set was checked to verify that the simulated speed distribution and empirical speed distribution were equally distributed, and to determine whether they displayed the same continuous distribution. The latter check was performed using the K-S and Wilcoxon tests. The set with the lowest value of MAPE (4.96%) passed the non-parametric tests (K-S and Wilcoxon), which meant that the empirical approach speed and simulated approach speed distributions originated from the same continuous distribution. Consequently, this parameter set was utilized for further evaluation. While the MAPE values for the calibrated and default parameter sets were similar, the distributions from the latter

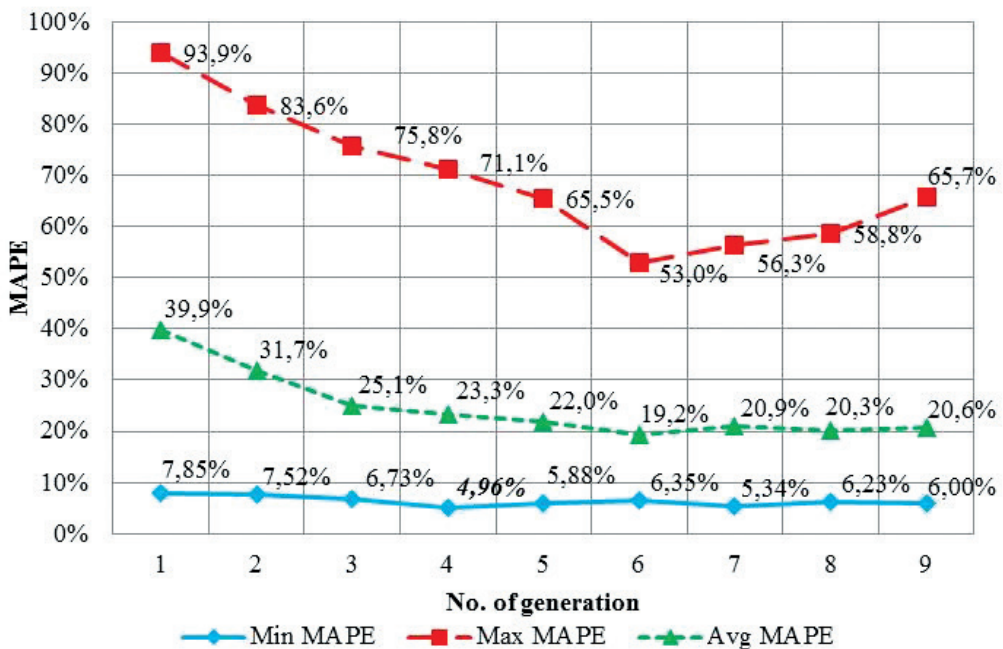


Fig. 9. Calibration results for the analyzed intersection

scenario did not differ statistically from the empirical distribution. Therefore, the calibrated parameter set was utilized in subsequent analyses.

Fig. 9 displays the results for each generation in the GA.

As can be seen, following nine generations, MAPE values were nearly steady; consequently, the algorithm was terminated.

3.6. Evaluation of the calibrated parameter set

Table 5 displays a comparison between the simulated results from the default parameter set and the calibrated parameter set.

It can be seen that the calibrated results had lower MAPE values, as calculated using the five measures of speed distribution. In addition, the calibrated model passed statistical testing, indicating that the simulated and empirical speed distributions were statistically equivalent. The calibrated VISSIM parameter set is displayed in Table 6.

Table 5. Evaluation of calibrated parameter sets

	Distribution Parameter	MICROSIMULATION MODEL	
		Default	Calibrated
MAPE [%]	Mean	2.4	1.8
	Median	0.0	0.0
	Mode	0.0	0.0
	Standard Deviation	23.2	22.3
	Kurtosis	1.6	0.8
TEST	Average MAPE [%]	5.40	4.96
	K-S test	Not Passed	Passed
	Wilcoxon test	Passed	Passed
	MAPE Improvement	8%	

Table 6. Calibrated parameter set for microsimulation model

BEHAVIOR	PARAMETER		DEFAULT VALUE	CALIBRATED MODEL
CAR FOLLOWING	Number of observed preceding vehicles	NUMB_PRECED	2	4
	Maximum look ahead distance	OBS_DISTANCE MAX	820.21	889
	CC0 (Standstill Distance)	CC0	4.92	3.5
	CC1 (Headway Time)	CC1	0.9	1.9
	CC2 ('Following' Variation)	CC2	13.12	1

LANE CHANGE	Waiting time before diffusion	T_DISAPPEAR	60.0	70.0
	Minimum headway (front/rear)	MIN_LC_GAP	1.64	3.0
DESIRED SPEED 1	% at 50mph	DESIRED_SPEED 1	0.150	0.210
	% at 60mph	DESIRED_SPEED 1	0.750	0.750
	% at 70mph	DESIRED_SPEED 1	0.900	0.910
DESIRED SPEED 2a	% at 40mph	DESIRED_SPEED 2a	0.070	0.120
	% at 50mph	DESIRED_SPEED 2a	0.400	0.300
	% at 60mph	DESIRED_SPEED 2a	0.900	0.820
DESIRED SPEED 2b	% at 30mph	DESIRED_SPEED 2b	0.040	0.130
	% at 40mph	DESIRED_SPEED 2b	0.300	0.340
	% at 50mph	DESIRED_SPEED 2b	0.800	0.810
SIGNAL CONTROL	Reaction to amber signal: α	AMBER_ALPHA	1.59	12.5
	Reaction to amber signal: β_1	AMBER_BETA1	-0.26	-0.25
	Reaction to amber signal: β_2	AMBER_BETA2	0.27	0.22

3.7. Model validation

As mentioned previously, validation was performed to determine whether the calibrated microsimulation model performed properly by comparing the output with data not utilized in the original calibration. The calibrated model was run 10 times to determine whether the model was capable of reflecting actual traffic conditions at the test intersection. Waiting time on the minor approaches was used as a validation parameter. Real values of waiting times were gathered during data collection. The mean value of waiting time for the EB and WB approaches, the output from 10 simulation runs, was compared to the values derived from the collected empirical data. The results are displayed in Table 7.

Table 7. Summary of calibrated model validation

Test intersection	Minor approach	Average waiting time [s]		t-test (p)	Result (Passed /Not passed)	Validation
		Empirical data	Simulated data			
US-77 & Pioneers Blvd	EB	21.6	22.2	0.397	Passed	YES
	WB	17.1	16.3	0.395	Passed	

Two statistical tests, the F-test, and the t-test, were performed to determine whether statistically significant differences existed between the analyzed values at the 5% level of confidence. It can be observed that the statistical tests were passed ($p = 0,397 > 0,05$ and $p = 0,395 > 0,05$), indicating that the microsimulation model could effectively mimic real traffic at the signalized intersection [14] by providing the values of the wait time on minor approaches that were proximate to the empirical values determined through data collection (see Table 7). Therefore, the calibrated model was acceptable, and was deemed eligible for further safety analysis.

4. Conclusions

This study resulted in a calibrated stochastic model of a signalized intersection located in Nebraska. The model was calibrated using a genetic algorithm with non-parametrical statistical tests. The GA approach provided a quick and effective method for finding the “best” set of VISSIM parameters, and seemed to be a very effective tool for the calibration of traffic and the development of a stochastic simulation model for the studied signalized intersection. An innovative approach described in this paper utilized speed distribution as the objective function in the calibration process. While other researchers have calibrated microsimulation models to measures of central tendency (e.g. mean), this paper proposed calibrating to approach speed distribution. Mean absolute percentage error, calculated for the speed distribution parameters, was used as a measure of fitness. Non-parametric statistical tests were then utilized to indicate whether the subsequent value of MAPE was acceptable at a 5% level of significance (see Table 2). The distributions of the observed and calibrated speeds were compared, and it was determined that no statistically significant differences existed at the 5% confidence level (see Table 5). The calibration process resulted in a microsimulation model (see Table 6), which was also validated to the wait time on the minor approaches (see Table 7).

The VISSIM model that was developed, calibrated, and validated in this paper could be used as the basis for a methodology to implement specific safety countermeasures at signalized intersections.

References

- [1] Burnett N. P., *Effect of information on driver's risk at the onset of yellow*, Master Thesis at University of Nebraska-Lincoln, July, 2011.
- [2] Cao Y., Wu Q., *Teaching Genetic Algorithm using Matlab*, International Journal of Electrical Engineering Education, Vol. 36/1999, 139–153.
- [3] Cunto F., Saccomanno F., *Calibration and validation of simulated vehicle safety performance at signalized intersections*, Accident, Analysis and Prevention, 40/2008, 1171–1179.
- [4] Duong D., Saccomanno F., Hellinga B., *Calibration of microscopic traffic model for simulating safety performance*, Compendium of papers of the 89th Annual TRB Conference held Jan. 10–14, 2010 in Washington, D.C., No. 10-0858/2010.

- [5] <http://maps.google.com> (access: 01.05.2010).
- [6] www.wavetronix.com/en/products (access: 01.05.2010).
- [7] Kim S-J., Kim W., Rilett L.R., *Calibration of microsimulation models using nonparametric statistical techniques*, Transportation Research Record, No. 1935/2005, 111–119.
- [8] Kim K., Rilett L.R., *Genetic-algorithm-based approach for calibrating microscopic simulation models*, 2001 IEEE Intelligent Transportation Systems Conference Proceedings, 2001, 298–704.
- [9] Mathew T.V., Radhakrishnan P., *Calibration of microsimulation models for nonlane-based heterogeneous traffic at signalized intersections*, ASCE Journal of Urban Planning and Development, Vol. 136/2010, Issue 1, 59–66.
- [10] Park B., Schneeberger J. D., *Microscopic simulation model calibration and validation: Case study of VISSIM simulation model for a coordinated actuated signal system*, Transportation Research Record, No. 1856/2003, 185–192.
- [11] Park B., Qi H., *Development and evaluation of a procedure for the calibration of simulation models*, Transportation Research Record, No. 1934/2005, 208–217.
- [12] Santhanam S., Park B., *Development of VISSIM Base Model for Northern Virginia (NOVA) Freeway System*, Report No. UVACTS-13-0-124, June 2008.
- [13] VISSIM 5.30-05 User Manual, PTV Planung Transport Verkehr AG, Innovative Transportation Concepts, 2011.
- [14] Wojtal R., *Development of a methodology for analyzing safety treatments at isolated signalized intersections*, Ph.D. Dissertation, University of Nebraska-Lincoln, Lincoln 2012.
- [15] Yu L., Lei Y., Xumei C., Tao W., Jifu G., *Calibration of VISSIM for bus rapid transit systems in Beijing using GPS data*, Journal of Public Transportation, No. 3/2006, 239–257.
- [16] Zalzala A., Fleming P., *Genetic algorithms in engineering systems*, The Institution of Engineering and Technology, February, 1999.

Jerzy Orlof (jerzy.orlof@pk.edu.pl)

Faculty of Physics, Mathematics and Computer Science, Cracow University
of Technology

POINT CLOUD BASED VIEWSHED GENERATION IN AUTOCAD CIVIL 3D

WYZNACZANIE WYKRESU WIDOCZNOŚCI NA PODSTAWIE CHMURY PUNKTÓW W PROGRAMIE AUTOCAD CIVIL 3D

Abstract

We propose a new method for a point cloud based viewshed generation. We start with a brief summary of existing methods for viewshed generation and then present ours based on point clouds. The method is implemented in AutoCAD Civil 3D. Moreover, we present an idea of how to deal with huge point clouds using AutoCAD Civil 3D and how to automatize this process.

Keywords: viewshed; point cloud, digital modelling; ray tracing; LiDAR

Streszczenie

W artykule została przedstawiona metoda tworzenia wykresu widoczności na podstawie chmury punktów. Zaprezentowano zarówno znane dotychczas metody wyznaczania wykresów widoczności jak i metodę bazującą na chmurze punktów zaimplementowaną w oprogramowaniu AutoCAD Civil 3D. Omówiono również nowe rozwiązanie problemu przetwarzania dużej chmury punktów za pomocą programu AutoCAD Civil 3D oraz propozycję automatyzacji tego procesu.

Słowa kluczowe: wykres widoczności, chmura punktów, model cyfrowy, śledzenie promienia, LiDAR

1. Introduction

Together with the development of technologies, it is important to consider evolving data types, such as point clouds [1]. The particular structure of point clouds provides new opportunities for 3D modelling and visualization. Nowadays, any professional project, for example of a building, should be tested to fit urban design and landscape planning. Thanks to point clouds such tests are easier to conduct. Both during design and construction, point clouds can be used for visibility purposes, from public safety to environmental protection, traffic or research. To check the visibility using a point cloud, a viewshed must be generated. A viewshed [2] is a geographical map generated for a fixed position. This map contains and differs between two types of zones: (1) in line-of-sight (visible) (2) and not visible from a selected position. Moreover, the grade of visibility can be shown.

Generation of a viewshed from a point cloud is not an easy task: point clouds can be huge and difficult to handle, and the viewshed calculation can be time-consuming and generate errors. In this article, we present a solution to this problem introducing a new method for point cloud based viewshed generation.

Viewshed applications depend on the area that they cover. A viewshed can contain areas such as a building with its terrain, a village or even a whole city. In particular, an example of a viewshed application for a car driver shows the part of the road that he/she can see from his/her point of view. Other examples include: change of the visibility due to the insertion of a new building, best placement of security cameras (to limit blind spots), both indoors as well as outdoors for city monitoring, where at least one camera should be able to follow a moving object. In marketing, viewsheds are used to decide the position of a banner advertisement so that it is visible from as many roads as possible. A combination of viewsheds (shadow analysis) with solar radiation and 3-D city models allows the sustainability of the rooftops for solar panels hosting to be evaluated. Finally, viewshed analysis is used in the visual impact assessment and landscape planning to find the most sustainable scenarios for the future

2. General concept

Historically, the visibility was checked through empirical methods, such as an architectural model with a light source.

Nowadays, the visualization of tridimensional data is done using numerical methods, based on GIS (Geographical Information System). These systems use algorithms to verify the visibility from a specific position and create a viewshed. In particular, GIS can visualize DEM (Digital Elevation Model) as heightmaps or TIN (Triangular Irregular Network) surfaces. A heightmap is a raster image made of a matrix of values (i.e. RGB) in a grid while a TIN is a vectorial surface created from triangles. If the resolution of the model (raster or vectorial) is insufficient in relation to the dimension of the target object, the resulting model can be irregular and has to be complemented by additional measurements points to increase the sampling resolution.

Following the development of digital techniques, visibility check algorithms based on rendering (computer graphics) were implemented in GIS programs. One of such algorithms is Ray tracing [3] created by A. Appel in 1968. This algorithm is used to simulate the propagation of the light waves and designate the shadows of objects. It generates a viewshed as an image with brightened/shadowed objects visible/invisible from the light source point.

The Ray tracing method works as follows. First, a ray is sent from a light source. The ray is reflected by any 3D object encountered along its path and travels back to the camera mounted above the light source. The ray can't penetrate objects, so only those visible from the light source are illuminated.

Recently, research has been directed towards the development of a new data structure: the point cloud. Point clouds contain information about the terrain that includes both natural elements and those created by people. A point cloud is a set of points with a structure assigned to each of them. This structure always stores information about the coordinates (x,y,z) , and optionally also colour (RGB), intensity, return time or echo number related to the landform. It is not possible to create a viewshed directly from a point cloud because the points are not connected in any way and so they do not describe a surface. To generate a viewshed, first one has to create a surface from the point cloud, usually a TIN [4]. The next step is to set up a light source. Finally, the viewshed can be generated from the Ray tracing method.

2.1. Point cloud generation methods

The performance of the methods for the generation of a point cloud depends on the size of the area that the point cloud should cover.

Point clouds can be generated using different types of scanners. Widely used are 3D scanners, which can be divided into two groups: contact and non-contact scanners.

Contact scanners require the physical touch of the rigid or articulated arm of the scanning machine with the object. That is why this method may result in damage to the target body and so can be too invasive for some measurements. The process itself is done through the gliding of the rigid arm scanner along the measurement line or through the rotation of the articulated arm. In both cases, carefully mounted sensors on the arms collect data points according to the shape of the scanned bodies.

Non-contact scanners are based on a completely different approach. These scanners emit radiation (light, ultrasound or x-ray) and collect it back after it's reflected by the target object. The time between the emission and collection of the radiation is directly related to the distance to the target. As the name suggests, a rangefinder is an example of such a scanner.

An example of non-contact scanning technique is LIDAR [5] (LIght Detection And Ranging, Fig.1). This scanner, installed on an airborne vector, is used to survey large areas. It is based on laser scanning from an aircraft, helicopter, drone or any other machine that is able to fly. The scanning gives the distance between the machine and a point on the ground. To get optimal results using LiDAR a combination of the scanner, GPS and an Internal Navigation System (INS) must be used to obtain a very precise point cloud.



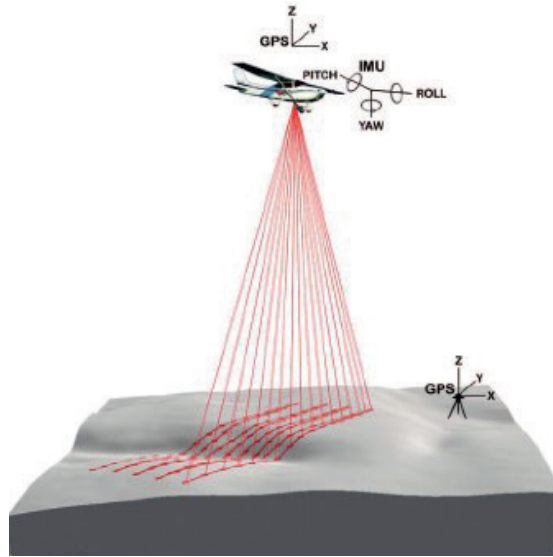


Fig. 1. Graphical visualization of the LiDAR airborne scanning method [6]

Another method for point cloud generation is photogrammetry (i.e. from the ReCap program) and requires very detailed georeferenced photographs of the terrain as an input. The photographs can be obtained by a photographer on the ground or, as in the laser scanning, by an airborne sensor platform. For this method to function correctly a specified overlapping among the photographs is needed.

Photogrammetry doesn't have to be used for terrain surveys. For example, in medicine, detailed and processed photographs of a body part can be used for the 3D printing of prosthesis.

Yet another method to be mentioned is based on structured light. Here a pattern of white lines of define sizes is projected on the scanned object (Fig. 2). From the computer analysis of the distortion of these lines a point cloud – and so the shape of the object – can be recovered.

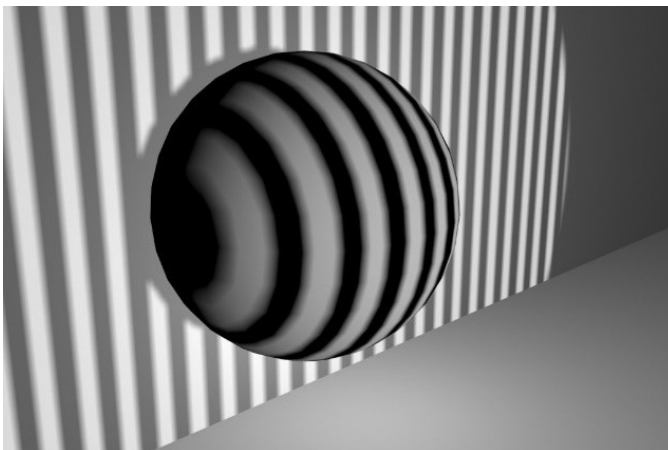


Fig. 2. Pattern of white lines projected on the scanned object

To get the whole 3D image of the object, it is enough to put it on the rotating plate and scan it from all sides.

The last method that we mention here is terrestrial 3D scanning [7].

This method works well for small areas like a room and it is not considered in the next subsection.

2.2. Viewshed generation from a big data set

As mentioned before, a viewshed cannot be generated directly from a point cloud. The intermediate step is to create a surface from a point cloud. This may be a difficult task if the number of the measured points is huge and the required computational power is not feasible. In the next part of the article, we give a solution to this problem and propose a method based on the generation of partial viewsheds and the way of connecting them.

In this article we present an example of a huge point cloud from which the generation of a viewshed was not possible. This example was a trigger to come up with a feasible solution for the surface and viewshed generation from huge point clouds. A viewshed from this huge point cloud was needed to another research project.

3. Input data

As input data we consider a point cloud of Krakow from 2012, within the project ISOK, Standard II – 12 points / m².

We present a set of assumptions, which were crucial for the choice of divisions of the example point cloud:

- ▶ The point cloud is in Standard II – 12 points / m²
- ▶ The structure containing information about a single point occupies 48 bytes
- ▶ The whole point cloud file occupies 12 GB, giving around 270833333 points
- ▶ The available computers cannot cope with such a big point cloud

The choice of the point cloud was made according to the data size and the reality of the location, otherwise was random.

3.1. Division of the point cloud

Once a central point (a light source) in the point cloud has been decided, it is natural to start the division at that point. That is why we propose a division into radial slices with a centre in the light source. All of the slices have the same, fixed angle and only 'x' and 'y' coordinates are relevant for the assignment a point into a pie slice. Note that the central point is not necessarily in the centre of the point cloud. The position of the central point in this example was not important for us, it was requested in another research project. Differently, the choice of the division method was crucial due to the necessity of the light ray continuation. That is why the methods such as



division into square grid would result in errors and were rejected. This is because viewsheds for every single square would be generated separately and the big objects, such as buildings in the squares closest to the light source, would not shed the smaller objects in the farther squares. Now, instead of our division into pie slices according to the cylindrical coordinates, one could consider pieces obtained from the division of a sphere according to the spherical coordinates. This division would be similarly good, however more difficult to handle due to the varied number of points in the pieces and more complicated overlapping zones. Finally, one could shrink the number of points by consideration of mean values. As we were interested in a very detailed viewshed, the idea of mean values was rejected due to the approximation errors.

In our pie division method, the number of slices can be chosen arbitrary and depends on the size of the point cloud and the computational power available. We conducted many iterations in the search for the minimal number of slices.

Let's consider an example of division of the point cloud into pieces

- ▶ Division of the point cloud into 10 pieces (Fig. 3)
- ▶ Whole pie = 360°
- ▶ A piece of a pie = $360^\circ / 10 = 36^\circ$

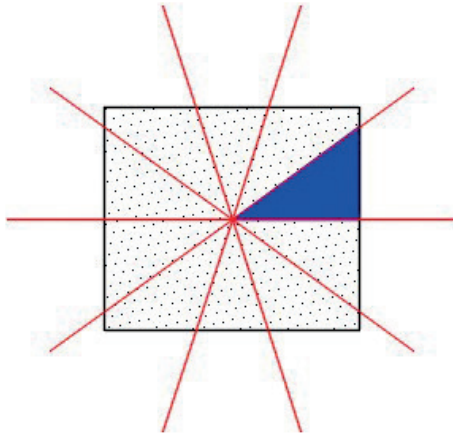


Fig. 3. Schematic representation of the division method. The square contains all the points of the considered point cloud. The division into pieces of pie (an example one is indicated in blue) is done from the central point

The resulting pieces contain a different number of points because the points in the point cloud are not uniformly distributed, the central point can be chosen arbitrary and the slices may have different areas. This has to be taken into account and accordingly the division would be called a success provided the generation of a surface from the biggest slice point cloud was possible.

In this sense neither the division into 10 pieces, nor in 30 led to success. Increasing the number of pieces, we found the first working division for 180 pieces, with a single slice of 2° . The surface generation time was related to the number of points contained in a single piece, where for bigger pieces it took way more time than for the smaller ones.

3.1.1. Point cloud division principle

A point cloud stores information about coordinates in Cartesian form (x,y,z) . To divide a point cloud into slices of a pie (as described in the previous section), it is necessary to change the coordinates (x,y) , into the polar form. As mentioned before the 'z' coordinate is irrelevant for the division. The change is performed according to the following formulas

$$x=r \cdot \cos(\alpha)$$

$$y=r \cdot \sin(\alpha)$$

$$r=\sqrt{x^2+y^2}$$

$$\text{For } x > 0 \quad \alpha = \frac{\operatorname{arctg}\left(\frac{y}{x}\right) \cdot 360^\circ}{(2 \cdot \pi)}$$

$$\text{For } x < 0 \text{ and } y \geq 0 \quad \alpha = \left(\operatorname{arctg}\left(\frac{y}{x}\right) + \pi \right) \cdot \frac{360^\circ}{2 \cdot \pi}$$

$$\text{For } x < 0 \text{ and } y < 0 \quad \alpha = \left(\operatorname{arctg}\left(\frac{y}{x}\right) - \pi \right) \cdot \frac{360^\circ}{2 \cdot \pi}$$

$$\text{For } x = 0 \text{ and } y > 0 \quad \alpha = \frac{\pi}{2} \cdot \frac{360^\circ}{2 \cdot \pi}$$

$$\text{For } x = 0 \text{ and } y < 0 \quad \alpha = -\frac{\pi}{2} \cdot \frac{360^\circ}{2 \cdot \pi}$$

$$\text{For } x = 0 \text{ and } y = 0 \quad \alpha = 0$$

where the arctangent is calculated in radians.

The subdivision of the points into slices of a pie was done according to the angle.

3.1.2. Implementation of the viewshed generation method into AutoCAD Civil 3D

After we have found the right division, we can focus on the viewshed generation from a smaller point cloud, namely our piece of a pie. This can be done in AutoCAD Civil 3D.

A viewshed generation can be performed in 7 steps (AutoCAD Civil 3D nomenclature is used here):

- 1) Set up a light source in a fixed position
- 2) Create a new empty TIN surface (the surface style must be renderable)
- 3) Import points from the point cloud to build the surface
- 4) Set up the orthogonal view from the top
- 5) Perform rendering, it will show the difference in illumination for visible and invisible zones
- 6) Save the file



The manual call of those functions for every single pie slice would take a lot of time. The solution is automatization of this process so that the program with steps 1-6 is applied in batch to all the slices of the pie. We perform the automatization writing a template in AutoCAD Civil 3D with a light source, correct TIN surface style and the orthogonal view.

The first step has to be performed manually by the user, namely one has to start program AutoCAD Civil 3D and open the template. Next, the steps 1:6 are performed automatically for all the pie pieces.

4. Results

Thanks to the division into 180 pieces, the generation of a viewshed for every single piece was possible. The next step was to connect the pieces and obtain the whole point cloud viewshed, which was done in Matlab. In Fig. 4, 5 and 6, we show one slice of the pie, a few connected pieces and the whole viewshed, respectively. The image shows a bridge over the Vistula River. The light source was placed in the centre of the image, around 1.5 m above the bridge so that it mimics the position of a human eye. The viewsheds are in grey scale, where the illuminated zones are brightened and the invisible ones are darkened. As our goal is to differ between visible and invisible zones, we transferred Fig. 6 into a binary scale, presented in Fig.7 with visible elements in white and invisible in black.

We encountered an issue while connecting the partial viewsheds to obtain the whole image. Every single viewshed was based on a triangular surface (TIN). Because of that there was always a piece of surface missing between two adjacent slices (Fig. 8). The solution was to increase the angle that a single viewshed covered adding 0.1° on both sides of the 2° slice (Fig. 9), to provide a partial overlapping between adjacent slices. The additional angle (0.1°) was chosen to be sufficiently big so that possible glitches on the overlapping surfaces were negligible. Note that, for the sake of clarity in the presentation, we show smaller number of points and division slices in Fig.8-10, than in our actual project presented in the paper.

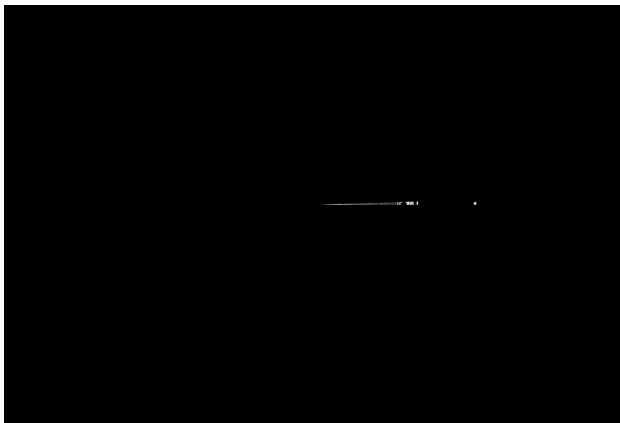


Fig. 4. The viewshed of a pie piece (2°) after rendering

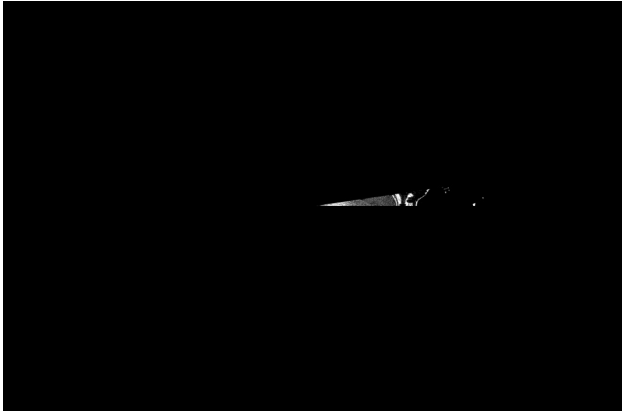


Fig. 5. he viewshed from a few connected rendered pieces



Fig. 6. The viewshed of the whole area, terrestrial view

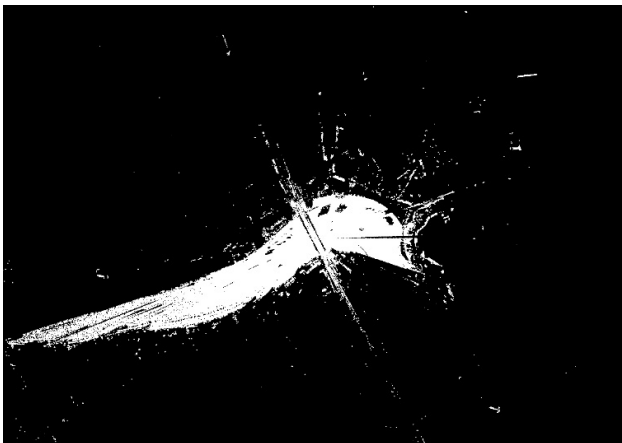


Fig. 7. The final binary viewshed





Fig. 8. The connection of two surfaces (in red and grey) with an empty space in the middle (in green)

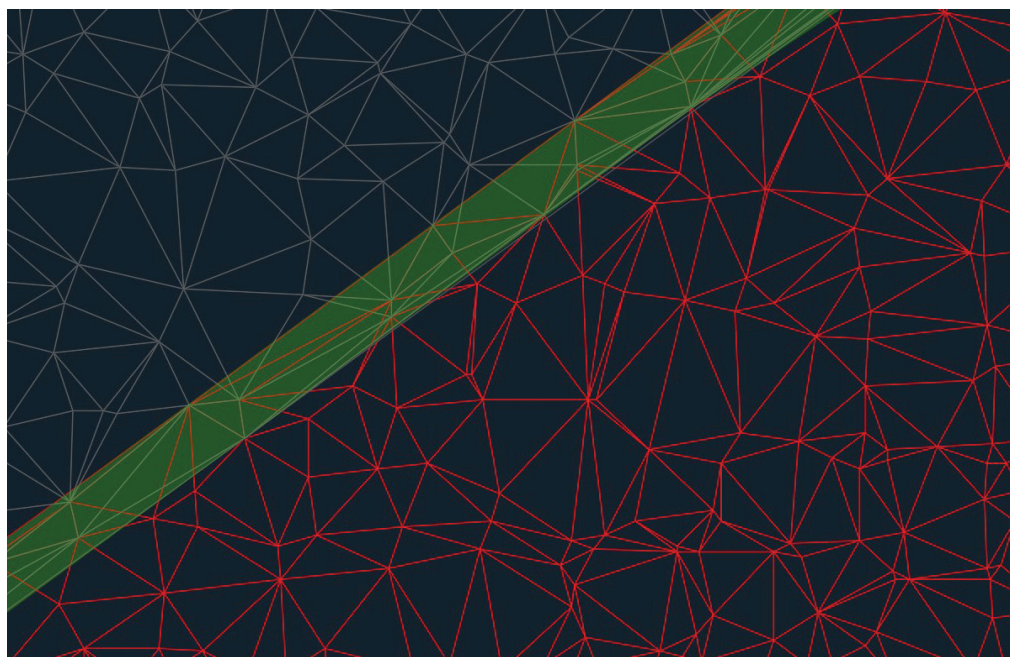


Fig. 9. Connection of two surfaces (in red and grey)
The surfaces have overlapping piece of 0.2° (in green)

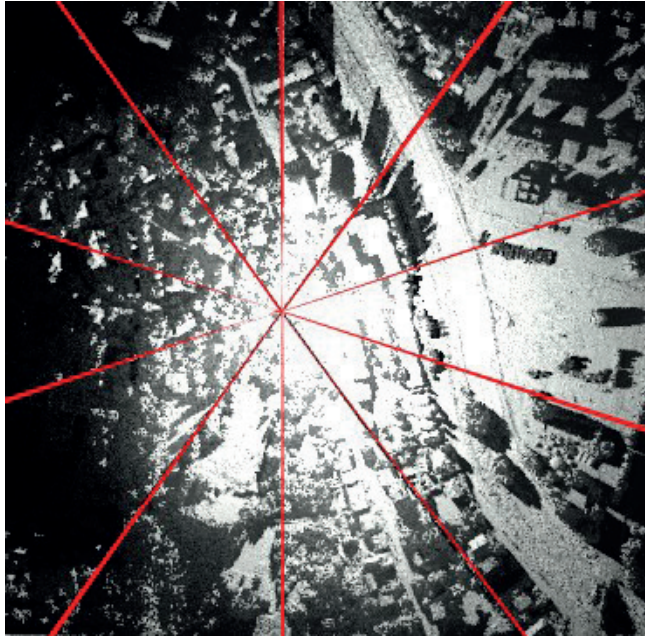


Fig. 10. The final viewshed with all the pie pieces connected. In red we indicate the overlapping part between the consecutive viewsheds

Another problem was a surface in the central point of the viewshed, where all the compounding surfaces should connect (Fig. 11). To deal with this problem, we generated an additional square surface covering the centre (Fig. 12–13).

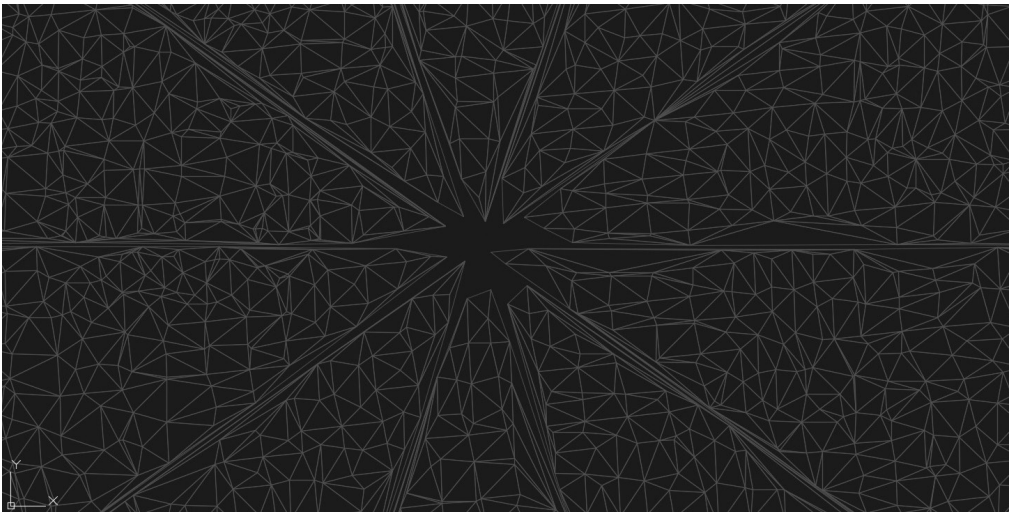


Fig. 11. Connection of all the compounding surfaces. Note the empty space in the centre and in between every two consecutive pieces of pie



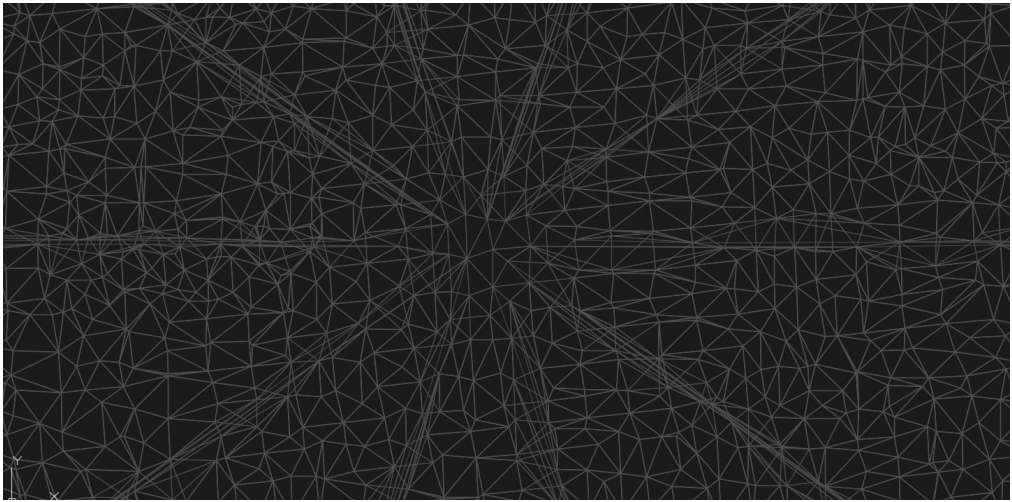


Fig. 12. Connection of all the compounding surfaces with the overlap). Note the overlapping part of the surface between the consecutive slices and additional surface in the middle

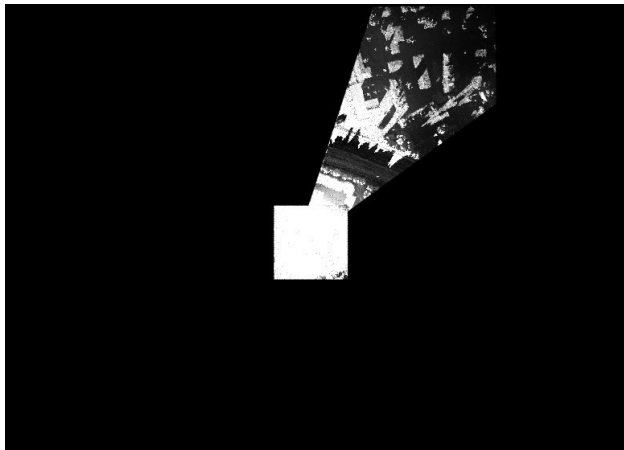


Fig. 13. Connection of between a single slice of a pie viewshed with the central square viewshed

5. Parallelization

In the future we plan to parallelize the programs for the division of the point cloud and for the generation of viewsheds in Autocad Civil 3D. It is possible to apply parallelization to divide a big point cloud into smaller ones because all the points are independent of each other. This independence of data is a necessary condition for parallelization. Thanks to the parallelization, a few points can be distributed into smaller point clouds at the same time. Similarly, the generation of viewsheds in Autocad Civil 3D can be performed with parallelization. The output viewshed from the first pie slice is not necessary to create the next ones. The viewsheds from different pieces of pie can be done simultaneously on different computer cores.

6. Summary

In this article, we presented a method for viewshed generation from a point cloud. For big point clouds, we proposed a method based on their subdivision according to pie slices starting from a central point. To apply this method, we considered an example of a point cloud, where we generated separate viewsheds for every single slice. Finally, while connecting the slices we encountered problems with empty space on the slice borders and near the central point, that we fixed using partial overlapping between slices. The generation of viewsheds was performed and automatized in AutoCAD Civil 3D. The next step is the parallelization of the division and viewsheds generation process.

References

- [1] Warchoń A., *Gęstość chmury punktów pochodzącej z mobilnego skanowania laserowego / Density of point clouds in mobile laser scanning*, Archives of Photogrammetry, Cartography & Remote Sensing, 27, 149–161, Jan. 2015.
- [2] Kim Y., Rana S., Wise S., *Exploring Multiple Viewshed Analysis Using Terrain Features and Optimisation Techniques*, Computers and Geosciences 2004, 30(9).
- [3] Kajia K.T., *The rendering equation*, ACAM SIGGRAPH Computer Graphic Homepage Vol. 20, Issue4, 1986.
- [4] Quan Y., Song J., Guo X., Miao Q., Yang., *Filtering LiDAR data based on adjacent triangle of triangulated irregular network*, Multimedia Tools & Applications, 8 May 2017.
- [5] Ptak A., *LiDAR w wielkim mieście*, Geodeta: Magazyn Geoinformacyjny, May 8, 2017.
- [6] Researchgate, https://www.researchgate.net/profile/Robert_Mcgaughey3/publication/233622295/figure/fig1/AS:299984089042947@1448533188764/Figure-1-Schematic-showing-LIDAR-data-collection-over-bare-ground.png (access: 03.07.2015).
- [7] Przywara J., *Skaner, czyli myśleć w 3D*, Geodeta: Magazyn Geoinformacyjny, May 8, 2017.



Krzysztof Tomczyk (ktomczyk@pk.edu.pl)

Marek Sieja

Department of Automatic Control and Information Technology, Faculty of Electrical and Computer Engineering, Cracow University of Technology

ASSESSMENT OF PROPAGATION OF MODELLING UN-CERTAINTY BY THE
PROCEDURES FOR DETERMINING MAXIMUM DYNAMIC ERRORS

OCENA PROPAGACJI NIEPEWNOŚCI MODELOWANIA PRZEZ PROCEDURY
WYZNACZANIA MAKSYMALNYCH BŁĘDÓW DYNAMICZNYCH

Abstract

The paper discusses the method for modelling linear analogue systems of the second order in the time domain. As a result of such modelling, the parameters of the model and the associated uncertainties are obtained. Procedures for determining the absolute error and the integral-square error are presented. These procedures make it possible to determine precisely such an input signal with one constraint that maximizes an error at the output of the system. Values of the parameters of an example model are determined and the propagation of the uncertainties of modelling results is assessed by the procedures for determining the maximum dynamic errors. The results of calculations presented in the paper were carried out in MathCad 15.

Keywords: uncertainties of model parameters, time domain modelling, maximum dynamic error

Streszczenie

W artykule omówiono metodę modelowania w dziedzinie czasu liniowych systemów analogowych drugiego rzędu. Jako wynik takiego modelowania uzyskano parametry modelu oraz związane z nimi niepewności. Przedstawiono procedury wyznaczania maksymalnych błędów dynamicznych dla przypadku kryteriów błędu: bezwzględnego i całkowokwadratowego. Procedury te pozwalają w sposób precyzyjny określić taki sygnał wejściowy z jednym ograniczeniem, który maksymalizuje błąd na wyjściu systemu. Wyznaczono wartości parametrów przykładowego modelu oraz oceniono propagację niepewności wyników modelowania przez procedury wyznaczania maksymalnych błędów dynamicznych. Wyniki obliczeń przedstawionych w artykule przeprowadzono w programie MathCad 15.

Słowa kluczowe: niepewności parametrów modelu, modelowanie w dziedzinie czasu, maksymalny błąd dynamiczny

1. Introduction

An analysis of errors generated by analogue systems intended for processing undetermined dynamic signals can be realized based on a previously determined mathematical model of such a system. This model can be presented by means of a transfer function, a complex frequency response and an impulse or step function. There are mutual mathematical relations between these types of models, which allow them to transform easily from one form into another.

In the measurement technique, in automation and signal processing theory, modelling in the frequency domain is most often carried out based on measurement of the amplitude and phase characteristics [1–6]. The method for determining both the model parameters and associated uncertainties is determined by the relevant standards, choice of which depends on the type and application of the system modelled [7]. However, there are practical applications, for example in biomedical measurements, for which it is impossible to apply the abovementioned modelling. In such cases, modelling in the time domain is applied on the basis of measurement of the response to step stimulus signal [8, 9]. For the order of model dynamics which is assumed according to the type of modelled system, the characteristic parameters of the step response are determined on the basis of its time recording. In order to minimize the error of modelling method, it is desirable to perform a series of readings of the response to successive switches of step stimulus signal. In this way, an accurate determination of the model parameters and associated uncertainties is possible. In this paper, modelling was performed on an example of a serial RLC circuit with the output signal from a capacitor. This system has the properties of a second-order model. The choice of such a system results from the ability to easily check the effectiveness of the applied modelling.

The mathematical model reflecting the dynamics of the system constitutes the basis for a series of theoretical studies that include, among other things, procedures for determining maximum dynamic errors. These errors are the response to an input signals with the constraints imposed on them, that is, the amplitude or both the amplitude and the rate of change [10, 11]. In [11, 12], it has been shown that any other input signal generates an error less than or at most equal to the maximum value. Depending on purpose of the dynamic system, one uses different error criteria. The most popular are the absolute error and the integral-square error.

In this paper, above criteria of the error for the case of input signals constrained only in amplitude, which have the 'bang-bang' shape are considered. For this type of input signal, it is possible to determine precisely the switching times for both error criteria. In the case of two simultaneous constraints, it is only possible for the absolute error [12, 13]. For both criteria, the values of maximum error are determined and the propagation of uncertainties associated with the parameters of the model is assessed using the procedure for determination of the errors. This was realized by checking all possible cases of increases or decreases of the model parameters by the values of associated uncertainties. Such research has been carried out so far on the basis of modelling in the frequency domain and only for the absolute error [14].

2. Modelling of second-order systems in the time domain

For the systems described by the following transfer function

$$K(s) = \frac{a\omega_0^2}{s^2 + 2\beta\omega_0s + \omega_0^2}, \quad (1)$$

where a is the amplification coefficient, ω_0 is the non-damped natural frequency and β is the damping factor, the step response is presented by

$$h(t) = \mathcal{L}^{-1}\left(\frac{K(s)}{s}\right) = a \left\{ 1 - \exp(-\beta\omega_0 t) \left[\cos(\omega_d t) + \frac{\omega_0\beta}{\omega_d} \sin(\omega_d t) \right] \right\}, \quad (2)$$

in which

$$\omega_d = \omega_0 \sqrt{1 - \beta^2} \quad (3)$$

is the damped natural frequency.

The step response becomes extreme at times when its derivative vanishes, that is, when the impulse response

$$k(t) = \frac{d}{dt}h(t) = \frac{a\omega_0^2}{\omega_d} \exp(-\beta\omega_0 t) \sin(\omega_d t) \quad (4)$$

is equal to zero.

The response $k(t)$ reaches zero if

$$\sin(\omega_d t_n) = 0 \quad (5)$$

that is, for the following case

$$\omega_d t_n = n\pi, \quad n = 0, 1, 2, \dots \quad (6)$$

Based on (3) and (6), we have

$$t_n = \frac{n\pi}{\omega_0 \sqrt{1 - \beta^2}}, \quad n = 0, 1, 2, \dots \quad (7)$$

Substituting t_1 into

$$\tilde{h}(t) = h(t) - a \quad (8)$$

one obtains the ratio of the maximum value of the step response to the steady state value which is called the overshoot and is calculated by

$$\Delta y = a \cdot \exp\left(\frac{-\beta\pi}{\sqrt{1 - \beta^2}}\right). \quad (9)$$

Transformation of (9) gives

$$\beta = \frac{\ln\left(\frac{\Delta y}{a}\right)}{\sqrt{\ln^2\left(\frac{\Delta y}{a}\right) + \pi^2}}. \quad (10)$$

Let us present the damped natural frequency in the form

$$\omega_d = \frac{2\pi}{T_d}, \quad (11)$$

where T_d is the period of damped oscillations.

Comparing the right-hand sides of (3) and (11), we finally have

$$\omega_0 = \frac{2\pi}{T_d \sqrt{1-\beta^2}} = \frac{2\pi}{T_d \sqrt{1 - \frac{\ln^2\left(\frac{\Delta y}{a}\right)}{\ln^2\left(\frac{\Delta y}{a}\right) + \pi^2}}}. \quad (12)$$

Figure 1 shows the step response and the principle of determination of the parameters a , Δy and T_d which are the basis for determination of parameters of the model (1).

The parameters a , Δy and T_d are calculated as arithmetic means from a series of measurements obtained for the positive step responses. Fig. 2 shows the a voltage step stimulus signal and the positive and negative responses.

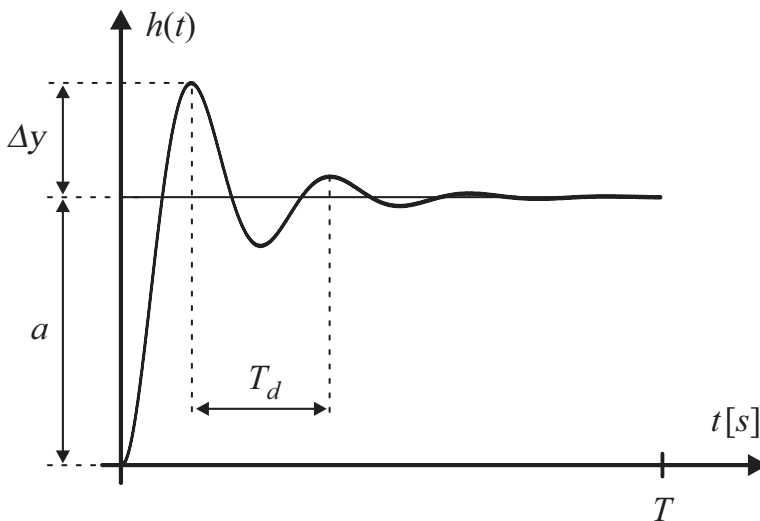


Fig. 1. Step response and its associated parameters

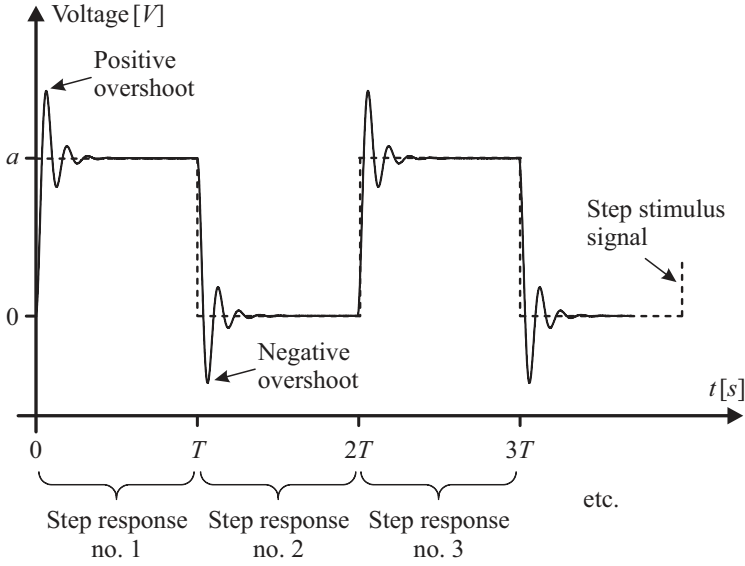


Fig. 2. Successive step responses

3. Determination of uncertainties associated with model parameters

The standard uncertainties $u(a)$, $u(\Delta y)$ and $u(T_d)$ associated with the parameters a , Δy and T_d as well as the expanded uncertainties related to the parameters of model (1) are calculated based on methods A and B [15, 16].

Utilizing method A, we have

$$u_A(x) = \sqrt{\frac{1}{M(M-1)} \sum_{m=1}^M (x_m - \bar{x})^2}, \quad (13)$$

where M is the number of measurement data, and

$$\bar{x} = \frac{1}{M} \sum_{m=1}^M x_m. \quad (14)$$

Value of M corresponds to the number of positive step responses.

Uncertainty of B type is calculated based on

$$u_B(x) = \sqrt{B_1^2 + B_2^2}, \quad (15)$$

where

$$B_1 = c_f \cdot \sigma_x / \sqrt{M} \quad (16)$$

and c_f and σ_x are the coverage factor and the standard deviation calculated for measurement data, respectively.

The variable B_2 calculated for a and Δy is

$$B_2 = \sqrt{u_a^2 + u_g^2 + u_o^2}, \quad (17)$$

where: u_a , u_g and u_o are the absolute, gain and offset uncertainties, respectively. The uncertainties under the root in (17) are calculated by

$$u_a = R_V \cdot u_g + R_V \cdot u_o + u_n \quad (18)$$

$$u_g = u_{rg} + g_t \cdot t_{chlic} + r_t \cdot t_{chlec}, \quad (19)$$

$$u_o = u_{ro} + o_t \cdot t_{chlic} + u_{INL}, \quad (20)$$

where: R_V , u_n , u_{rg} , g_t , t_{chlic} , r_t , t_{chlec} , u_{ro} , o_t and u_{INL} are the reading voltage range [V] the noise uncertainty [μV] the residual gain uncertainty [ppm of range] the gain tempco [ppm/ $^{\circ}C$] the temperature change from last internal calibration [$^{\circ}C$], the reference tempco [$^{\circ}C$], the temperature change from last external calibration [$^{\circ}C$], the residual offset uncertainty [ppm of range] the offset tempco [ppm of range/ $^{\circ}C$] and the integral nonlinearity uncertainty [ppm of range], respectively.

The noise uncertainty is calculated, as follows

$$u_n = r_n, \quad (21)$$

where r_n is a standard deviation of random noise [μV]

In the case of parameter T_d the variable B_2 is

$$B_2 = 1/s_r. \quad (22)$$

where s_r is the sampling rate.

The standard uncertainties $u(a)$, $u(\Delta y)$ and $u(T_d)$ are calculated based on (13) and (15), utilizing the following formula

$$u(x) = \sqrt{u_A(x)^2 + u_B(x)^2}. \quad (23)$$

The combined uncertainties associated with β and ω_0 are calculated utilizing the total derivative, as follows

$$u_c(\beta) = \sqrt{\left[\frac{\partial \beta}{\partial a} u(a) \right]^2 + \left[\frac{\partial \beta}{\partial \Delta y} u(\Delta y) \right]^2} \quad (24)$$

and

$$u_c(\omega_0) = \sqrt{\left[\frac{\partial \omega_0}{\partial a} u(a) \right]^2 + \left[\frac{\partial \omega_0}{\partial \Delta y} u(\Delta y) \right]^2 + \left[\frac{\partial \omega_0}{\partial T_d} u(T_d) \right]^2}. \quad (25)$$

Finally, the expanded uncertainties associated with parameters of the model (1) are

$$U(a) = c_f \cdot u(a), \quad (26)$$

$$U(\beta) = c_f \cdot u_c(\beta) \quad (27)$$

and

$$U(\omega_0) = c_f \cdot u_c(\omega_0). \quad (28)$$

4. Procedure for determining the absolute error

The absolute error represents the maximum possible value of the output signal $y(t)$ over the interval $[0, T]$ for the linear system described by the total impulse response

$$k(t) = k_r(t) - k_s(t), \quad (29)$$

where $k_r(t)$ and $k_s(t)$ are the impulse responses of the real system and its standard.

The standard is represented by a high-order low-pass filter with a transfer function defined by

$$K_s(s) = \frac{a}{\prod_{l=1}^L \left(\frac{s}{\omega_c} - e^{\frac{j(2l+L-1)\pi}{2L}} \right)}, \quad (30)$$

where L is the order of the filter.

The impulse responses $k_r(t)$ and $k_s(t)$ are obtained as the inverse Laplace transform of (1) and (30), respectively.

The maximum transient value of $y(t)$ can be achieved only for $t=T$. However, it is necessary to determine the input signal $x_0(t)$ of “bang-bang” type which generates the output $y(T)$. This output signal represents the absolute error denoted below by D .

The signal that maximizes the error D is determined based on a simple formula, as follows

$$u_0(t) = A \cdot \text{sign}[k(T-t)]. \quad (31)$$

The absolute error is calculated by means of

$$D = A \int_0^T |k(t)| dt, \quad (32)$$

where A denotes the magnitude constraint of signal $x(t)$ [12–14]. This constraint was assumed as equal to the amplification coefficient of the model (1).

5. Procedure for determining the integral-square error

Let us define by

$$\varsigma[k, x, y] = \frac{E[y]}{E[x]} \quad (33)$$

the energy transfer ratio over the interval $[0, T]$ and corresponding to the linear system described by the impulse response $k(t)$ defined by (29), input signal $x(t)$ and output $y(t)$.

The energy of the output signal $y(t)$ is represented by

$$E[y] = \int_0^T y(t)^2 dt = \int_0^T \left[\int_0^T K(t, \tau) x(\tau) d\tau \right] x(t) dt, \quad (34)$$

where

$$K(t, \tau) = \int_0^T k(t, \nu) k(\nu, \tau) d\nu \quad (35)$$

is the energy kernel which is the autocorrelation function calculated based on the impulse response $k(t)$.

It is obvious that there is such a signal $x_0(t) \in L^2$ that maximizes ζ . This signal has the “bang-bang” shape and satisfies the Fredholm integral equation of the second kind, as below

$$\int_0^T K(t, \tau) x(\tau) d\tau = \zeta x(t) \quad (36)$$

with the kernel defined by (35) [17].

The procedure for determining the signal $x_0(t)$ is carried out in three main steps:

1. Calculate the autocorrelation function

$$K(t) = \int_0^T k(\tau) k(t + \tau) d\tau. \quad (37)$$

2. Determine the initial input signal

$$x^0(t) = A \cdot \text{sign} [K(t)]. \quad (38)$$

3. Determine the signal $x_0(t)$ which has one constraint and maximizes the output energy based on the iteration algorithm, as follows

$$x^{i+1}(t) = A \cdot \text{sign} \left[\int_0^t K(t - \tau) x^i(\tau) d\tau \right] \text{ for } i = 0, 1, 2, \dots \quad (39)$$

This output energy corresponds to the integral-square error I_2 of a linear system [18].

The iteration algorithm is terminated when any switching times of signals $x^{i+1}(t)$ and $x^i(t)$ differ by the value of the assumed discretization step Δ . Based on the maximizing signal $x_0(t)$ the integral-square error can be calculated by

$$I_2 = \int_0^T \left[\int_0^t k(t - \tau) x_0(\tau) d\tau \right]^2 dt. \quad (40)$$

6. Results

The parameters of the model (1) were determined based on measurement of 34 positive step responses for a equals $1[V]$. The step stimulus signal was generated by NI-6221 measuring card with a sampling rate s_r equals $100[kS/s]$. This card was also used to determine the responses shown in Fig. 2.

Table 1 includes the values of the parameters a , Δy and T_d obtained in accordance with Fig. 1. The values of these parameters were tabulated with an accuracy of three significant digits $M=1, 2, \dots, M$, where M was assumes as equal to 34. The last row in Table 1, denoted as **Mean**, contains the mean values of all three parameters.

These values are assumed to be the estimates of parameters a , Δy and T_d and are indicated below as \tilde{a} , $\tilde{\Delta y}$ and \tilde{T}_d . The standard uncertainties associated with these parameters, determined based on (13), are as follows: $u_A(\tilde{a})=27[\mu V]$, $u_A(\tilde{\Delta y})=45[\mu V]$ and $u_A(\tilde{T}_d)=1,5[\mu s]$.

Table 1. Parameters based on successive step responses

m	a [mV]	Δy [mV]	T_d [ms]	m	a [mV]	Δy [mV]	T_d [ms]
1	994.026	440.113	1.290	18	994.288	439.689	1.290
2	994.041	439.613	1.280	19	994.302	439.513	1.280
3	994.051	440.250	1.280	20	994.311	439.666	1.280
4	994.056	439.760	1.300	21	994.337	439.478	1.290
5	994.076	440.063	1.290	22	994.349	439.628	1.280
6	994.115	439.862	1.280	23	994.362	439.453	1.270
7	994.113	440.025	1.280	24	994.387	439.267	1.290
8	994.113	439.703	1.290	25	994.397	439.257	1.300
9	994.151	440.150	1.290	26	994.406	439.410	1.270
10	994.163	439.652	1.290	27	994.428	439.710	1.290
11	994.158	439.657	1.300	28	994.430	439.548	1.280
12	994.202	439.452	1.280	29	994.465	439.512	1.290
13	994.200	439.616	1.290	30	994.488	439.489	1.270
14	994.232	439.422	1.270	31	994.493	439.484	1.280
15	994.260	439.556	1.290	32	994.506	439.148	1.280
16	994.261	439.554	1.300	33	994.533	439.283	1.300
17	994.261	439.393	1.290	34	994.559	439.418	1.280
				Mean	994.280	439.612	1.286

The coverage factor c_f was assumed as equal to 2, according to a 95% confidence level. The standard deviations calculated based on Table 1 are: $\sigma_a=0,16$ mV, $\sigma_{\Delta y}=0,26$ mV and $\sigma_{T_d}=8,8$ μ s.

The following values: $t_{chili}=1$ [°C], $t_{chle}=25$ [°C], $u_a=95$ [ppm of range], $u_{rg}=25$ [ppm/°C], $r_t=5$ [°C], $u_{ro}=25$ [ppm of range], $o_t=79$ [ppm of range/°C], $u_{INL}=76$ [ppm of range] and $r_n=30\mu V$ are provided by the manual of measuring card [19].

For the above data and based on (15–22), the uncertainties calculated by method B are: $u_b(\tilde{a})=1.984$ [mV], $u_b(\Delta\tilde{y})=1.985$ [mV] and $u_b(T_d)=11$ [μ s]. Then the standard uncertainties, calculated based on (23), are as follows: $u(\tilde{a})=1.984$ [mV], $u(\Delta\tilde{y})=1.985$ [mV] and $u(T_d)=11$ [μ s]. It can be seen that these uncertainties are affected by the uncertainty determined by method B. The combined uncertainties associated with β and ω_0 are: $u_c(\beta)=0,0014$ [mV] and $u_c(\omega_0)=42$ [rad/s]. Finally, the expanded uncertainties associated with the model parameters are calculated. These uncertainties and the model parameters are tabulated in Table 2. The parameter \tilde{a} corresponds to the mean value included in Table 1. The parameters $\tilde{\beta}$ and $\tilde{\omega}_0$ were calculated on the basis of relations (24) and (25).

Table 2. Estimates of the model parameters and associated uncertainties

Model parameters			Associated uncertainties		
\tilde{a} [mV]	$\tilde{\beta}$ [-]	$\tilde{\omega}_0$ [rad/s]	$U(\tilde{a})$ [mV]	$U(\tilde{\beta})$ [-]	$U(\tilde{\omega}_0)$ [rad/s]
994.280	0.2514	5050	3.968	0.0029	83

Table 3 contains the values of errors D and I_2 determined based on (32) and the iterative procedure executed utilizing relations (37)–(40), respectively. The fifteenth-order low-pass filter with a cut-off corresponding to a 10% decrease of the amplitude-frequency characteristic relative to its value for $\omega=0$ was adopted as a standard. This characteristic was determined based on the model (1) for the parameters contained in Table 2. The order of the model adopted constitutes the maximum value for which it is possible to carry out the calculations in the program MathCad15. The value of the amplitude constraint A was assumed to be equal to the value of estimate \tilde{a} .

The first row of Table 3 contains the errors without taking into account the influence of the uncertainties, whereas other rows include the errors for any case increase or decrease in the parameters by the values of the associated uncertainties.

Table 3. Parameters of the model and the maximum values of errors D and I_2

Change of parameters			D [Vs]	I_2 [mV ² s]
\tilde{a}	$\tilde{\beta}$	$\tilde{\omega}_0$	4.001	24.499
$\tilde{a}-U(\tilde{a})$	$\tilde{\beta}-U(\tilde{\beta})$	$\tilde{\omega}_0-U(\tilde{\omega}_0)$	3.995	24.586
$\tilde{a}-U(\tilde{a})$	$\tilde{\beta}+U(\tilde{\beta})$	$\tilde{\omega}_0-U(\tilde{\omega}_0)$	3.944	24.015
$\tilde{a}-U(\tilde{a})$	$\tilde{\beta}-U(\tilde{\beta})$	$\tilde{\omega}_0+U(\tilde{\omega}_0)$	3.996	24.747

Change of parameters			D [Vs]	I_2 [mV ² s]
$\tilde{a}-U(\tilde{a})$	$\tilde{\beta}+U(\tilde{\beta})$	$\tilde{\omega}_0+U(\tilde{\omega}_0)$	3.944	24.065
$\tilde{a}+U(\tilde{a})$	$\tilde{\beta}-U(\tilde{\beta})$	$\tilde{\omega}_0-U(\tilde{\omega}_0)$	4.060	24.981
$\tilde{a}+U(\tilde{a})$	$\tilde{\beta}+U(\tilde{\beta})$	$\tilde{\omega}_0-U(\tilde{\omega}_0)$	4.007	24.401
$\tilde{a}+U(\tilde{a})$	$\tilde{\beta}-U(\tilde{\beta})$	$\tilde{\omega}_0+U(\tilde{\omega}_0)$	4.061	25.145
$\tilde{a}+U(\tilde{a})$	$\tilde{\beta}+U(\tilde{\beta})$	$\tilde{\omega}_0+U(\tilde{\omega}_0)$	4.007	24.452

The highest values of both errors are obtained for the case when the estimates of parameters \tilde{a} and $\tilde{\omega}_0$ are increased by the values of associated uncertainties, while the estimate of parameter $\tilde{\beta}$ is decreased by the uncertainty. These values are 1.50% and 2.64% higher than the values without taking into account the uncertainties for the errors D and I_2 respectively.

7. Conclusions

The paper presents an assessment of the propagation of modelling uncertainty in the time domain by the procedure for determining both the absolute error and the integral-square error for the input signals constrained only in amplitude.

Based on 34 identification tests, the estimates of the model parameters are determined and the associated uncertainties are calculated. Then, the maximum dynamic errors were determined both for the model parameters only and for all cases of increases or decreases in the model parameters by the values of associated uncertainties. These errors were determined based on modelling of a serial RLC circuit as an example of a second-order system. As a standard that constitutes the reference for determination of the errors, the fifteenth-order low-pass filter was adopted.

The results showed that modelling uncertainty has an influence on dynamic errors, particularly for the integral square criterion. In comparison with modelling results in the frequency domain presented in [14], the time-domain modelling increases the uncertainty of parameter \tilde{a} more than 16-fold, but decreases the uncertainties of the parameters $\tilde{\beta}$ and $\tilde{\omega}_0$ more than 100-fold. Thus, in the case of modelling in the time domain presented in this paper, the influence of uncertainties on the absolute error was approximately 80 times lower. This is due to the fact that the method of modelling in the time domain is more accurate but cannot always be used, in particular for examination of sensors intended for measurement of vibration. However, the most important conclusion is that for both methods of modelling, the uncertainties associated with the model parameters have an essential influence on the value of errors and this effect should be taken into account during the calibration process based on the maximum dynamic errors.

References

- [1] Kollar I., *On Frequency-Domain Identification of Linear Systems*, IEEE Transactions on Instrumentation and Measurement, Vol. 42, Issue 1, 1993, 2–6.
- [2] Ljung L., *Some results on identifying linear systems using frequency domain data*, Proc. 32nd. IEEE Conf. Decis. Control, San Antonio, 1993, 3534–3538.
- [3] Guillaume P., *Frequency Response Measurements of Multivariable Systems using Nonlinear Averaging Techniques*, IEEE Transactions on Instrumentation and Measurement, Vol. 47, Issue 3, 1998, 796–800.
- [4] Pintelon R., Schoukens J., *System identification: A Frequency Domain Approach*, IEEE Press, Piscataway, New York 2001.
- [5] Isermann R., Münchhof M., *Identification of Dynamic Systems*, Springer-Verlag, Berlin Heidelberg, Dordrecht, London, New York 2010.
- [6] Tomczyk K., Sieja M., *Parametric Identification of System Model for the Charge Output Accelerometer*, Technical Transactions 2-E/2015, 235–245.
- [7] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, *Evaluation of measurement data – Supplement 1 to the “Guide to the expression of uncertainty in measurement” – Propagation of distributions using a Monte Carlo method*, 2008.
- [8] Assambo C., *Determination of the Parameters of the Skin-Electrode Impedance Model for ECG Measurement*. Proceedings of the 6-th WSEAS Int. Conf. on Electronics, Hardware, Wireless and Optical Communications, 2007, 90–95.
- [9] Tomczyk K., *Procedure for Correction of the ECG Signal Error Introduced by Skin-Electrode Interface*, Metrology and Measurement Systems, Vol. XVIII, No. 3, 2011, 461–470.
- [10] Rutland N.K., *The Principle of Matching: Practical Conditions for Systems with Inputs Restricted in Magnitude and Rate of Change*, IEEE Transaction on Automatic Control, Vol. 39, 1994, 550–553.
- [11] Layer E., *Non-standard Input Signals for the Calibration and Optimisation of the Measuring Systems*, Measurement, Vol. 34, 2003, 179–186.
- [12] Layer E., Tomczyk K., *Measurements. Modelling and Simulation of Dynamic Systems*, Springer-Verlag. Berlin Heidelberg, 2010.
- [13] Layer E., Tomczyk K., *Determination of Non-Standard Input Signal Maximizing the Absolute Error*. Metrology and Measurement Systems, Vol. 17, 2009, 199–208.
- [14] Tomczyk K., *Impact of Uncertainties in Accelerometer Modelling on the Maximum Values of Absolute Dynamic Error*. Measurement, Vol. 49, 2016, 71–78.
- [15] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML 2011. Supplement 2 to the Guide to the Expression of Uncertainty in Measurement, Extension to any number of output quantities JCGM 102:2011.
- [16] Kubisa S., *Intuicja i symulacja Monte Carlo Podstawa Analizy Niedokładności Pomiarów*, PAK, No. 9, 2007, 3–8.
- [17] Wyner A., *Spectra of Bounded Functions in Open Problems in Communication and Computation*, T.M. Cover and B. Gopinath, Eds. Springer Verlag, New York 1987, 46–48.

- [18] Honig M.L., Steiglitz K., *Maximizing the Output Energy of a Linear Channel with a Time and Amplitude Limited Input*, IEEE Transaction on Information Theory, Vol. 38, No. 3, 1992, 1041–1052.
- [19] NI 622x Specifications, National Instruments, ref. 372190G-01, jun. 2007
- [20] <http://www.ni.com/datasheet/pdf/en/ds-15>

Marta Bożym (m.bozym@po.opole.pl)

Agnieszka Bok

Department of Environmental Engineering, Faculty of Mechanical Engineering,
Opole University of Technology

ADVANTAGES AND DISADVANTAGES OF THE SOLAR DRYING OF SEWAGE SLUDGE IN POLAND

ZALETY I WADY SOLARNEJ SUSZENIA OSADÓW ŚCIEKOWYCH W POLSCE

Abstract

Solar drying has become a popular method of sewage sludge processing in Poland over recent years. This is due to the possibility of using a renewable source of energy and the fact that solar dried sludges can be used for energy purposes. Over the last ten years in Poland, almost thirty solar or hybrid greenhouses of sewage sludge have been built. The aim of this work is to present the advantages and disadvantages of the solar drying method of processing sewage sludge. The advantage of using solar greenhouses are the use of renewable energy sources for sludge drying, their low operational costs, the reduced sludge volume, their preparation for thermal utilisation, simple construction and easy operation. The disadvantages are that drying efficiency depends on climatic conditions and season, the need to use additional heat sources, the fact that facilities occupy a large area and the possibility of odors emission released during drying of the sludge.

Keywords: sewage sludge, treatment, solar drying

Streszczenie

Suszenie słoneczne jest coraz popularniejszą metodą suszenia osadów ściekowych w Polsce w ostatnich latach. Wynika to z możliwości wykorzystania odnawialnego źródła energii oraz faktu, że osuszone słońcem osady mogą być wykorzystywane do celów energetycznych. W ciągu ostatnich dziesięciu lat w Polsce zbudowano prawie trzydzieści słonecznych lub hybrydowych suszarni osadów ściekowych. Celem niniejszej pracy jest przedstawienie zalet i wad suszenia słonecznego stosowanego w przetwarzaniu osadów ściekowych. Zaletą zastosowania szklarni słonecznych jest wykorzystanie odnawialnych źródeł energii do suszenia osadów, niskie koszty operacyjne, uzyskanie zmniejszonej objętości osadu, przygotowanie do energetycznego wykorzystania, prosta konstrukcja i łatwa obsługa. Do wad należy zależność skuteczności suszenia od warunków klimatycznych i pory roku, konieczności wykorzystania dodatkowych źródeł ciepła, faktu, że obiekty zajmują duży obszar i możliwość emisji odorów emitowanych podczas suszenia osadu.

Słowa kluczowe: osady ściekowe, przetwarzanie, suszenie słoneczne

1. Introduction

The preferred direction in the application of sewage sludge in Poland is its agricultural utilization. Sewage sludge which cannot be utilised in such a way, may be thermally used. According to assumptions of the National Plan of the Waste Disposal 2022 [1] in the coming years in Poland the proportion of sewage sludge that undergoes thermal treatment will increase. For thermal treating methods of sewage sludge are mainly meant combustion, co-combustion in cement plants as well as alternative processes, so as the pyrolysis or gasification [2–4]. Sewage sludge is characterized by a calorific value similar to that of brown coal, usually 9-13 MJ/kg. Their calorific value can be increased by mixing with higher calorific value wastes [5–6]. The calorific value of fuels, also sewage sludge, depends on their moisture content. Sewage sludge contains a significant percentage of water and should be dehydrated prior to thermal utilization [7, 8]. The thermal disposal of sewage sludge is related immediately to the dry matter content (d.m.) achievement of required level – at least 40%. However, cement producers usually require a degree of drying of sludge to more than 90% d.m. [7]. The drying process of sewage sludge can be conducted in a conventional manner in a dedicated dryer, where energy source is meant to be: fossil fuel such as natural gas or use of alternative methods, eg. use of sunlight in solar greenhouses. However, the efficiency of such kinds of drying processes is usually insufficient for obtaining a high proportion of dry matter; therefore, wastewater treatment plants use additional energy sources such as biogas for underfloor heating or infrared radiators. It is recommended that solar dried sewage sludge should be used for thermal disposal rather than for agricultural purposes. [10]. The conditions of applying sewage sludge as alternative fuel, e.g. in cement industry, are: a high percentage of dry matter, the physico-chemical stability, appropriately low content of pollutants. Under Polish law [9], all installations for waste co-burning are obliged to monitor the release of pollutants (Table 1).

Table 1. Emission standards applicable to the installation for co-combustion of waste [9]

Contamination	Value	Unit
total dust	30	mg/m ³
HCl	10	
HF	1	
NO _x	500	
SO ₂	50	
TOC	10	
CO	2,000	
Cd + Tl	0.05	
Hg	0.05	
the sum of the content of heavy metals	0.5	
dioxins and furans	0.1	µg/m ³

Because of the need to meet emission standards, thermal installations for co-combustion of waste, including sewage sludge, the requirements concerning the quality of fuels combusted in such installations, that means maximum values of pollutants are determined. Cement plants require high quality substrates and fuels as they are incorporated into the cement [6, 11]. In order to fulfil requirements concerning emissions from waste co-combustion installations, it is necessary to determine the content of pollutants within the input fuel. It is a requirement that the levels of heavy metals, sulphur and chlorine concentration be identified, because these substances affect the formation of gaseous pollutants. Cement plants also require the determination of the alkali content (K and Na) in the waste fuel, as they can adversely affect the cement kiln operation and cement quality. Table 2 shows the characteristics of sewage sludge taken in 2012 (dewatered sludge), 2014 and 2016 (solar dried sludge) from the waste water treatment plant in Antoniów, Poland [7, 11] compared with the formal requirements concerning co-combustion in Polish cement plants. The characteristic of the wastewater treatment plant in Antoniów is presented below (Chapter 2).

Table 2. Characterisation of sewage sludge from the waste water treatment plant in Antoniów compared with the formal requirements concerning co-combustion in Polish cement plants

Parameter	Sewage sludge(Antoniów)	EURITIS [10]	Lafarge groups [12]	Polish cement plants [13]
dry matter content, %	no data (2012) 75 (2014) 90 (2016)	NA	NA	>70%
dust, %	26 (2012) 25 (2014) 24 (2016)	5	NA	NA
calorific value , MJ/kg	12.7*(2012) 10.5 (2014) 13.3 (2016)	15	>14	>13
heavy metals, mg/kg d. m.	688 (2012) 837 (2014) 392 (2016)	NA	<2500	<2500
Cl / S	0.07/0.25 (2012) no data (2014) 0.16/0.89 (2016)	0.5 /0.4	<0.2/ <2.5	<0.3/ <2.5

NA - no available

* - calculated for 95% d.m.

Sewage sludge from the sewage treatment plant in Antoniów meets the specified requirements for the cement plants with regard to the contaminant content (metals, sulphur, chlorine). However, the low dry matter content influences its calorific value. The Cement Plant located in Góraźdże (Poland), which is a potential recipient of dried sludge, requires a dry matter content of sewage sludge of more than 85%.



2. Technologies used in the solar drying of sewage sludge

The solar drying process relies upon the occurrence of a greenhouse effect within the drying areas. Solar radiation is reflected back into the drying room from the roof, thus increasing the internal temperature. The covering of roof and walls of greenhouse should be made of materials which allow the free permeation of radiation into the interior, these materials are rarely glass but rather, polycarbonate and polythene foils. Greenhouses for sewage sludge should also include ventilators or overblowing exhausters as part of their construction. A crucial element within the greenhouse hall is a rotary sludge tedder. The function of the tedder is sewage sludge granulation and its aeration, homogenization and acceleration of drying process. The thickness of the sewage sludge layer and its degree of dehydration affect the tedder efficiency. In Poland, solar greenhouses for the drying of sewage sludge, rotary tedders made by the German company WENDEWOLF are often installed. The tasks of these devices are the turning over, cutting and forming granules of dried sewage sludge inside the hall. Another solution is presented by a nave tedder of German HUBER company, which enables flipping the layer of drying sewage sludge over on the entire width of the hall [14]. Tedders recirculate the sewage sludge by transporting dried sludge from the rear of the drying hall to the front. Other tasks of the tedder also include: partitioning and moving the transfer of dried granules to bulk channels. Other technology – so-called THERMO-SYSTEM, relies on tedders of type “Electric Mole” or “Sludge Manager”. The first one is an automated electric vehicle, which moves along a track in drying area. It is possible to install this device on an uneven floor. “Sludge Manager” is a device which moves along a beam putted crosswise and along the storage surface. Its efficient functioning does not require a perfectly equal level of area, because it’s able to create the map of floor unevenness. Both devices are designed for small size greenhouses. A novel solution was applied by SOLIA technology, using the SOLIAMIX robot. The device gathers fresh sewage sludge into piles parallel to itself along the entire length of greenhouse hall [15]. Special feeders for the transport of sewage sludge, a digger and belt or voluted conveyors are usually used in those technology [16]. During the drying process, ventilation is a very important consideration. In solar greenhouses, the drying of sewage sludge is also supported by ventilators in the interior. Deodorisation systems are also used in the greenhouses to control odours. In order to accelerate the drying process in solar greenhouses, an additional source of heat can be used in the form of underfloor heating and infrared radiants [17–19]. The construction of drying halls with underfloor heating and the associated infrastructure require almost double the investment of halls of similar size without such facilities [19]. The drying rate in solar greenhouses depends on the thickness of the layer of sewage sludge and climatic conditions [20]. In the summer period, a layer of 1 to 10 cm dries in around 10 days; however, in the autumnal period, this process lasts four times longer [16–18]. In the winter time, in many solar greenhouse which don’t have an underfloor heating installation, sewage sludge is kept in piles to spring [11].

Over the last 10 years in Poland, almost 30 solar or hybrid greenhouses for the processing of sewage sludge have been built. Solar greenhouses are located in such Polish towns as Kozienice, Tuchów, Łañcut, Żary, Myszków, Antoniów, Kłodzko; hybrid greenhouses are found in Iława, Końskie, Chełm, Morąg, Żarów, Żagań, among others.

3. Characteristics of the sewage greenhouses in the Wastewater Treatment Plant in Antoniów (Poland)

The wastewater treatment plant in Antoniów produces around 2400 Mg of sewage sludge at 20% s.m. per year. After dehydration on the filter press, sewage sludge is dried in two greenhouses with a total area of 2,750 m² (width 12 m, length 113.5 m) (Fig. 1a). In each greenhouse, a gravity ventilation system and 12 mechanical ventilators are used. When the air humidity growth inside the greenhouse the roof windows are automatically opened and after exchanging the air they are closed. At the bottom of the walls of greenhouse, there are 10 cm expansion joints which support the operation of gravity ventilation (Fig. 1.2). A rotary tedder from Huber Technology mixes and distributes the sewage sludge on the hall and finally converts it into granules (Fig. 1.3). The weight of the sludge after drying is about 650 t/a. The weight reduction is about 73% and the average dry matter content is about 75%. It was planned that the granules of dried sludge be co-combusted in Cement Plant in Górażdże (Poland). However, the tested sludge does not meet the dry matter content requirements specified by these cement plant. Currently, the wastewater treatment plant uses dry sludge for agricultural purposes. The problem in this case is incomplete hygienisation, this is why the *Salmonella* bacteria is present in the dried sewage sludge.



Fig. 1. Solar greenhouse in the wastewater treatment plant in Antoniów: a) outdoor view of the solar greenhouse; b) gravity ventilation system; c) rotary tedder (Huber) with granules of dried sewage sludge

4. Advantages and disadvantages of the solar drying process of sewage sludge

The effectiveness of the drying process depends mainly on the degree of solar exposure on drying location, therefore locations in the area of south-east Poland are the most profitable. The effectiveness of the solar drying process also depends on the initial method of sludge dehydration, its amount and the type of technology applied. Due to the fact that drying by using sunlight is most effective in the spring-autumn period, the use of an additional source of heat, e.g. underfloor heating, is a recommended solution. The described process is profitable for wastewater treatment plants located a short distance from installations which enable the co-combusting of dried sewage sludge, e.g. in cement plants. Dried sewage sludge should meet the requirements for alternative fuels and specified by the legal regulations for co-incineration plants. Drying sludge causes odour emission, this is linked with an additional investment in a system of filters. The use of dried sewage sludge in agriculture seems unreasonable because the soil can be fertilized with sewage sludge in liquid form. After drying, granules of sewage sludge are characterised by their very compact structure, similar to ceramic materials. This may suggest that, after application to the soil, fertiliser components will be very slowly released. Most often, other methods, such as liming or composting, are used to treat and fertilizer production from sewage sludge.

In conclusion, the advantage of solar greenhouse is the use of renewable energy sources, which reduces the cost of drying. Dried sewage sludge are characterized by a smaller volume and a ceramic structure, which facilitates storage and transport. Nowadays in Poland the simplicity of getting funding of greenhouse build investment partially from the EU may be one of the advantage. The disadvantage of solar drying of sewage sludge is the necessity of building a dryer, in which should be use additional energy sources, in order to obtain a high content of dry matter. In addition, the drying efficiency depends on the degree of sunlight and temperature, which varies throughout the year. Another disadvantage is the need to find a co-combustion instalation plant for dried sludge. Therefore, it is a cost-effective method for wastewater treatment plants located near the cement plant, which reduces transport cost. This method should be applied at sewage treatment plants which plan to thermal utylisation of dried sewage sludges, rather than their agricultural use.

5. Alternative approach for the use of the solar drying of sewage sludge in Poland

The main direction of sewage sludge drying is the energy recovery; therefore, the most important point is leading sewage sludge to the maximum participation of dry matter what causes increase in calorific value. Drying capacity in solar greenhouses may be insufficient to achieve optimal drying of sludge as it depends on sunlight. Therefore, wastewater treatment plants which use solar greenhouses for sewage sludge drying, should consider the additional option of their heating. Polish wastewater treatment plants show that the maximal value of dry substance content that is possible to achieve in the summer months

is at a level of 75–80%. In Poland, sun-dried sludge is usually co-combusted in cement kilns. Recently in Poland an alternative approach for the use of dried sewage sludge is its environmental application, mainly as fertilizer. It seems obvious that the use of dried fertiliser in the farming sector doesn't make sense because for environmental targets, it is also possible to apply sewage sludge in liquid or mud form. Macronutrients are more easily and quickly absorbed by plants if the fertilizer is introduced into the soil in a hydrated form. Moreover, drying of sewage sludge used as fertilizer seems economically unprofitable and too time consuming. Another disadvantage of using dried sludge as fertilizer is incomplete hygienisation. In 2014, a piece of research on the value of sewage sludge granules as a fertiliser was conducted, these granules were obtained from a solar greenhouse at the sewage treatment plant in Antoniów in the Opole region of Poland [11]. The research has shown that the drying process has provided to reduction content of assimilable forms of macro- and microelement. There is a suggestion that after applying dried sewage sludge to the soil, all fertiliser nutrients are released at a slower rate than from hydrated sewage sludge. It may hinder the calculation of nitrogen dose entered into the soil with sewage sludge and that may lower the nutrient bioavailability for crops [21]. It is interesting that during this study, after 24h of water extraction the dried sewage sludge granules did not completely not fall apart. Another consideration with the agricultural use of this granules is reduce the amount of useful microorganisms at higher temperatures during drying process what can lead to limitation of decomposition of organic matter and availability of nutrients. Summarizing, other methods of sewage sludge stabilization and hygienisation eg liming or composting, should be considered for their natural use than solar drying, which is time consuming and requires a greenhouse.

6. Summary

In summary, the advantages of using solar greenhouses are the use of renewable energy sources for sludge drying, the low operating costs, the reduced sludge volume and correct preparation for thermal disposal, simple construction and ease of operation of instalation. The disadvantages are the fact that drying efficiency depends on the climatic conditions and the season, the need to use an additional heat source, the fact that the required facilities occupy a large area and the possibility of odour emissions released during the drying of the sludge. The main direction of using dried sludge is their use for energy purposes. Co-combustion in cement kilns is the most ecological solution in view of waste-free technologies, due to the possibilities of heavy metals binding to cement and the limitation of the emission of gaseous pollutants. The thermal utylisation of sludge is limited by its high hydratation level, its adverse consistency, its low calorific value and pollution content.

However the solar drying of sewage sludge does not seem to be the most appropriate method of treatment for the purpose of use in agriculture, which is preferred way of using sewage sludge in Poland again.



References

- [1] Resolution No. 88 of the Council of Ministers of 1 July 2016 on the “National Waste Management Plan 2022”, *Monitor Polski* poz. 284.
- [2] Bień J.D., *Zagospodarowanie komunalnych osadów ściekowych metodami termicznymi*, *Inżynieria i Ochrona Środowiska*, 15(4), 2012, 439–449.
- [3] Bień J., Neczaj E., Worwąg M., Grosser A., Nowak D., Milczarek M., Janik M., *Kierunki zagospodarowania osadów ściekowych w Polsce po roku 2013*, *Inżynieria i Ochrona Środowiska*, 14(4), 2011, 375–384.
- [4] Lutzner K., *Uzdatnianie ścieków i osadów*, *Ekologia i Technika*, 2(8), 1994, 14–17.
- [5] Wzorek M., Król A., *Analysis of selected properties of sewage sludge within the consideration of its energetic application in forming alternative fuels*, *Polish Journal of Environmental Studies*, 6, 2009, 132–136.
- [6] Wzorek M., *Characterisation of the properties of alternative fuels containing sewage sludge*, *Fuel Processing Technology*, 104, 2012, 80–89.
- [7] Bożym M., *Wymagania jakościowe stawiane osadom ściekowym spalany w krajowych cementowniach/Sewage sludge quality standards required by Polish cement plants*, *Chemik* 67, 2013, 10, 1019–1024.
- [8] Burzała B., *Termiczne przekształcanie osadów ściekowych jako jedna z metod ich utylizacji*, *Nowa Energia*, 1(37), 2014, 29–32.
- [9] Regulation of the Minister of the Environment of 4 November 2014 on emission standards for certain types of installations, combustion sources and combustion or co-incineration facilities, *Dz.U.* 2014 nr 0 poz. 1546.
- [10] Bożym M., *Możliwości wykorzystania osadów ściekowych w przemyśle ceramicznym*, *Materiały Budowlane* 12(460), 2010, 1–3.
- [11] Bożym M., *Ocena możliwości wykorzystania suszonych solarnie osadów ściekowych/Evaluation of the possibilities of using of solar dried sewage sludge*, *Chemik* 69(10), 2015, 666–669.
- [12] Sarna M., Mokrzycki E., Uliasz-Bocheńczyk A., *Paliwa alternatywne z odpadów dla cementowni*. Doświadczenia Lafarge Cement Polska S.A., *Zeszyty Naukowe Wydziału Budownictwa i Inżynierii Środowiska Politechniki Koszalińskiej*, 21, Seria: *Inżynieria Środowiska* Wydawnictwo Uczelniane Politechniki Koszalińskiej, Koszalin 2003, 309–316.
- [13] Duda J., *Współspalanie węgla i paliw alternatywnych w cementowych piecach obrotowych*. *Prace Instytutu Mineralnych Materiałów Budowlanych*, 35/36, 2003, 7–24.
- [14] Sobczyk R., Dembińska J., *Suszarnia słoneczna osadów w Antoniowie*, *Forum Eksploatatora*, 4, 2013.
- [15] Sobczyk R., Sypuła M., *Słoneczne suszenie osadów ściekowych w technologii Solia*, *Forum Eksploatatora*, 3, 2012.
- [16] Trojanowska K., *Energetyczne aspekty solarnego suszenia osadów ściekowych*, [in:] Z. Sadecka, *Oczyszczanie ścieków i przeróbka osadów ściekowych*, OWUZ, Zielona Góra 2010, 159–165.

- [17] Trojanowska K., *Ogrzewanie podłogowe w solarnych suszarniach osadów*, *Wodociągi i Kanalizacja*, 9, 2010, 24–27.
- [18] Trojanowska K., *Suszenie osadów ściekowych energią słońca*, *Przegląd Komunalny* 10, 2010, 46–48.
- [19] Sadecka Z., *Suszenie osadów – hybrydowe?*, *Materiały III Ogólnopolskiej Konferencji Szkoleniowej “Metody zagospodarowania osadów ściekowych”*, Chorzów, 2012, 86–96.
- [20] Yilmaz E., Wzorek M., *Assessment of the impact of various parameters on solar drying process in Aydin region in Turkey*, 3rd International Conference on Sustainable Solid Waste Management, 2–4 July 2015 Tinos.
- [21] Regulation of the Minister of the Environment of February 6, 2015 on municipal sewage sludge, *Dz.U.* 2015 poz. 257.



Marek Kubala (qmq@vistula.pk.edu.pl)

Institute of Water Supply and Environmental Protection, Faculty of Environmental Engineering, Cracow University of Technology

THE USEFULNESS OF CLUSTER ANALYSIS IN THE ANALYSIS OF DATA
OBTAINED IN THE MONITORING OF THE WATER ENVIRONMENT

PRZYDATNOŚĆ ANALIZY SKUPIEŃ W ANALIZIE DANYCH POZYSKANYCH
W MONITORINGU ŚRODOWISKA WODNEGO

Abstract

Data obtained through the monitoring of the water environment often includes a number of indicators, and is frequently collected from a large area or over a long period of time. Analysis of such data can be problematic. The division of elements which have a certain degree of similarity into subgroups may facilitate data analysis and provide indications as to the direction of the analysis. One tool for the separation of such groups of similar elements is cluster analysis.

This paper describes the two most commonly used cluster analysis algorithms and summarises the results of several applications of cluster analysis in water monitoring.

Keywords: water environment monitoring, cluster analysis, hierarchical clustering, k-means clustering

Streszczenie

Dane monitoringu środowiska wodnego zawierają często pomiary wielu wskaźników, a także bywają zbierane z dużego obszaru czy w długim okresie. Analiza takich danych może być utrudniona. Podział ich na podgrupy, których elementy wykazują pewne podobieństwo, może przyczynić się do łatwiejszej analizy oraz dostarczyć przesłanek co do jej kierunku. Jednym z narzędzi wydzielenia takich grup podobieństwa jest analiza skupień. Praca przedstawia opis dwóch najczęściej stosowanych algorytmów analizy skupień oraz streszcza rezultaty kilku zastosowań analizy skupień w monitoringu środowiska wodnego.

Słowa kluczowe: monitoring środowiska wodnego, analiza skupień, metody hierarchiczne, algorytm k-średnich

1. Introduction

Monitoring the water environment is often a matter of the collection and analysis of data (indicators) obtained at many points across a large area or over a long period of time. In addition, in many cases monitoring is not limited to measuring only one indicator but operates in a multidimensional metric space. In such cases, the relationship between them may be difficult to establish.

Generally, complex environmental processes may depend on many factors that are not always recognised. Measuring only selected indicators can lead to revealing completely different relationships between them in one area than in another area. Similarly, measurements conducted over a long period of time may lead to situations in which one set of data is influenced by different processes occurring in the environment leading to completely different dependencies than in another set of data. If such disparate spatial or temporal areas are not properly isolated, the correlation analysis between the different indicators in the set will, at best, not produce any positive results; at worst, this leads to the drawing of erroneous conclusions. Even if data diversity is expected, a strict spatial or temporal criterion may not be available to separate the relevant data groups. In addition, in the case of measurements with too little spatial density, it may not be possible to separate zones with different characteristics, especially when the different zones are mutually intertwined and one measurement point lies in one zone.

The first approach in this situation may be the graphical analysis of data, such as the analysis of the distribution of a selected indicator. If its distribution is characterised by several distinctly separated modes, this gives the basis for exploring the reasons for such a distribution [1]. A similar analysis of the density of measurements can be made either in two-dimensional space (on the basis of two indicators) or three-dimensionally. For larger dimensions, this analysis is no longer possible. In addition, usually the groups are not clearly separated, or the data set contains too few measurement points to make the division clear. In this situation, appropriate tools are needed to support the decision-making process with regard to division. Such tools include methods of cluster analysis.

2. Cluster analysis

Cluster analysis, or clustering, is the process of dividing a set of objects into subsets (clusters) in such a way that objects belonging to the same cluster are somehow more similar to one another than to objects from other clusters. This task is not defined by a unique algorithm, and its choice may depend not only on the type of data we want to analyse but also on the choice of the cluster definition itself. This term is usually understood as referring to a group of objects that are not too far apart, or areas of data space with a particular density of measurement points or intervals of statistical distribution. The appropriate clustering algorithm depends on the type of data being analysed and the clustering goal. This also applies to parameters such as the distance function, the density threshold and the number of expected clusters. Therefore, cluster analysis

is not only limited to the automatic application of an appropriate algorithm but also involves a gradual decision-making process in subsequent stages of its implementation, starting with the choice of variables in which the data set is divided (if the data set is multidimensional), and finally evaluating the real value of obtained division [1].

Based on several typical cluster models, the following main algorithms can be distinguished: hierarchical clustering algorithms, centroid algorithms, and algorithms based on the spatial density of objects and statistical distribution functions [2]. We will discuss the first two types of algorithms.

3. Hierarchical clustering

Hierarchical cluster analysis is designed to build a hierarchy of clusters with different levels of fragmentation. Following the method of building the hierarchy, the algorithms are divided into agglomerative and divisive types. Agglomerative algorithms start from the level where each observation has its own cluster and gradually, progressing up the hierarchical ladder, the individual clusters are joined together. By contrast, divisive algorithms take the lead from one cluster which is the set of all observations and the next steps of the hierarchy depend on successive divisions; however, they are not universally available and rarely applied. This is necessary to determine which clusters should be connected in the next step (if the agglomeration method is used). To accomplish this, the clustering criterion and the metric that will calculate the distance between clusters must be selected. There are three main criteria: complete linkage, single linkage, and average linkage.

The complete-linkage criterion (also called the diameter or the maximum method) interprets the distance between clusters as the maximum distance from any element of one cluster to any element of the other cluster. By contrast, the single-linkage criterion (also called the connectedness or minimum method) interprets the distance between clusters as the minimum distance from any element of one cluster to any element of the other cluster. Finally, the average linkage criterion takes on the average value of distances between elements of clusters. A variant of this criterion is the median criterion, which is more proof to outliers.

To calculate the distance, one has to measure it. The most commonly chosen measurements are the Euclidean distance, the square Euclidean distance, the taxicab distance and the maximum distance. The Euclidean distance is simply the straight-line distance between two points. This value can be squared, as the square Euclidean distance, to place greater weight on objects that are further apart. The taxicab distance is the sum of absolute differences between the coordinates of two points. The maximum distance is defined as the maximum value of the absolute differences between the coordinates of two points.

All these distance measures are sensitive to the scale of the values of the coordinates of the points. In order to eliminate the dominance of a parameter with high values relative to others, the parameters involved in calculating the distance should be normalised. Another way to solve this problem is to replace the Euclidean distance with the Mahalanobis distance [3].



Having already selected the linking criterion and the metric, we use the following algorithm [4]:

- 1) Assign a separate cluster to each object. In the beginning, there are as many clusters as there are points. In this case, the distance between clusters is equal to the distance between points.
- 2) Identify the two nearest clusters and combine them into one. The number of clusters is reduced by one.
- 3) Calculate the distance between the new cluster and all others using the clustering criterion.
- 4) Repeat steps 2 and 3 until you get a single cluster containing all of the points.

The algorithm is very simple and intuitive; its drawback is that its computational complexity is of the order of $O(n^3)$.

4. Centroid clustering

In this method, the cluster is represented by a point of gravity (centroid), which is not necessarily a component of the data set. One of the simplest algorithms of this method is the k-means algorithm. It leads to the simple and easy division of the data set into a given number of clusters [5]. The first step is to set the centroid for each cluster. It is very important to deploy them appropriately because this algorithm finds only the local minimum of the target function, so the clustering depends on the initial centroid spacing. The best choice is to place them as far apart as possible. In this case, an objective function is a squared error function.

The next step is to assign each point belonging to the data set to the closest centroid. After the first grouping, the centroids' positions are recalculated as the gravity points of the clusters formed in the previous step. After creating a new set of centroids, we introduce new relationships between data points and the closest new centroid. We repeat this operation as long as the gravity points change their position.

Formally, this algorithm aims to minimise the objective function, which in this case is a square error function. Of course, the behaviour of the algorithm depends on the metric we assume and on the initial distribution of gravity points, as has already been said. Therefore, it is recommended to repeat it several times with different starting points to avoid stopping at the local minimum.

5. Examples of applications

The following cases are examples in which cluster analysis gave interesting results in water monitoring.

The first example is the application of cluster analysis to investigate the impact of floods on the quality of waters of the Goczałkowice Reservoir [6]. This study analysed the

values of 22 water quality indicators measured 42–48 times in 2010; during this year, the reservoir has received three floods. This occurred 17/05/10, 02/06/10 and in the first days of September. For all measurements from 2010, cluster analysis was performed with both methods described previously for physicochemical indicators. Both methods separated a group of three measurements, namely those taken on 19/05/10, 25/05/10 and 07/09/10, the remaining measurements were attributed to the second group. These dates correspond to the aforementioned floods. To investigate whether this division is real, the significance of the differences between the groups for these indicators was checked by the Mann-Whitney U test. In 16 out of 22 cases, the differences were statistically significant. Similar results were obtained for bacteriological water quality indicators.

Another example is the study of the differentiation of heavy metal content in bottom sediments of dam reservoirs [7–9]. For example [8], cluster analysis enabled the separation of data due to similar proportions of granulometric fractions – for the Czorsztyński Reservoir – 5, and for the Goczałkowice Reservoir – 4. It has been found that for such a division, there is also a statistically significant difference in the content of lead in the sediment. This makes it clear that cluster analysis helps the isolation of sediment groups, the parameters of which differ significantly.

It was also possible to isolate similarity groups of photosynthetic dyes extracted from *Chlorella vulgaris* algae grown in waters from individual points of the Goczałkowice Reservoir collected at regular intervals [10, 11]. The similarity cluster identified three groups of measurements [11]. The first of these corresponded to the water quality during the occurrence of flooding and just after the flood. The second group corresponded to the water quality from the pre-flood period and from the measuring points located near the Vistula tributary from the whole period, with the exception of the first group. The third group concerned the measurements from the autumn period, when the vegetation clearly disappears.

Such a division seems to be very intuitive and obvious after its discovery, but no other standard method was able to obtain it.

6. Conclusions

The above examples show that cluster analysis is capable of separating data areas with different environmental quality indicators. This is possible both for data differentiated over time [6] and for spatially differentiated data [8]. This is also possible when differentiation of indicators simultaneously takes place both in time and in space [11].

The above examples clearly show that cluster analysis is a useful tool for data analysis in water monitoring.

I am deeply grateful to Ph.D. D.Sc. Anna Czaplicka and Ph.D. Zbigniew Ślusarczyk for making their work available to me and giving me valuable explanations. I am also grateful to Prof. Agnieszka Generowicz, whose persistent encouragement and help contributed to the publication of this work.



References

- [1] Han J., Kamber M., Pei J., *Data Mining. Concepts and Techniques*, Elsevier, 2012.
- [2] Jain A.K., Murty M.N., Flynn P.J., *Data clustering: a review*, ACM Computing Surveys, Vol. 31, No. 3, September 1999, 264–323.
- [3] Mahalanobis P.C., *On the generalized distance in statistics*, Proceedings of the National Institute of Sciences of India, Vol. 2, No. 1, 1936, 49–55.
- [4] Johnson S.C., *Hierarchical Clustering Schemes*, Psychometrika, Vol. 32, issue 3, 1067, 241–254.
- [5] MacQueen J.B., *Some Methods for Classification and Analysis of Multivariate Observations*, Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1, University of California Press, Berkeley 1967, 281–297.
- [6] Ślusarczyk Z., Czaplicka-Kotas A., *Wpływ powodzi w roku 2010 na jakość wód Zbiornika Goczałkowice*, Czasopismo Techniczne, 2-Ś/2012, 261–270.
- [7] Czaplicka A., Bazan S., Szarek-Gwiazda E., Ślusarczyk Z., *Spatial distribution of manganese and iron in sediments of the Czorsztyn Reservoir*, Environment Protection Engineering, Vol. 42, No. 4, 2016, 179–188.
- [8] Czaplicka A., Szarek-Gwiazda E., Ślusarczyk Z., *Factors influencing the accumulation of Pb in sediments of deep and shallow dam reservoirs*, Oceanological and Hydrobiological Studies, Vol. 46, issue 2, 2017, 174–185.
- [9] Czaplicka A., Ślusarczyk Z., Szarek-Gwiazda E., Bazan S., *Rozkład przestrzenny żelaza i manganu w osadach dennych Zbiornika Goczałkowice*, Ochrona Środowiska 3/2017, 47–54.
- [10] Czaplicka-Kotas A., *Badania wpływu jakości wody na wytwarzanie barwników fotosyntetycznych w komórkach glonów Chlorella vulgaris na potrzeby biomonitoringu wód powierzchniowych*, Ochrona Środowiska 1/2007, 27–33.
- [11] Czaplicka-Kotas A., Lodowska J., Wilczok A., Ślusarczyk Z., *Changes of photosynthetic pigments concentration in the synchronous culture of Chlorella vulgaris as an indicator of water quality in Goczałkowice Reservoir*, Archives of Environmental Protection, Vol. 35, No.1, 2009, 65–73.

Renata Sikorska-Bączek (sikorska@pk.edu.pl)

Institute of thermal Engineering and Air Protection, Faculty of Environmental Engineering, Cracow University of Technology

THE PROPOSED DESIGN OF A REFRIGERATION SYSTEM WITH
AN ENERGY-EFFICIENT CHILLER WITH ADJUSTABLE CAPACITY
FOR A FRUIT STORAGE CHAMBER

INSTALACJA CHŁODNICZA Z ENERGOOSZCZĘDNYM AGREGATEM
Z REGULOWANĄ WYDAJNOŚCIĄ NA POTRZEBY KOMORY
DO PRZECHOWYWANIA OWOCÓW

Abstract

This article presents the design of a refrigeration system with an energy-efficient chiller for a fruit storage chamber. This is a direct evaporating refrigeration system with an energy-efficient, multi-compressor chiller with an air-cooled condenser. Refrigeration systems with multi-compressor chillers, adjustable efficiency and system components that enable efficient and stable operation will reduce electricity consumption and increase reliability of the system during operation.

Keywords: fruit storage chamber, multi-compressor chiller, energy-efficient chiller

Streszczenie

W artykule przedstawiono projekt instalacji chłodniczej z energooszczędnym agregatem chłodniczym na potrzeby komory do przechowywania owoców. Jest to instalacja chłodnicza na bezpośrednie odparowanie z energooszczędnym wielosprężarkowym agregatem chłodniczym ze skraplaczem chłodzonym powietrzem. Instalacja chłodnicza z wielosprężarkowym agregatem chłodniczym z regulowaną wydajnością i elementami instalacji, które pozwolą na jego efektywną i stabilną pracę, spowoduje obniżenie zużycia energii elektrycznej oraz zwiększenie niezawodności instalacji w trakcie jej eksploatacji.

Słowa kluczowe: komora chłodnicza do przechowywania owoców, wielosprężarkowy agregat chłodniczy, energooszczędny agregat chłodniczy

1. Fruit storage conditions

A fundamental requirement of fruit storage is to limit both losses in terms of the quantity of produce and the degradation of produce quality in the period from harvest to sale. Processes taking place in the fruit, such as respiration and transpiration, demonstrate that after harvest, fruit items remain as living organisms. As a result of the use of storage chambers, consumers have access to fresh fruit produce throughout the year; however, despite the use of modern technology, produce losses within storage amount to around 15–40% [2]. It is possible to reduce these losses; however, this entails high investment costs.

The technique used in refrigeration chambers maintains stable and optimal levels of both temperature and humidity. Apples should be harvested about 1 to 2 weeks before they reach maturity. The storage period of fruit depends on both the produce and the technology used; for apples, it amounts to 3–10 months. The storage conditions for apples are as follows: temperature -1 to 3°C; relative humidity 85 to 90% [2]. The respiration of stored products depends on temperature as well as the oxygen and carbon dioxide content. If the temperature in the chamber is low, the respiration process slows down, resulting in lower consumption of hydrocarbons and less heat generated. The same effect can be achieved by reducing the oxygen content and increasing the carbon dioxide content. The respiration of fruit generates heat – this is important when calculating the heat balance of the chamber. The amount of heat emitted depends on the variety of fruit being stored. The amount of moisture produced depends on the humidity and temperature of the space where the fruit is located. It is therefore very important that the humidity in the storage chamber is kept as high as possible because the higher the relative humidity, the smaller the loss of weight from the product. This can be achieved through the use of large coolers that allow the maintaining of lower temperature differences.

When selecting a cooler, it should be remembered that the lower the temperature difference the smaller the product losses. Thus, the optimal level of humidity is the level at which there is sufficient inhibition of the growth of fungus-causing rotting, and also, a level at which the loss of weight is not too high. The most important parameter in storing fruit is temperature; by lowering it, we can slow down the rate of ripening, which increases the storage period. Fruit should be placed in a low temperature chamber as soon as possible after harvest; however, it is important not to freeze the fruit. Therefore, the chamber thermostat is set at 1–2°C above freezing point.

In order to speed up the cooling process of fruit after loading, the refrigeration system is switched on a few days before delivery so that the walls, floor and ceiling lose their accumulated heat and the temperature in the chamber stabilises at around 0°C. Fruit at the temperature of 15–20°C should be cooled to 2–6°C within 24 hours after being harvested. It should take a minimum of 8–10 days to load the chamber. The fruit contains a large amount of water – this is easily lost by transpiration when the relative humidity of the air in the chamber is low. The smallest weight loss occurs when the fruit is stored at the correct temperature and relative humidity. Therefore, the refrigeration unit must be properly selected and operated.

A sufficiently large surface of the cooler's heat exchanger is important and often underestimated; a large surface area of the cooler enables the maintaining of a higher

relative humidity. In such conditions water vapour deficiency decreases, resulting in reduced transpiration intensity and reduced weight loss of apples. The air flowing through the refrigeration chambers receives the heat that is emitted by the products stored within and ensures a uniform temperature throughout the chamber.

Proper air circulation has a strong impact on proper fruit storage. Uneven temperature in the chambers is often caused by improper cold air distribution. Incorrect loading of the chamber often results in the ineffective cooling of products stored in crates. Correct loading of the chamber should be carried out so that the air from the cooler flows freely around the products [6]. In chambers with good air circulation, the temperature difference at various points in the chamber should not exceed 0.5°C. The air flow follows the path of least resistance, so the distance between the crates and the walls and ceiling should be equal. To ensure good circulation of chilled air, the distance between the rows of pallet units must be 5 to 10 cm, and 10 to 20 cm from the side walls. The rows of pallets should always be aligned according to the direction of the flow of cooling air.

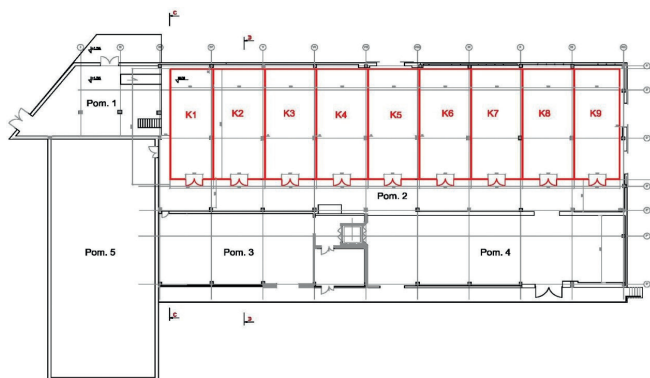


Fig. 1. Ground plan of the facility

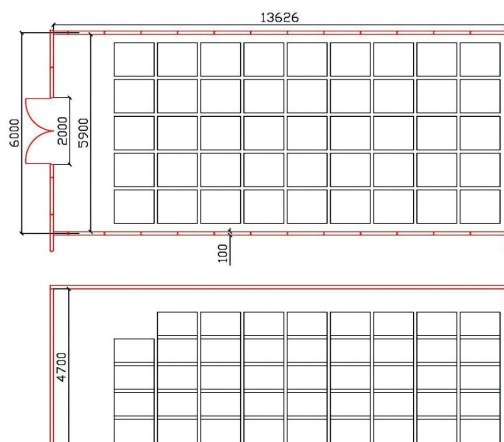


Fig. 2. Plan and cross section of the fruit storage chamber

2. Description and characteristics of the cold storage facility

The facility for which the refrigeration system has been designed is an existing building adapted to be a fruit storage facility, it is a one-storey building divided into five rooms. Room 1 is used as the cooling plant, room 2 features nine refrigeration chambers (K1-K9) built by an outside construction company. Rooms 3-5 will to be used and adapted by the investor. The building features gravity ventilation. Room 2 has nine identical refrigeration chambers, each with a usable area of 80.4 m^2 and a volume of 378 m^3 , in total, this equates to 723.6 m^2 of usable area and $3,402 \text{ m}^3$ of volume. The dimensions of the chambers are shown in Fig. 2. The chambers will be used for storing fruit in wooden crates with dimensions of $1.2 \times 1 \times 0.8 \text{ m}$. Each chamber will contain 220 crates, which gives a total of 1,980 crates in all chambers. Assuming that each crate holds 500 kg of apples, each chamber has a maximum load of 110 tons and 990 tons for the entire cold storage facility. Crates will be arranged in the chamber in columns and stacked to a height of five crates at a distance of 20 cm from the walls and 10 cm between each crate. In the first row, all the columns directly in front of the air cooler have a height of four crates so that the air directly blowing on them does not freeze the fruit. The arrangement of the crates in the chambers is shown in Fig. 2. In order to reduce the demand placed on the cooling system, whilst at the same time lowering the running costs, chambers will be loaded gradually so that the process of fully loading the entire chamber will take nine days. The daily loading of a single chamber will amount to 12.5 tons and will occupy one row of the chamber. The initial temperature of the fruit to be cooled to the required temperature is 25°C . Apples should be stored under appropriate conditions in the chambers. Therefore, the temperature that should be maintained in the chamber is -1°C to 3°C . The relative humidity of the air should be maintained at 85–90%. Under these conditions, apples can be stored in the chambers for 3 to 10 months [6].

3. Heat balance of the cold storage

The basic purpose of thermal calculations is to determine, for each cold storage room, the required refrigeration capacity of the devices installed therein. This capacity should be sufficient for the removal of heat from the room and thus maintain the thermal parameters of the refrigerated environment at the assumed required level. With these calculations, we can also determine the required efficiency of the cooling plant which will chill the cold storage rooms [1].

We can distinguish between two types of heat balance – evaporator heat balance and compressor heat balance. The evaporator heat balance determines the required heat output of the device, usually an evaporator, which cools the room. This balance is equal to the sum of the heat fluxes fed into the room. The sum of the heat fluxes fed to a set of rooms supported by a single chiller is called the compressor heat balance. If one room is supplied from one refrigerating system, the evaporator and compressor heat balance is the same. The size of the refrigeration system components is calculated on the basis of the heat balance, i.e. the combination of the values of the heat flux delivered in the chilled room from all sources.

The equation for daily heat gains is:

$$Q_{ch} = Q_p + Q_{wp} + Q_T + Q_k + Q_{u1} + Q_{u2} + Q_o + Q_L + Q_{sz} + Q_n \text{ [kWh/day]} \quad (1)$$

where:

- Q_p – heat gains through baffles [kWh/day];
- Q_{wp} – heat gains brought about by air exchange [kWh/day];
- Q_T – heat gains brought from stored goods [kWh/day];
- Q_k – heat gains brought together with packages and pallets [kWh/day];
- Q_{u1} – heat gain from equipment (forklift trucks) [kWh/day];
- Q_{u2} – heat gains from cooler fans [kWh/day];
- Q_o – heat gain from lighting [kWh/day];
- Q_L – heat gains from people [kWh/day];
- Q_{sz} – heat gains due to defrosting [kWh/day];
- Q_n – other heat gains [kWh/day].

The value of Q_{ch} , which is the sum of all components of the evaporator heat balance determined for individual rooms, is the basis for the selection of the air cooler for the cold storage rooms. After calculating Q_{ch} , we can calculate the total heat flux of the compressor heat balance, which consists of the evaporative heat loads of all the refrigeration chambers fed from one refrigeration system in this design [5]. The total compressor heat balance is the basis for the selection of the chiller and the rest of the cooling system. The summary of all the heat gains on the first loading day for chamber K2 is shown in Table 1:

Table 1. Heat gains on the first loading day for chamber K2

Lp.	Source of heat gains	Value [kWh / day]	Percentage [%]
1	through baffles	26.2	4.8
2	by air exchange	25.2	4.6
3	from stored goods	262.4	47.8
4	packages and pallets	24.3	4.4
5	equipment (forklift)	66.0	12.0
6	cooler fans	64.6	11.8
7	lighting	9.0	1.6
8	people	1.4	0.3
9	defrosting	43.8	8.0
10	<i>other</i>	26.2	4.8
Total		549.1	100.0



As we can see from the data in Table 1, the main source of heat in the chamber is the goods stored there. The heat brought to the chamber from the stored goods is 47.8% – this accounts for almost half of the total heat gain of the whole chamber. To properly select the efficiency of the cooler, we must consider the scenario where the heat gains in the chamber are the highest. In the case of the fruit cold storage chamber, we have to take into account the heat produced through respiration that occurs throughout the fruit storage period. As shown in Fig. 3, the highest heat gains for the analysed chamber occur on the eighth day of loading, these amount to **592.5 [kWh/day]** and the air coolers for this chamber should be selected with reference to this value. Assuming the working time $\tau_d = 16\text{h}$, the cooler capacity will be:

$$\dot{Q}_{ch} = \frac{Q_{ch}}{\tau_d} = \frac{592.5}{16} = 37.0 \text{ kW} \quad (2)$$

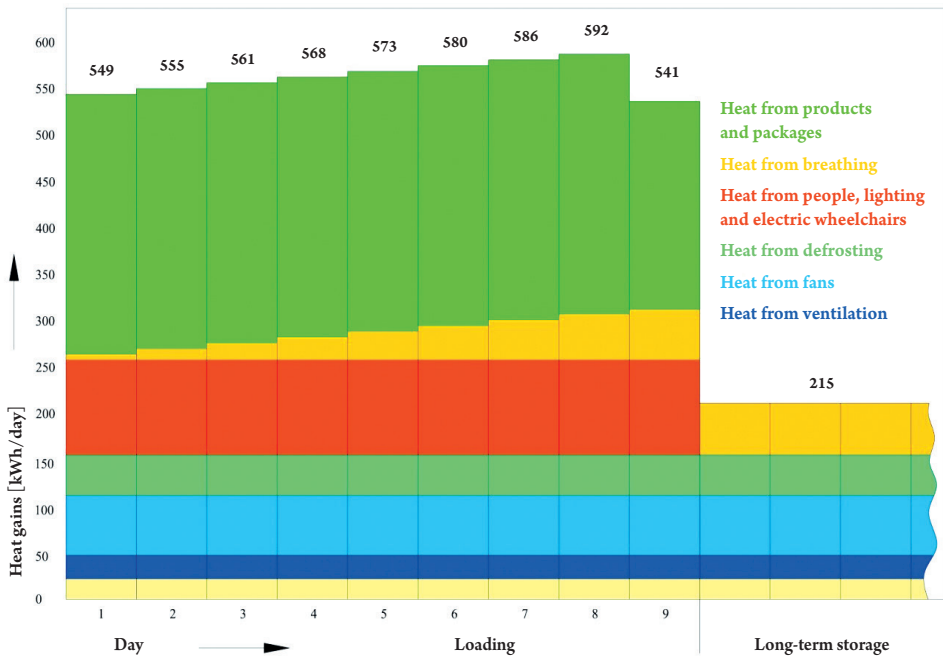


Fig. 3. Heat gains over a period of several days for fruit storage chamber

4. Calculation of cooling capacity of chiller

The cooling capacity of the chiller is determined by the sum of the heat flux values brought to the cold storage rooms that will be supported from the temperature controlled by the chiller. Table 2 shows the heat gains brought to the chambers during their loading and the heat gains after loading during long-term storage of fruit.

Table 2. Heat gains brought to the chambers during their loading and long-term storage

Number of cold room storage	Heat gains from loading [kWh / day]	Heat gains from loading [kW]	Heat gains from storage [kW]
K1	602.4	37.7	13.7
K2	592.5	37.0	13.5
K3	592.5	37.0	13.5
K4	592.5	37.0	13.5
K5	592.5	37.0	13.5
K6	592.5	37.0	13.5
K7	592.5	37.0	13.5
K8	592.5	37.0	13.5
K9	602.4	37.7	13.7
Sum:	5 352.3	334.4	121.9

Based on aggregate heat gains from the chambers of **5,352.3 [kWh/day]**, we can calculate the chiller's cooling capacity during fruit loading and the cooling demand for the long-term storage of fruit. Assuming a cooler working time of $\tau_d = 16\text{h}$, the cooling capacity of the chiller will be:

$$\dot{Q}_{ch} = \frac{Q_{ch}}{\tau_d} = \frac{5352.3}{16} = 334.5 \text{ kW} \quad (3)$$

Cooling demand for long-term storage:

$$\dot{Q}_{ch} = \frac{Q_{ch}}{\tau_d} = \frac{1950.4}{16} = 121.9 \text{ kW} \quad (4)$$

5. The cooling system and its technical concept

When the cooling capacity of the system is determined, system components will be designed. Firstly, the chiller together with compressors and oil drainage systems are selected. Secondly, the refrigeration system is selected – this is comprised of the condenser, the air cooler, the expansion valve and the ducts [4]. One of the most commonly used cooling systems is a direct evaporating refrigeration system with an air-cooled condenser. This type of system has a simple design but requires additional energy to drive the cooling fans and the condenser. In an air system, the heat received from the facility is transferred to the atmospheric air by means of a refrigerant. The refrigerant exchanges heat between the lower and the upper heat sources via continuous circulation within a closed system. Refrigerant

R404A was used in the analysed system. The choice of evaporation temperature primarily depends on the type of cooled goods and the required temperature in the chamber. The rule is that the smaller the temperature difference Δt , the greater the chiller's heat transfer surface and greater relative humidity. When selecting the evaporation temperature for our system, we must remember that high relative humidity of the air is required for optimum storage of fruit, so that the fruit does not dry out, this means that Δt must be small. The evaporation temperature can be calculated using the following equation:

$$t_o = t_R - \Delta t \text{ } ^\circ\text{C} \quad (5)$$

where:

t_R – temperature in chamber [$^\circ\text{C}$];

Δt – assumed temperature difference [$^\circ\text{C}$].

When the temperature in the chamber is 2°C , the evaporation temperature should be -3°C .

The designed chiller will have an air cooled refrigerant condenser. For this type of condenser, the condensing temperature can be calculated from the equation:

$$t_k = t_{amb} + \Delta t_1 \text{ } ^\circ\text{C} \quad (6)$$

where:

t_{amb} – ambient temperature [$^\circ\text{C}$];

Δt_1 – assumed temperature difference [$^\circ\text{C}$].

When the ambient temperature is 30°C , the condensing temperature should be 45°C .

6. Multi-compressor chiller

A variable-output compressor unit was used in the designed refrigeration system. The simplest way to regulate chiller capacity is to use more compressors. Each compressor has a specific capacity and the total capacity of all compressors is the total capacity of the chiller. Switching on and off individual compressors causes changes to the current capacity of the chiller [3]. The more degrees of control we can achieve, the more accurately the chiller's capacity will match the temporary cooling demand. This results in more efficient system operation and reduced energy consumption as well as reduced failure rate and greater reliability of the whole refrigeration system. Ten COPELAND refrigeration compressors of type ZB 114KCE-TFD were used to obtain the required degree of chiller control. Each compressor will cover 10% of the chiller's capacity. An air cooled condenser with axial fans and vertical air flow from GEA KUBA type CAV N08-2x3A was also selected, as well as nine direct evaporating air coolers with electric defrosting from GEA KUBA type SGBE 56-F84A. Air coolers must be installed so that the primary air stream is not directly aimed at the cooled goods. It should be directed along the ceiling above the goods – this causes a low-velocity suction of a secondary air jet and then slowly encircling the goods.

7. Refrigerants used in the system

One way to reduce the power consumption is to use a suitable refrigerant. The above system uses R404A. This product is used mainly in industrial refrigeration facilities such as cold stores and freezers, it is also non-combustible and non-explosive. The global warming potential of R404A is $GWP=3,300$, while the ozone depletion potential $ODP=0$. A new refrigerant R407F has recently appeared on the market. It can be used to achieve significant improvements in energy efficiency in systems compared to other refrigerants. An additional advantage of this refrigerant is its GWP of 1,824, which is significantly lower than the commonly used R404A. The use of R407F will make it possible to avoid additional operating charges. In the designed systems, calculations were made by replacing refrigerant R404A with refrigerant R407F. By analysing the results of the comparison of refrigerants, we can see that the use of refrigerant R407F instead of R404a will increase the EER cooling capacity from 3.24 to 4.21 (increase by 30%). This means that from one kW of electrical power, we will get 4.21 kW of cooling power, which will result in a significant reduction in demand for electricity.

The use of R407F will also reduce the refrigerant stream, which will enable a reduction in the diameter of the refrigerant piping. Reducing the refrigerant stream will also reduce the condenser's heat output. This will reduce the investment cost of the cooling system by using smaller piping and smaller condenser heat exchanger surfaces. R407F is a new refrigerant, consequently, it costs more than the standard R404A; however, cost of buying the refrigerant is incurred only once and the benefits in the form of reduced demand for electricity will be observable throughout the years of use of the refrigeration system – this will result in the rapid return on the invested refrigerant.

8. Conclusions

Designed for use with fruit storage chambers, the refrigeration system with an energy-saving variable-output chiller will be characterised by a high level of system reliability and low power requirements. The use of a multi-compressor chiller will provide 10 degrees of control, allowing the chiller to be adjusted to the temporary load of the refrigeration system. In order to achieve 100% of the cooling capacity, ten compressors will be in use, and when the chambers are loaded and the storage period starts, the output of the chiller will drop to 37%, which will require only four compressors to be operating. If the chiller is running on several compressors, they will work in rotation so that each compressor's lifespan is similar. Using multiple compressors increases the reliability of the chiller and the entire refrigeration system. Failure of one compressor will reduce the cooling capacity by only 10%. There are also benefits when replacing one small compressor. The starting current for a small compressor is low – this results in a significant reduction in compressor drive energy. A properly designed and selected air cooler with a low temperature difference does not cause excessive air drying – this reduces transpiration and water loss and increases fruit quality. This allows for an increase in the evaporation temperature, which results in an increase in the efficiency of the



chiller and a reduction in the energy consumption of the compressor drive. Using a wider lamella spacing will result in less frosting of the air cooler exchangers, which will reduce both the number of defrost cycles and the demand for electric power supplied to the cooler's heaters. The cost of such a cooler is greater than for a standard cooler; however, it is a one-off cost that is incurred during installation, and the loss of weight and quality of the fruit and the excessive energy required to defrost the exchanger will be visible throughout the years of use. Another element that reduces the cost of electricity consumption is the expansion valve in the regulation system. Using it will enable a reduction in the temperature difference in the air cooler. This will result in a higher evaporation temperature than is obtained with conventional expansion valves. Higher evaporation temperature will result in lower compressor power consumption. The use of the new R407F refrigerant contributes to a 30% increase in the EER cooling capacity. This will also allow us to get even greater cooling power from 1 kW of electrical power. The use of R407F will improve the energy efficiency of the refrigeration system and will reduce its operating costs.

References

- [1] Czapp M., Harun H., *Bilans cieplny pomieszczeń chłodni*, WSL, Koszalin 1995.
- [2] Gaziński B., *Technika chłodnicza dla praktyków- przechowalnictwo żywności*, Systherm, Poznań 2013.
- [3] Nowacki B., *Regulacja wydajności układów sprężarkowych*, Chłodnictwo i Klimatyzacja, 11/2011.
- [4] Recknagel, Sprenger, Schramek, *Kompendium Ogrzewnictwa i Klimatyzacji*, OMNI Scala, Wrocław 2008.
- [5] Ulrich H., *Technika Chłodnicza - poradnik*, Tom 2, IPPU Masta, Gdańsk 1999.
- [6] Zalewski W., *Systemy i urządzenia chłodnicze*, Politechnika Krakowska, Kraków 2014.

Augustyn Lorenc (alorenc@pk.edu.pl)

Logistics Systems Division, Cracow University of Technology

Małgorzata Kuźnar

Reliability and Technical Exploitation Division, Cracow University of Technology

THE COST OF RISK EVALUATION IN INTERMODAL TRANSPORT

– VARIANT ANALYSIS

ANALIZA KOSZTÓW I RYZYKA W TRANSPORCIE INTERMDALNYM – ANALIZA WARIANTOWA

Abstract

In the article, the impact of monitoring systems on the possibility of lowering costs and improving the flow of information in the transport chain has been analysed. For the analysed transport chain, three variants have been taken into consideration: variant I – without using systems for monitoring location, variant II – container transportation with using systems for monitoring location, variant III – container transportation with using location monitoring and cargo parameters. The conducted tests made it possible to confirm the benefits of implementing this type of system, especially for high-value goods. In the article, a simulation for 50,000 cases of container transportation was done. The simulation and cost analysis has shown that it is possible to reduce the cost of risk significantly.

Keywords: intermodal transport; containers; positioning systems; localization monitoring; cargo parameters monitoring; risk and costs performance analysis

Streszczenie

W niniejszym artykule przebadano wpływ systemów monitorowania ładunków na możliwość zmniejszenia kosztów i poprawy przepływu informacji w łańcuchu transportowym. Dla analizowanego łańcucha transportowego rozważono trzy warianty: wariant I – bez systemu lokalizacji, wariant II – z systemem do monitorowania lokalizacji, wariant III – z systemem do monitorowania lokalizacji i parametrów ładunku. Wykonane badania pozwoliły stwierdzić korzyści wynikające z zastosowania tego typu systemów, zwłaszcza dla produktów o wysokiej wartości. W artykule wykorzystano dane dla 50 000 przypadków przewozu ładunków. Przeprowadzone symulacje i analiza kosztów pozwoliły potwierdzić, że możliwe jest w dużym stopniu zmniejszenie kosztów ryzyka.

Słowa kluczowe: transport intermodalny; kontenery; systemy monitoringu, lokalizacja, monitorowanie parametrów ładunku, analiza ryzyka i kosztów

1. Introduction

Multimodal transport involves the carriage of goods using at least two different modes of transport. Intermodal transport is the transport of goods in one and the same transport unit using successively several modes of transport without reloading the goods themselves when changing the mode of transport. An intermodal transport unit may be, for example, a container or a swap body. Intermodal transport is therefore a particular type of multimodal transport.

The risk analysis for each element of the supply chain in an intermodal transport, including the identification of the occurring risks, is extremely important for the functioning of the supply chain [1]. The implementation of effective risk management in related chain links has a significant impact on the security of the entire flow process. Both in the transportation and in other links of the chain it is necessary to identify susceptible to fault places and recognise current and potential risks [2, 3]. The appropriate level of reliability of the supply chain is exposed to various types of risk due to the presence of many types of risks associated with changes in the internal and external environment [4]. Figure 1 presents the global cargo theft risk in 2014.

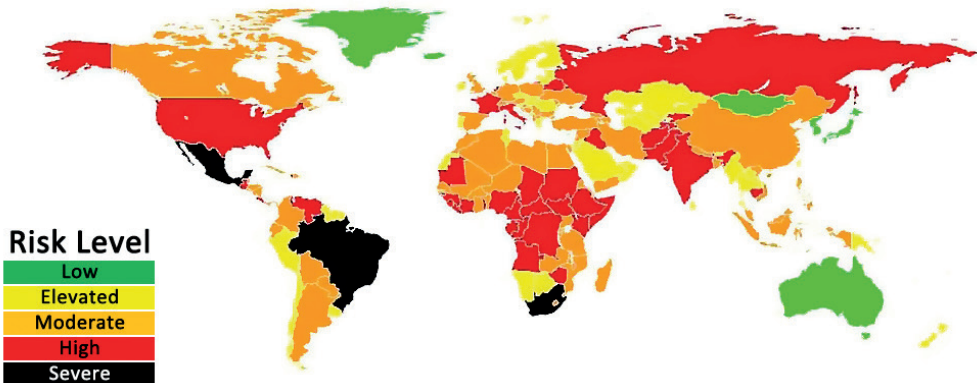


Fig. 1. Global cargo theft risk in 2014 [5]

In the context of the security risk of the transported cargo, we can distinguish the following threats and adverse events [6]:

- ▶ loss of cargo,
- ▶ reduce the quality of the cargo,
- ▶ absorption of water by the load,
- ▶ contamination of the load,
- ▶ changes in weather conditions,
- ▶ fluctuations in temperature,
- ▶ accident or incident during the transportation,
- ▶ maritime piracy,
- ▶ failure to comply with procedures for cargo handling,
- ▶ environment pollution,
- ▶ fire, explosion.

Among the container cargo damage, can be distinguish physical and temperature related harms. Bad stowage is one of the contributory causes of cargo damage. Among the most common container cargo claims in the maritime transport 25% applies to physical damage, 14% of the claims are temperature related, 11% concern containers overboard lost, 9% theft and 8% shortage of the load [7]. Figure 2 shows percentage share of large cargo Claims broken down by types of damage.

The causes of container cargo damage can be different. For example, we can mention lack of export packaging, inadequate ventilation, wrong choice of container, poor condition of container, lack of clear carriage instructions, ineffective internal cleaning, contaminated floors (taint), wrong temperature settings, condensation, overloading, poor distribution of cargo weight, wrong air flow settings, organised crime, fragile cargoes stored in areas of high motion, damaged securing equipment, poor monitoring of temperatures, wrong use of temperature controls [8].

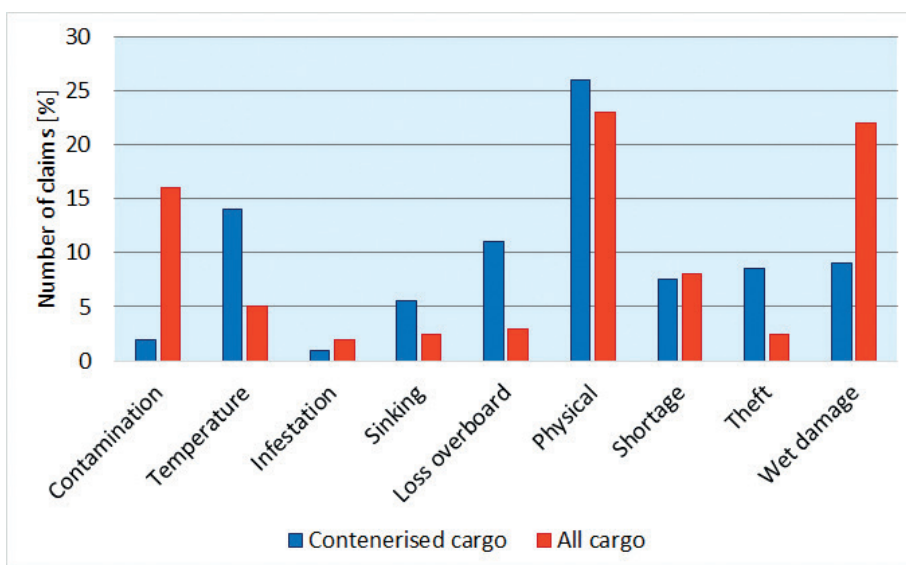


Fig. 2. Large cargo claims – type of damage

2. Location Systems - solutions overview

Intermodal Freight Transport is carriage of cargo using more than one type of transport using only one unit of load on the whole route of transport, without reloading the goods themselves when changing the mode of transport. Such kind of transport is often characterised by a high value of transported goods, mass character and large transport distances, moreover, often require a change in the carrier and even the mode of transportation. Therefore, the owners of intermodal freight and rail operators want to have information not only about its exact location but also the conditions of transport [9]. Most of the technological solutions such as GPS modules are universal and adapted to the requirements of the customer, regardless of whether it comes to industrial tools, dangerous chemicals, expensive drugs, exotic or high-

quality food products. Most commercially available units can monitor whether the specified temperature range is respected, measure the strength of shock or register moisture inside the load unit, record exposure to solar and electromagnetic radiation.

More advanced modules for monitoring parameters of containers also allow the cargo protection i.e., if the container doors are open, check if inside of the container there was no movement, lights or other deviating from standard events. In short, they can monitor everything that can have an impact on transported cargo. Available modules allow for current [10, 11]:

- ▶ monitoring the cargo location,
- ▶ recording of transport conditions,
- ▶ guarding the permitted range of conditions,
- ▶ temperature, humidity and shock measurement,
- ▶ container movement reporting,
- ▶ cargo protection by detecting unwanted actions of third parties,
- ▶ history of the cargo movement.

The units used in intermodal transport are designed specifically for monitoring containers. They are usually mounted on the outer wall of the container, have sensors on the door or the inside of the container and are sensitive to atmospheric conditions.

The installed device receives location data from the GPS satellites. The information is then sent via data transmission by mobile network to the server, which is then elaborated. When a vehicle is located out of range of a mobile network, e.g. in the middle of the ocean during a storm or other reasons, history of the position does not lose. If the GPS module cannot send position data, save it in its internal position and send it when possible.

The modules used in the monitoring of cargo transportation are highly prevalent on the market and are usually in the form of independent modules. According to destination and monitoring capabilities can be distinguished following systems:

- ▶ **Dedicated systems for port and container terminals**

This is usually an online application dedicated to seaports and their clients, offering a wide range of functionality (TOS type systems - Terminal Operating System). Position location of containers in the harbour is one of the basic system options. It is not required to mount a GPS module to container because the application is based on a virtual model, i.e. every container which is inserted on storage square has assigned a specific location. Linking the container number to a specific place of storage allows the system to visualise the stored units in the terminal. This solution provides high accuracy, but in the case of reading errors make it difficult to locate the container. Apart from indicating the location on the square storage system allows you to show the status of the container, for example placing on the ship, unloading, storage, loading wagon/trailer, etc [12].

- ▶ **Systems to monitor the approximate position**

Such systems are based on information exchange between users. Do not use the GPS modules or GPRS data transmission. Therefore, they do not allow the precise position of the load, but also did not require the installation of additional equipment for intermodal unit [13].

► **Systems to monitor only the locations of units**

The location device with a GPS receiver installed in the vehicle collects data on its current position. Then, using GSM/GPRS network, data is sent to the monitoring server, for which the user receives access. The application allows electronic tracking of cargo by monitoring its location and safety. By using SaaS solution - Software as a Service and capability to support almost every tagging technology (GPS, GPRS, RFID and others), it is possible to track and monitor cargo units almost all over the world [14].

► **Systems to monitor the specific parameters of the load**

Such a system makes it possible not only to identify the exact location of the cargo unit but also allows you to specify the specific parameters of the load. With the ability to define custom POIs - Points of Interests, and assign to them the maximum time within which the cargo should leave the defined area, it is possible to automatically generate notifications. This solution makes it possible to supervise costs resulting from too long waiting time for transshipment in a terminal or warehouse. The system offers automatic notification of temperature changes in relation to the established range. The system also includes the measurement of various parameters of the load, for example, temperature, pressure, leaks and other safety-related parameters [11].

► **Systems to monitor the location and parameters of the mode of transportation**

The system helps supervise individual units by using GPS locators. It is also possible to apply additional sensors that increase the number of monitored parameters. Data handling from the locators and using a special set of statistical tools may be held from a web browser. Sometimes, such a system does not provide direct possibilities for monitoring the position of the intermodal unit. Thanks to the information concerning the means of transportation and cargoes situated on it, it is possible to link the means of transportation location to the location of the container. Additional services and functionality of the system may permit inter alia: the registration of constant parameters (registration number of the vehicle, driver and vehicle data), registering exploitation parameters (vehicle mileage, pressure, digital signals), information about the faults and device status information, recording service information for the diagnosis [15].

3. Transport chain – considered variants

To be able fully to assess the benefits resulting from the application of location and cargos parameters monitoring systems, the transport chain has been established. As primary elements of that transport chain could be counted: road transport, sea transport, transshipment and storage in container terminals in ports. For each transportation process risk, which could appear in a process, has been identified. Transport chain with potential risk is shown in Fig. 3.

For presented transport chain, tree variants have been taken into consideration:

- Variant I, container transportation without using systems for monitoring location,
- Variant II, container transportation with using systems for monitoring location,



- Variant III, container transportation with using systems for location monitoring and cargo parameters, like temperature, humidity, shock, opening the door, the light inside.

What is more, the costs for each transportation processes and the cost of undesirable incidents were established with a currency conversion at the rate of 4.407 for Euro/PLN 1.118 and USD/Euro (the state from 2017-03-04).

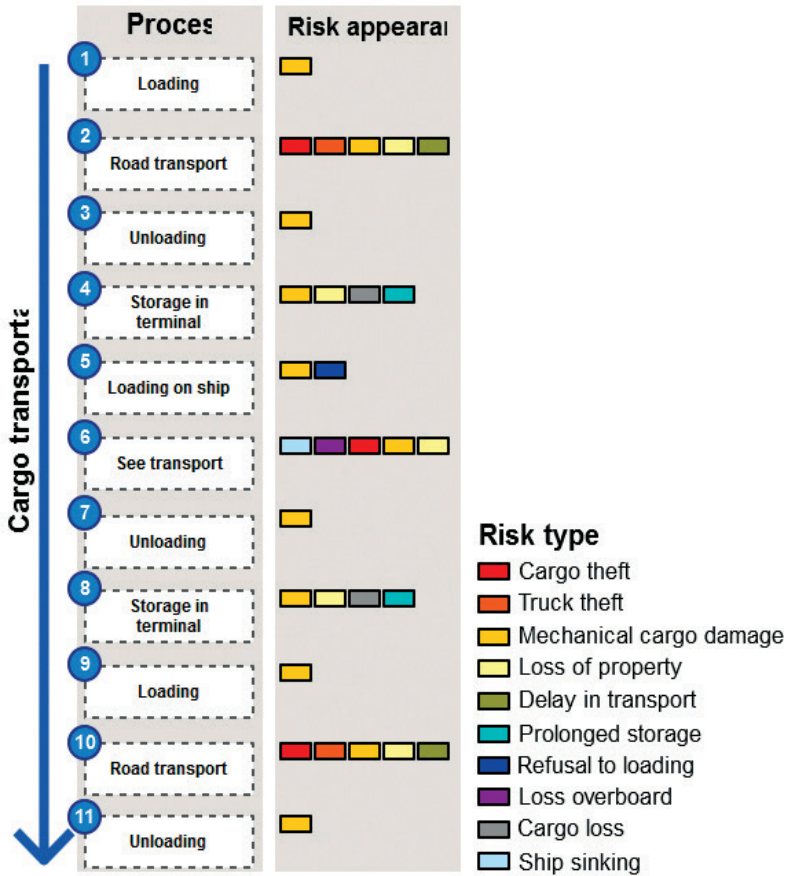


Fig. 3. Risk appearance in transport chain

Both the costs and the responsibility of loading during loading in the shipped company are on side of the shipper. Because of that, the cost of cargo damage was not taken into consideration in that analysis. If, at that stage, cargo damage would be stated, the manufacturer shall replace it. But in case of cargo damage in such a way that it is not possible to notice it during loading, the damaged cargo will be supplied to the customer. Those products will be replaced by new one in the future, so it will generate extra costs.

For cargo transportation by the road, the cost equal to 1.10 Euro/km was established. That rate base on standard transport costs in Europe.

Transshipment costs for container terminal were established by average costs in that type of services. So, it is 42.50 Euro by loading/unloading [22]. Cost of storage in the terminal is 4 Euro per day, and it complies with polish rail company PKP CARGO CONNECT [23].

Cost related to sea transport was established by calculation for transport from Shanghai (China) to Gdańsk (Poland) [22].

The probability of risk appearance was estimated by real statistical data. In 2014, there were 2 007 926 see transportation processes and 1 069 202 full load road transportation processes in Europe [22]. Base on statistic information of transportation and risk estimation and information about undesirable incidents in see transport [7, 17–19] and road transport [20, 21] the probability of risk appearance was estimated. In Table 1 are shown probability of risk appearance for separated processes.

Table 1. Probability of risk appearance

Process	Risk	Probability [%]
1. Loading	Mechanical cargo damage	0.00623
2. Road transport	Cargo theft	0.09353
	Truck theft	0.00954
	Mechanical cargo damage	0.00830
	Loss of property	0.01539
	Delay in transport	0.05000
3. Unloading	Mechanical cargo damage	0.00623
4. Storage in terminal	Mechanical cargo damage	0.00623
	Loss of property	0.01494
	Cargo loss	0.00050
	Prolonged storage	1.50000
5. Loading on ship	Refusal to loading	0.05000
	Mechanical cargo damage	0.00623
6. See transport	Ship sinking	0.00498
	Loss overboard	0.00996
	Cargo theft	0.00770
	Mechanical cargo damage	0.02490
	Loss of property	0.03033

Process	Risk	Probability [%]
7. Unloading	Mechanical cargo damage	0.00623
8. Storage in terminal	Mechanical cargo damage	0.00623
	Loss of property	0.01494
	Cargo loss	0.00050
	Prolonged storage	1.50000
9. Loading	Mechanical cargo damage	0.00623
10. Road transport	Cargo theft	0.09353
	Truck theft	0.00954
	Mechanical cargo damage	0.00830
	Loss of property	0.01539
	Delay in transport	0.05000
11. Unloading	Mechanical cargo damage	0.00623

4. Description of Variant I

In variant I, the containers are transported without using systems for location monitoring and cargo parameters. So, in this case, there is no possibility to determine in real time where an exactly container is. It is possible only in an approximate way (as the stage in transport chain) by the identification number of a container. Taking into account the above, there is no possibility to respond to delays in transportation, which could result in refusal to loading the cargo to ship, prolonged storage time in the terminal, the problems with the synchronisation with other means of transport, etc. Likewise, in the case of truck/trailer or cargo theft, ship sinking or container loss overboard, it is not possible to quickly determine what happened with the cargo. That incident could cause additional costs related i.e. to contractual penalties.

5. Description of Variant II

In variant II, the containers are transported using systems for location monitoring but without additional functionality. In this case, it is possible to determine where the exactly container is. Cause of that, costs of undesirable incidents like theft or container loss, could be reduced. Moreover, it is possible to synchronise transshipment process, reduce the time of

storage in terminal – in the case when vessel arrives faster than expected, and reduce problems with determining loading and unloading terms.

In this variant, it is not possible to identify the container state and other parameters of the cargo. What is more, it is not possible to identify partial cargo theft from the container - these cases represent the majority [9]. Noticing if the cargo is complete is possible practically only during transshipment in the presence of a man - checking the condition of the seals, door and side walls of the container.

Location monitoring systems allow to partial reduction of costs related especially with contractual penalties and synchronisation between transportation processes [23].

6. Description of Variant III

In variant III, the containers are transported using systems for location monitoring and additional functionality – monitoring cargo parameters. For that with basic module were integrated extra sensors, for checking if container doors are open, check if the inside of the container there was no movement, lights or other deviating from standard events and sensors for monitoring cargo conditions like temperature, humidity or shock. Because of that, in the case of transshipment, if the sudden increase or decrease of monitored parameters out of the permissible range was noticed, it could be a proof to prove the guilt of a subcontractor (carrier or shipper). So, therefore, the load is in the responsibility of the subcontractor; he bears the costs of damaged cargo. Such a solution ensures costs reduction related not only to delays in transport and loss of the cargos, but also in case of cargo theft, damaged or loss of property.

7. Assumptions for simulation

For the simulations, the above-presented assumptions were made:

- ▶ Route: Shanghai (China) to Gdańsk (Poland)
- ▶ Transport modes: road, see
- ▶ Costs:
 - ▷ The cost of loading/unloading is 42.50 Euro for each,
 - ▷ The cost of storage in terminal is 4 Euro per day,
 - ▷ The cost of see transport for 20' container is 727 Euro,
 - ▷ The cost of preparing necessary documents and bill of landing is 66 Euro,
 - ▷ The cost of customs procedure is 44 Euro,
 - ▷ The cost of road transport is 1.10 Euro/km.
- ▶ Contractual penalties:
 - ▷ delivery of damaged / incomplete cargo: 20% of charge,
 - ▷ delivery of damaged / incomplete cargo in the case of prior notify the customer: 10% of the load,
 - ▷ delay: 1% for each day of delay.

8. Evaluation of risk for variants

A data generator was developed for the simulation. In the first step, an identification of potential risks in a transport chain was made, which is shown in Fig. 3. Then, on the basis of the transport chain stage and the risk, it was written in the form of mathematical equations, which allow to draw data from the range corresponding to the probability of occurrence of the risk. Data were generated using normal distribution assuming the probabilities presented in Table 1. The excel function returned a value of 1 in the case of an adverse event appearance associated with a particular type of risk and a value of 0 if it did not occur. Then, based on the type of risk that occurred, a percentage share of the damaged load was generated in relation to the entire carried cargo. If such risks occurred, for example, in the case of a sinking ship, the percentage share of the damaged cargo was taken as 100%.

Costs arising from the occurrence of the risk were calculated on the basis of the percentage share of damaged cargo and other costs arising from the adverse event. The higher the stage of transport, the higher the cost due to the occurrence of contractual penalties or the need to substitute a new cargo for undelivered or damaged cargo.

In case of the simulation of 50 000 transports cases, in 1713 of them risky incidents appear. For all case was established the value of cargo equal to 7142 Euro. The result of the simulation of risk costs is graphically shown in Fig. 4.

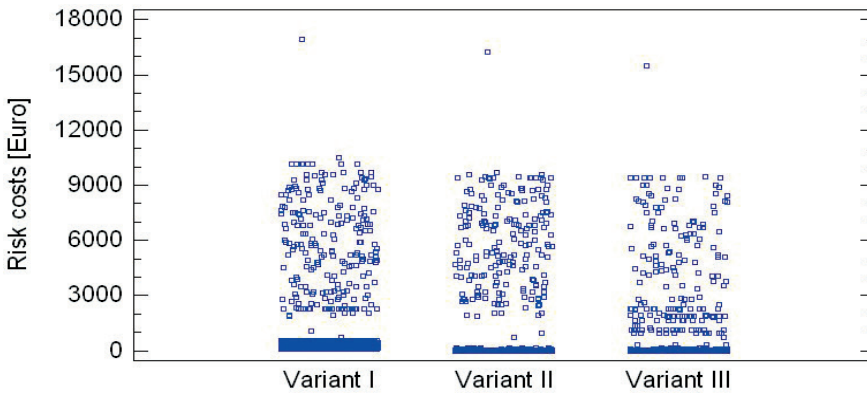


Fig. 4. Risk costs, depending on the variants

On the basis of Fig. 4, it can be stated that Variant I and Variant II produce the results with the greatest range and standard deviation. This is also noticeable, based on the statistical characteristics presented in Table II. As could be expected, with the usage of systems to monitor the location and parameters of the products, the reduction of the cost of risk can be noted. This is due to the fact that the freight forwarder has the option to react to the situation, which causes that the problem is not pulled in time and therefore do not bear the additional costs arising from it. In Fig. 4 it can be seen that variant III also reduced the number of cases of risk associated with variant II. The costs of the risks themselves have not been fully offset but significantly reduced – as can be seen in the increased number of cost cases up to € 3,000. Values that exceed the average are marked with colour.

Table 2. Descriptive statistics of the performed analyses [euro]

	Average	Standard deviation	Median	Min	Max	Range
Variant I	1021.4	2081.6	377.1	71.4	16927.9	16856.4
Variant II	704.8	2063.5	8.0	0.0	16213.6	16213.6
Variant III	483.5	1618.9	8.0	0.0	15499.3	15499.3
Total	736.6	1945.4	24.0	0.0	16927.9	16927.9

For the conducted analyses, analysis of variance (ANOVA) was performed in order to test the significance of differences between average values. The F-ratio, which in this case equals 33.5, is a ratio of the between-group estimate to the within-group estimate. Since the P-value of the F-test is less than 0.05, there is a statistically significant difference between the means of the 3 variables at the 95.0% confidence level.

In order to determine which groups differ statistically from each other, multiple comparisons were performed, the so-called post hoc test. To perform post-hoc tests, Scheffe, T Tukey (HSD), Fisher (LSD), Bonferroni, Newman-Keuls and Duncan tests are applied. In the analysed case, the Scheffe test, which is considered one of the most conservative post hoc tests, has been used. The results of the test are shown in Table III and in Fig. 5.

Table 3. Multiple comparison using Scheffe test at 95% of confidence interval [euro]

	Mean	Std. error	Lower limit	Upper limit	Homogeneous Groups
Variant III	483.5	46.7	940.5	1102.2	X
Variant II	704.8	46.7	623.9	785.6	X
Variant I	1021.4	46.7	402.7	564.4	X

This table shows the mean for each column of data. It also shows the standard error of each mean, which is a measure of its sampling variability. The standard error is formed by dividing the pooled standard deviation by the square root of the number of observations at each level. The table also displays an interval around each mean. The intervals currently displayed are based on Scheffe's multiple comparison procedures. They are constructed in such a way that if all the means are the same, all the intervals will overlap at least 95.0% of the time.

The last part of the analysis was to make a comparison of medians using the Kruskal-Wallis test. Which tests the null hypothesis that the medians within each of the 3 columns are the same. The data from all the columns are first combined and ranked from smallest to largest. The average rank is then computed for the data in each column. The result of the test was 2134.1 and the P-value equalled 0.00, which demonstrates that the groups differ significantly from each other. The results coincide with the preliminary assessment of the graph shown in Fig. 4. To determine which medians are significantly different from which others, the Box-and-Whisker plot was prepared (Fig. 6).

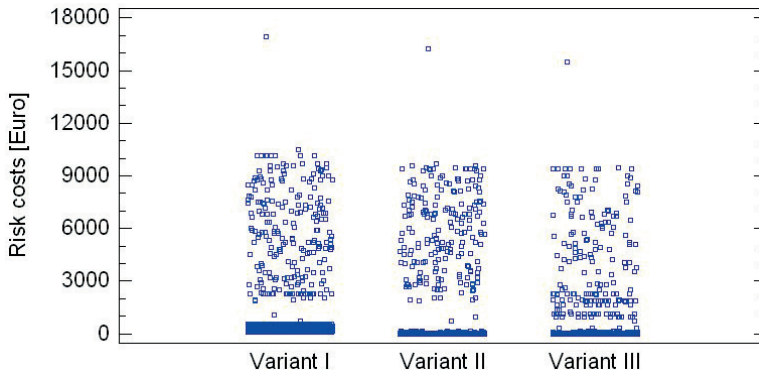


Fig. 5. Result of the comparison of average with Scheffe interval with 95% confidence interval

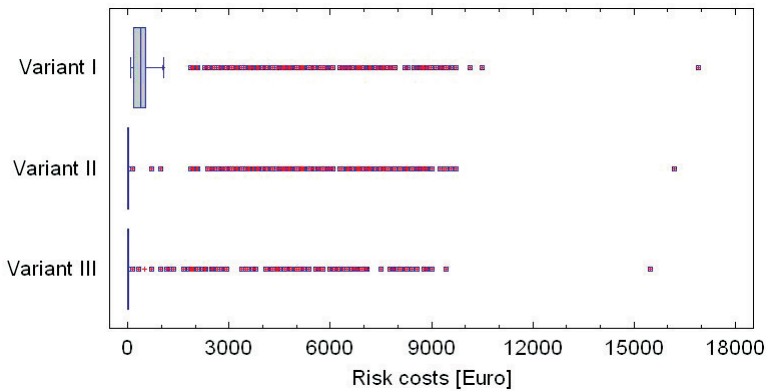


Fig. 6. Box-and-Whisker Plot for variants

Making a comparison of average and median values, it was found that the best results were obtained for Variant III where devices for monitoring localisation and cargo parameters was used. This method allows obtaining lower costs than in the case of container transport without any monitoring by an average of 52.7%. What is more, using the system for only monitoring localisation helps to reduce costs by 31.0%.

Figure 7 shows the cost of an adverse event, which was come out for the simulation. As could be noticed, the higher costs are in road transport in last transportation process. It is noticeable that for stage 10 the risk costs are lower than for stage 2. Both stages have the same type of carriage, while the cost difference results from higher contractual penalties. The disruptions in later transport are spotted the greater cause financial losses. This is due to lack of supply time, what in the occurrence of an adverse event is the greatest issue.

Costs on the stage of carriage by sea or first road transport and storage in the terminal are similar. Using systems for position monitoring helps to a slight reduction of risks cost in each process in transport chain, and can considerably reduce the cost of storage in containers terminal. Systems, which permit to monitor additional parameters, help to noticeably reduce costs in each stage of the transportation chain.

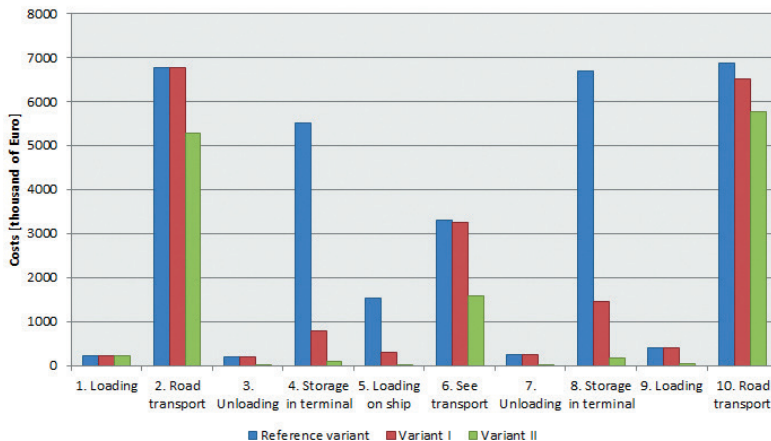


Fig. 7. Costs of adverse event by processes in presented supply chain

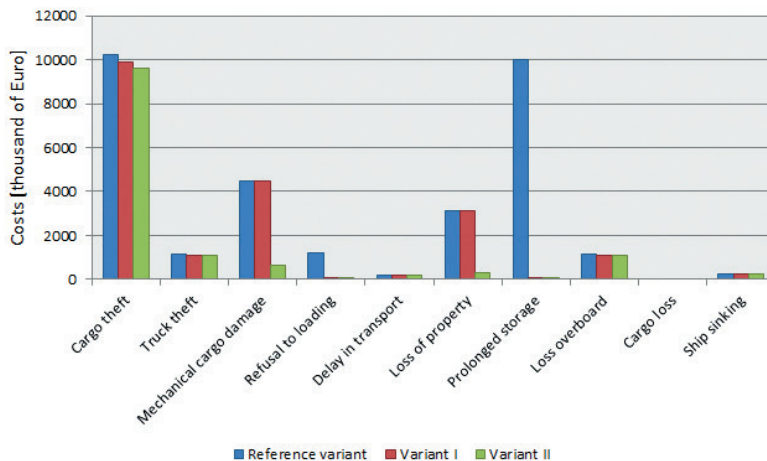


Fig. 8. Costs of adverse event by risk type in presented supply chain

Figure 8 shows the cost of an adverse event by risk type. For the performed simulations, the biggest costs were for theft cases. Per 50000 transport cargo theft occurred 102 times and truck/ trailer theft 11 times. Cost of that incident could be reduced by additional cargo insurance, which is not obligatory. The problem could appear for high-risk countries. For that country, the insurance rate is significantly higher than normal. What is more for transport through sea territory of that country (i.e. Somalia or Republic of South Africa) insurer could refuse cargos insurance because of frequent cases of piracy. Cargo insurance could cover the costs associated with its mechanical damage or loss of cargo property. Increased costs connected with the storage time in terminal could be decreased if cargo forwarder is informed about a position of container [24]. It allows to synchronise each process in transport chain and reduce the time of storage between cargo movement [25].

9. Conclusion

Conducted tests made it possible to confirm the benefits of implementing a system for positioning and monitoring cargo parameters, especially for high-value goods. Devices for monitoring location and cargo parameter allows obtaining lower costs than in the case of container transport without any monitoring systems by an average of 52.7%. What is more, using system for monitoring only location can be helpful to reduce costs by 31.0%. These results showed that using those systems could bring great benefits. The simulations show that the riskiest stage is a road transport - for theft and mechanical damage cases. On the other hand, sea transport is the riskiest in terms of cargo properties lost or lost cargo. For this type of adverse event, it is advisable to carry out further research and develop prognostic models. According to statistics, consumers bear the cost of 20% of the value of products - this is due to the disruption in transport chain. Improving the efficiency of transport operations and reducing the risk can thus reduce the final costs of goods. What is more besides reduction of costs those could increase customer service levels.

In the next stage of research, it is planned to develop some intelligent forecasting methods to predict risk. It is planned to use artificial neural networks and genetic algorithms for this purpose. If developed methods will allow for effective risk forecasting, it is planned to implement them in own type of Transport Management System (TMS).

References

- [1] Gajewska T., Szkoda M., *Ocena konkurencyjności transportu intermodalnego*, Logistyka (3), 2015.
- [2] Szkoda M., *Assessment of Reliability, Availability and Maintainability of Rail Gauge Change Systems*, Eksploatacja i Niezawodność – Maintenance and Reliability 16(3), 2014.
- [3] Lorenc A., Szkoda M., *Customer Logistic Service in the Automotive Industry with the Use of the SAP ERP System*, Proceedings of: 2015 4th IEEE International Conference on Advanced logistics and Transport (ICALT), 20–22 May 2015.
- [4] Hall P.V., *Institutional Challenges to Intermodal Transport and Logistics*, Journal Of Transport Geography 47, 2015.
- [5] http://www.iumi.com/images/gillian/Spring2015/IUMI_spring_2015_Cargo_presentation.pdf (access: 8.02.2016).
- [6] Ciesielski M., *Instrumenty zarządzania łańcuchami dostaw*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2009.
- [7] *Container matters*, UK P&I CLUB, <http://www.ukpandi.com> (access: 28.02.2016).
- [8] Dotoli M., Epicoco N., Seatzu C., *An improved technique for train load planning at intermodal rail-road terminals*, Proceedings of 20th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Luxembourg 2015.

- [9] Lam J.S.L., Gu Y., *A market-oriented approach for intermodal network optimisation meeting cost, time and environmental requirements*, International Journal Of Production Economics 171 (2), 2016.
- [10] <http://www.mecomo.com/en> (access: 8.02.2016).
- [11] <http://www.intermodaltracker.eu> (access: 8.02.2016).
- [12] <http://www.navis.com> (access: 8.02.2016).
- [13] <https://www.cesar-online.com> (access: 8.02.2016).
- [14] <http://www.savi.com> (access: 8.02.2016).
- [15] <http://www.novacom-services.com> (access: 8.02.2016).
- [16] <http://ec.europa.eu/eurostat/data/database> (access: 8.02.2016).
- [17] <http://www.actuarialey.com/2014/03/30/how-many-ships-disappear-each-year> (access: 8.02.2016).
- [18] <http://www.statista.com/statistics/264024/number-of-merchant-ships-worldwide-by-type> (access: 8.02.2016).
- [19] *Carefully to Carry: Cargo damage – the causes*, UK P&I CLUB.
- [20] *Road transport: Prevent the risk of theft*, AXA Corporate Solutions, 2009.
- [21] *Crime in road freight transport*, OECD, France 2002.
- [22] <http://kontener.pl> (access: 8.02.2016).
- [23] *Oferta PKP CARGO CONNECT*, <http://www.tradetrans.eu> (access: 8.02.2016).
- [24] Clott C., Hartman B.C., *Supply chain integration, landside operations and port accessibility in metropolitan Chicago*, Journal of Transport Geography (51), 2016.
- [25] Stuart E., *Excellence in warehouse management – how to minimise costs and maximise value*. Chichester, John Wiley & Sons Ltd., 2005.



Karol Miądlicki (karol.miadlicki@zut.edu.pl,)

Mateusz Saków

West Pomeranian University of Technology

THE USE OF MACHINE VISION TO CONTROL THE BASIC FUNCTIONS OF A CNC MACHINE TOOL USING GESTURES

ZASTOSOWANIE SYSTEMU WIZYJNEGO DO STEROWANIA PODSTAWOWYMI FUNKCJAMI OBRABIARKI CNC ZA POMOCĄ GESTÓW

Abstract

This paper presents a concept of a vision system which can simplify the way in which some basic functions of CNC machines can be controlled. The proposed system enables the operator to control a machine tool using gestures. The developed solution is based on Microsoft Kinect for a Windows v2 sensor with a time-of-flight camera. A gesture recognition module was implemented in the VC 760 milling machine with an open control system (O.C.E.A.N.). To conduct tests of the proposed interface, a set of gestures used to control a CNC machine was developed. Furthermore, the concept, the structure of the system and the test results are discussed. In summary, the advantages and potential problems of the proposed control system and plans for future development are discussed.

Keywords: gesture control, Microsoft Kinect, time of flight, machine tool, machine vision

Streszczenie

W artykule przedstawiona została koncepcja systemu wizyjnego umożliwiającego kontrolowanie i programowanie obrabiarki CNC za pomocą gestów. Opracowane rozwiązanie ułatwia obsługę obrabiarki CNC poprzez rozpoznawanie gestów wykonywanych przez operatora. Do realizacji systemu wykorzystany został kontroler ruchu Microsoft Kinect for Windows v2. System rozpoznawania gestów zastosowano w otwartym systemie sterowania obrabiarki VC 760 (O.C.E.A.N.). W ramach badań opracowane zostały zestawy gestów pozwalających na sterowanie obrabiarką CNC. W artykule omówiono koncepcję i budowę systemu oraz wyniki przeprowadzonych testów. W podsumowaniu wskazano zalety oraz potencjalne problemy związane ze strukturą i zastosowaniem systemu, a także zarysowano plany jego dalszego rozwoju.

Słowa kluczowe: sterowanie gestami, Microsoft Kinect, obrabiarki CNC, widzenie maszynowe, systemy wizyjne

1. Introduction

1.1. Current systems of CNC machine control

One of the most sought-after skills on the job market is expertise related to controlling CNC machine tools and programming machining operations. In recent years, productivity and accuracy of machining has increased in the machine tool industry. Numerous CNC control algorithms have been developed and machine vision systems have been introduced to improve precision and to support the reference positioning of the workpiece. However, methods of communicating with the machine through the operator interface have not been developing at the same rate. Despite developments in the graphic display of CNC interfaces or the introduction of touch panels, the CNC control systems that are currently available on the market are not very intuitive. Manual interfaces still rely on complicated control panels and require knowledge of how to use control systems, Fig. 1 (Siemens, Heidenhain, Fanuc).



Fig. 1. Operator panel of the most popular CNC systems. From left Sinumerik, Heidenhain, Fanuc

As a consequence of the current nature of manual interfaces it is necessary to learn a set of procedures and key sequences, memorise the functions of diodes and switches, and what actions are triggered by what procedures. Since machines, CNC drivers and operator panels are often configured differently, this only adds to the current difficulties. As a result, the interface between the operator and the machine has become increasingly complex [1]. Only qualified operators can be employed to operate CNC machines – they must go through advanced training to improve their efficiency and minimise the risk of errors.

1.2. Smart control systems

With the increasing popularity of computers with huge processing capacity, the development of virtual reality technology and 3D imaging, the application of gestures and human senses to remotely control various machines (haptic feedback) [2, 3] seems inevitable [4]. This is evidenced by the current trends on the market of consumer and industrial electronics. While traditional control methods (remote control units, buttons, keys and

joysticks) seem to be disappearing – new control solutions allowing the use of gestures, voice [5, 6] and eyes [7] are more frequently used in game consoles, smartphones and TV sets [8]. The next step in the development of control methods is the full integration of a user with a device through ordinary conversation. This is already being tested on the latest smartphones and computers using voice assistants, for example, Cortana (Microsoft), Siri (Apple) and Google Now (Google Inc.). It is only a question of time before new systems can be implemented into industrial machines.

1.3. Gesture control systems

While attempts to modify currently used control systems are being made, other attempts are being made to design virtual environments to use them for remote control purposes in modern factories [9]. These are intended to provide easy and remote operator/machine interaction methods. Gesture control and virtual reality systems might be of interest to the industry. The proper selection of gestures used to give precise commands enables effective and intuitive interaction with machinery without having to be close to it or to learn how to operate a complicated control panel. Gestures and hand-motion control systems can be used for mobile robots, humanoid robots [9], flying robots [10], industrial manipulators [11], electro-hydraulic manipulators [12] and cranes [13–15]. Attempts have been made to use them in surgical robots [16] for controlling data displayed on a monitor without having to touch the keyboard or the mouse [17] as this would ensure maintaining a sterile environment within the operating theatre. In another attempt [18], a virtual environment was used to program a CNC machine using gestures.

Three dimensional (3D) gesture control systems may be implemented using one of the following methods: vision [19], sound [20], magnetic [21], mechanical [22]. The most commonly used vision method is stereovision. However, active methods based on laser scanning heads (RADAR, LIDAR) [23, 24] cameras depth (ToF) or light structure are increasingly being used in many projects. In active vision methods, it is particularly important to obtain information relating to depth. Besides recognising gestures, depth is used to analyse the environment around the machine or locate the position of objects. Despite its popularity

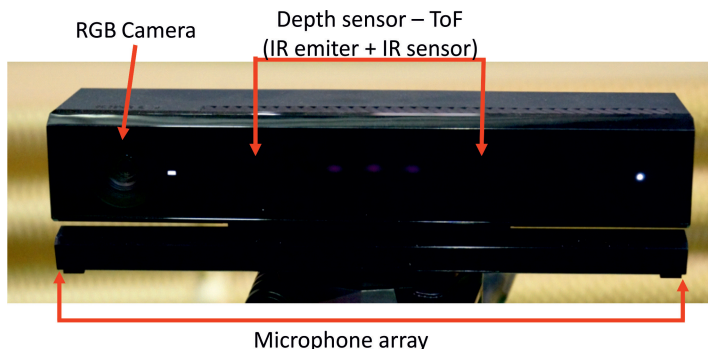


Fig. 2. Microsoft Kinect v2

and many advantages, optical methods for gesture recognition have several major drawbacks: they require many calculations (stereoscopy), they are expensive (time of flight) and sensitive to external factors (light structure). The effectiveness and reliability of the gesture control system, depends on the precision of the performed gesture and the processing of the gesture signal on each of four or five stages of operation of the system (in the case of tracking systems) [19]. The first important stage is how the user performs a gesture. Here, of particular importance is the accuracy, speed and environment in which the gesture is performed. The gesture is then captured by the sensor (for example, by using a camera, an accelerometer and a magnetic field sensor), which transforms it into a digital signal. The purpose of the following steps (algorithms) is to recognise and classify the received signal (image). Firstly, the image is subjected to segmentation and the feature extraction process.

The segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as super-pixels). The purpose of segmentation is to simplify and change the representation of an image into a form that is more meaningful and easier to analyse. Feature extraction is the process of defining a set of features, or image characteristics, which will most efficiently or meaningfully represent the information that is important for analysis and classification. *The* real-time tracking system are required to perform these processes and locate a tracking object in every frame of the video sequence. The most commonly used methods for segmentation and feature extraction are based on *the* skin colour model [25] and static [26] or dynamic [25] background subtraction. These are popular because of their simple implementation and applicability to real-time systems [27]. Some researchers use the infrared camera [28] to extract human I think you should insert another noun here such as 'forms', 'elements' or 'features' here from the frame based on body temperature, as well as more advanced methods based on filters [29].

The next task of the system is to classify the gesture performed by the operator. Therefore, features are used both before and after object segmentation to provide statistical parameters for classification by a recognition algorithm. A classification or recognition mechanism is a computational algorithm that takes the representation of an object and classifies it as some known class type. The most commonly used techniques for classification are the HMM (hidden Markov model) [30, 31] and other statistical methods, ANN (artificial network neuron) [32], methods based on state space [33] wavelet transform [32], matching the curves [34], and database methods [35]. Since the issue of recognition of gestures is currently the subject of intensive research, there are many review papers in which authors present their research in detail and compare algorithms for segmentation, extraction, tracking and classification systems used in gesture recognition.

1.4. Concept of the proposed control system

This paper presents the concept of a CNC control system extension for a milling machine based on gesture monitoring and recognition. Its concept is based on tracking operator's gestures using a Microsoft Kinect v2 motion controller (Fig. 2) with a built-in, time-of-flight camera. Data acquired by the system is processed by control algorithms and used for gesture

recognition. The operator's hand, used as the controller, provides a natural and intuitive control interface. The environments of Kinect Studio, Microsoft Visual Studio, Visual Gesture Builder, Kinect SDK & Developer Toolkit libraries and C++ programming language were implemented in the system.

2. Depth sensor

2.1. Microsoft Kinect v2

The Microsoft Kinect sensor was unveiled by Microsoft in November 2010 and it was dedicated for the Xbox 360 game console. This then novel solution enabled its users to interact with their console without having to use pads or other control systems which were common at that time. The system implemented a new interface using hand gestures, body movements and spoken commands. Two years later, Microsoft released Kinect SDK (software development kit) enabling the sensor to cooperate with a PC, and in November 2013, the company launched Kinect v2 together with the new Xbox console.

The sensor is sold with a software package which contains Kinect Studio, Visual Gesture Builder and Kinect SDK 2.0 libraries. The new version of the sensor is more accurate owing to a camera with higher resolution and a ToF camera. This is an important feature since the sensor is used in a variety of professional and semi-professional applications, including advanced medical diagnostics, 3D image analysis (in computer tomography), robotics and IT.

The Kinect sensor consists of six modules: a multi-array microphone, a time-of-flight camera, an infrared emitter, an angularity controller, a USB 3.0 connection lead, and a colour RGB camera. The Kinect sensor has a range of 0.5–4.5 meters. The lens has an angular field of view of 60° vertically and 57° horizontally. The system features four microphones and a 24-bit analogue-to-digital converter which are able to isolate the voices of several users from the ambient noise. The controller uses 48-kHz, 24-bit mono pulse code modulation (PCM).

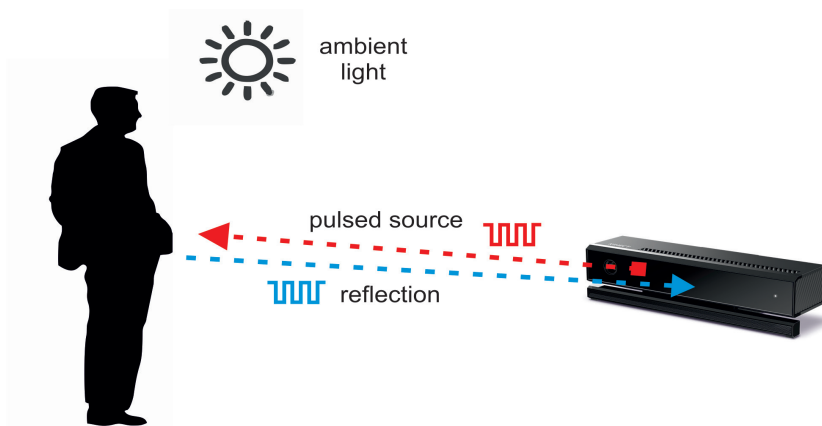


Fig. 3. Operational principles of ToF camera [36]

The RGB camera features a CMOS sensor and enables user facial recognition and image processing. It operates at 1920x1080px at 30 fps. For depth data input, the sensor interprets the data stream from the time-of-flight camera. It can detect the position of points in 3D space with a resolution of 512x424 at 30 fps. The 3D time-of-flight camera projects structured light and then gathers the reflected light. Distance is measured by calculating the phase shift between the emitted and received modulated light. The principles of its operation are presented in Fig. 3.

Either a semiconductor laser or LED diodes are usually used as near-infrared light sources (~850 nm). A modified image sensor then detects reflected light and converts the photonic energy to electrical current [36]. Since Kinect v2 uses a USB 3.0 interface, it supports 2 Gb/s data transfer rates. It can simultaneously track six people and 25 skeletal joints per person — including thumbs. It is also capable of detecting a user's pulse, face expression and muscle tension.

3. System description

3.1. Test stand

A VC 760 milling machine with an open control system (O.C.E.A.N.) [37] developed at the Centre for Mechatronics of the West Pomeranian University of Technology, Szczecin was used for the implementation and testing of the proposed system. O.C.E.A.N system is a conception, of the open, modular, fully reconfigurable and modifiable control system of the CNC cutting machine. The main goal of the system is to introduce an open interface to develop diagnostic functions of the milling machine, implement dynamic and static corrections within the cutting process, extend and test system kernel functions, and introduce new functions to users. The novelty of the system lies in its flexibility and the possibilities for extensions which would improve machining quality.

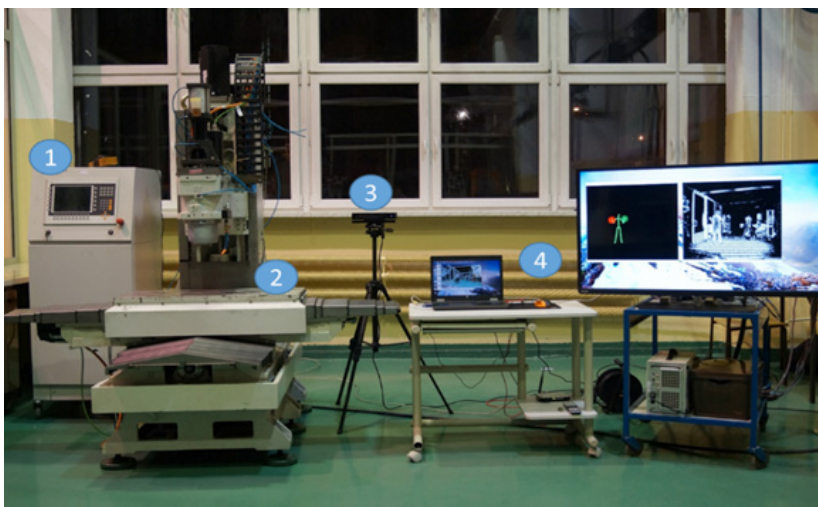


Fig. 4. Test stand

The VC 760 machine has two kinematic chains of body element movement: object and tool branches. It has screw couplings in 3 axes, with power applied directly from the motor. The test stand is presented in Fig. 4.

This equipment consists of:

- 1) O.C.E.A.N. control system, based on Acopos intelligent servo drives manufactured by B&R;
- 2) VC 760 milling machine;
- 3) Microsoft Kinect v2 sensor;
- 4) PC used for image processing and communication with the CNC machine.

3.2. Communication

The Kinect controller communicates with the milling machine according to the scheme presented in Fig. 5. The control system of the VC 760 milling machine communicates with external devices through an OPC server, operating in the PLC driver (manufactured by B&R). Positioning data can be sent in two modes:

- ▶ Absolute – the table or the workpiece is moved directly to a position defined by (x, y, z) coordinates;
- ▶ Incremental – the table or the workpiece is moved by a pre-defined vector from its current position ($\Delta x, \Delta y, \Delta z$).

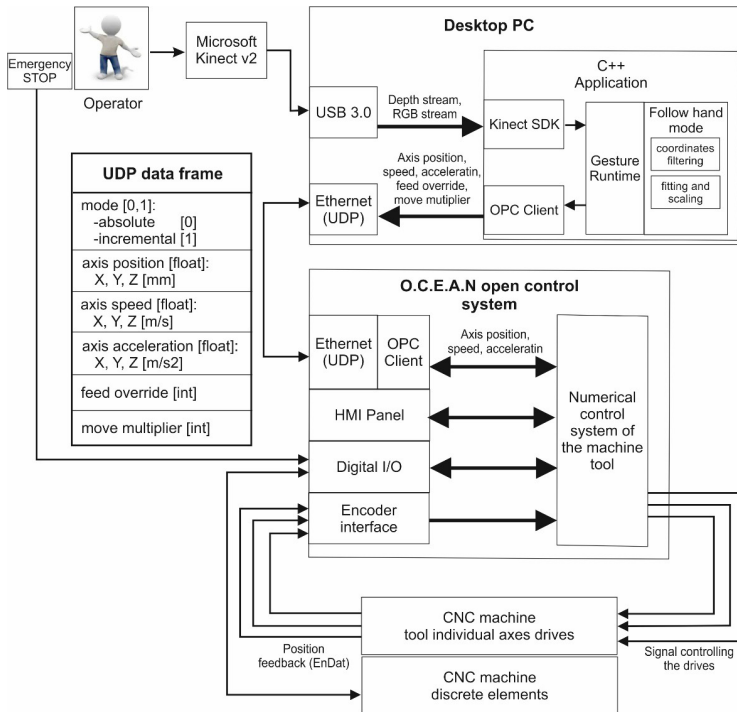


Fig. 5. The communication scheme of the gesture control system

Coordinates are fed using UDP protocol in specially designed frames. Every frame contains: coordinates, velocity data and acceleration data for each operation, global feed override and *the* move multiplier. Trajectory is automatically generated by the O.C.E.A.N. control system, based on received points or translations. For safety purposes and to avoid collision, the upper and lower motion ranges for all the axes are fixed. The second part of the program responsible for the Microsoft Kinect sensor operation, gesture recognition and feeding coordinates to the O.C.E.A.N. controller, was written in the Visual Studio 2013 environment using the Kinect for Windows Runtime package: Kinect SDK & Developer Toolkit library, Kinect Studio and Visual Gesture Builder. Visual Gesture Builder (VGB) is an interactive tool for building models of body gestures using the random forest regression (RFR) machine learning classifier. VGB utilises two detection technologies – AdaBoostTrigger (used to detect discrete gestures) and RFRProgress (used to detect continuous gestures).

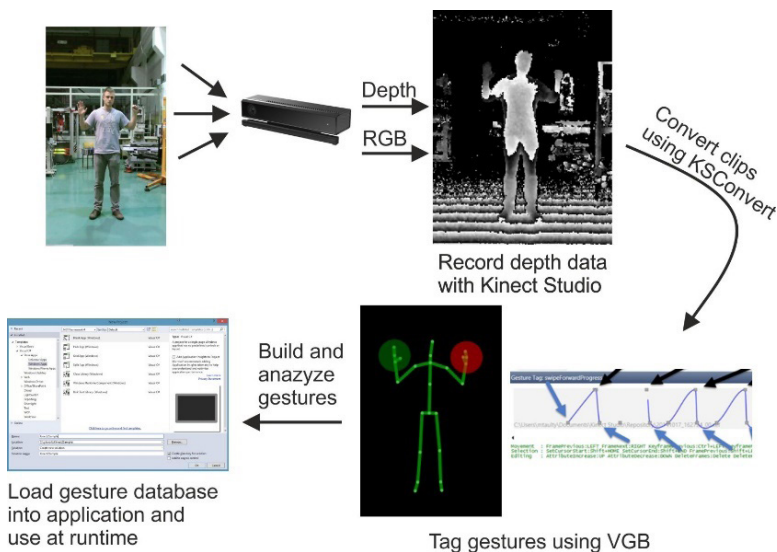


Fig. 6. The data-driven process of creating a gesture detector using Visual Gesture Builder

The developed program consists of five modules responsible for:

- ▶ communication with Kinect v2 sensor – receiving data about the position of tracking points
- ▶ filtering of received coordinates – smoothing of tracking points using a simple average filter
- ▶ gesture recognition – we pre-programmed the gestures described in 3.3. Algorithms implemented directly in Kinect SDK and Visual Gesture Builder were used to detect two gestures. To improve the repeatability of gesture detection, the program was taught by gestures performed by three people of different height and weight. The process of ‘learning’ gestures is presented in Fig. 6.
- ▶ Position, fitting and scaling – this part of the program is responsible for a calibration of the machine’s coordinates and the operator’s hand. This is critical for safety reasons, as

without the function, the machine would first make an unknown move to assume the position signalled by the operator's hand.

- ▶ Communication with the O.C.E.A.N. controller – sending, receiving and scaling of data about the position of the table and the spindle.

3.3. Control methods

Programming of the VC 760 can be performed in manual or semi-automatic mode. Both modes use the appropriate buttons on the control panel, on the remote pedant or G-code commands (numerical control programming language). To perform simple adjustment movements in manual mode, the operator must perform the following operations:

- 1) move close to the operator panel (HMI)
- 2) set the following parameters with the appropriate button or command:
 - c) operating mode – manual
 - d) axis that will be controlled
 - e) movement speed
 - f) positioning accuracy
- 7) use the appropriate buttons to move the selected axis

Use of the semi-automatic mode in which it is possible to move multiple axes at the same time requires the following steps:

- 1) move close to the operator panel (HMI)
- 2) set the following parameters with the appropriate button or command:
 - c) operating mode – manual
 - d) axis that will be controlled
 - e) movement speed
 - f) positioning accuracy
- 7) write a program in G-code
- 8) confirm the code execution with the appropriate buttons

Operation of the CNC in each mode is inconvenient (due to the distance from the machining zone). For this reason, the control pedant is commonly used. This is connected by a flexible cable to the control system of the machine. The remote pedant allows the operator to control the axes movement close to the workpiece but only in manual mode. However, performing positioning moves requiring multiple axes simultaneously, forces the operator to write a simple program in G-code. Using buttons to move the axis is not intuitive because it causes problems such as mistaking different buttons for each other. The use of buttons based interface for the operation of more than one axis is not intuitive and may lead to dangerous situations. A deliberate procedure by constructors is therefore limiting the possibility of using the buttons for one axis only. For the aforementioned reasons, works have been initiated on the development of an intuitive system to control a machine tool using gestures. The idea of a manipulation system for the manual movement of body units of a CNC tool is based on

gestures sets. This movement should be possible to perform within all the available axes of the machine tool simultaneously. The speed of movement of body elements should be dependent upon the movement speed of the operator's hand.

Three sets of gestures were developed for the purposes of the study. The first set includes four basic gestures to be performed by the left hand: start, stop, confirm and change precision (Fig. 7).



Fig. 7. First set of gestures - basic gestures (for left hand)

The second set is used to control the table and the spindle – the machine follows the right-hand gestures. Hand gestures performed in the air are followed by the table. Up and down gestures control the spindle. The velocity of movement is proportional to the velocity of gestures performed by the operator. The tracking of the operator's hand starts when they clench their right hand into a fist and the operator confirms it by performing the start gesture with their left hand. From now, every hand movement is followed by movement of the machine along the appropriate axis (Fig. 8). An example of the gesture sequence from the second set is shown in Table 1.

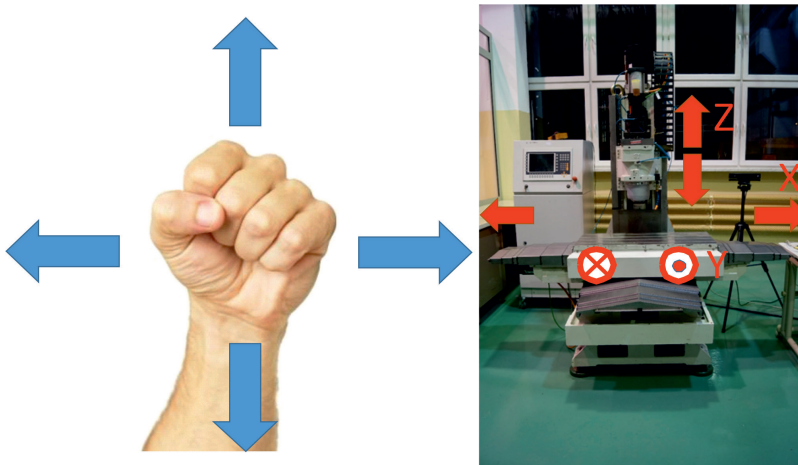













Fig. 8. Axis control in follow mode

Table 1. An example second set gesture sequence








Step	Left Hand	Right Hand	Preformed operation
step 1		insignificant	stop machine and cancel all active tasks
step 2	insignificant		set machine to be ready and cancel alarms
step 3			activate drives and start following right hand
step 4	axis movements – following right hand mode (Fig. 7)		
step 5		insignificant	stop machine and cancel all active tasks
step 6	insignificant		set machine to be ready and cancel alarms
step 7			change the precision of movement
step 8			confirm change of precision



step 9		activate drives and start following right hand with other movement precision
step 10	axis movements – following right hand mode (Fig. 7)	

The third set of gestures is used to select the axis to be controlled. The axis is selected by indicating its number with the fingers. After selecting the axis and confirming it with a ‘confirm’ gesture, it is possible to control the machine with the right hand clenched into a fist in the tracking mode. An example of the gesture sequence from the third set is shown in Table 2.

Table 2. An example third set gesture sequence

Step	Left Hand	Right Hand	Preformed operation
step 1		insignificant	stop machine and cancel all active tasks
step 2	insignificant		set machine to be ready and cancel alarms
step 3	insignificant		select the indicated axis
step 4			confirm axis selection
step 5			activate drives and start movement in the selected axis
step 6	axis movements – control selected axis		

4. EXPERIMENT RESULTS

The proposed system was tested on a test stand in the technology hall of the Institute of Materials Science and Engineering, West Pomeranian University of Technology, Szczecin. Experiments revealed that the proposed control system for the VC 760 milling machine was much more intuitive than the standard control panel or the joystick. In the tested system, the operator does not have to know how to use the control panel or command line, or know specific commands. The developed solution only requires the operator to know a basic set of gestures. When the operator takes position in front of the Microsoft Kinect sensor, they are ready to control the VC 760 milling machine. During the operation of the system, the background did not seem to affect its stability. The system could efficiently control the milling machine both with a homogeneous static white background and in a more standard environment (people moving around, objects in the background). The system worked correctly in complete darkness and in poor light conditions. However, a few problems were detected. Firstly, the system was sensitive to intensive light directed at the Kinect v2 sensor. The readings of hand coordinates and gesture detection were either random or assumed maximum values thus rendering proper control of the system impossible. Secondly, the repeatability of gesture detection worsened when the distance between the operator and the sensor was increased or when the operator was located not in front of the sensor. This issue was observed in gestures involving fingers (axis choice, precision). In some instances, the sensor could not detect gestures when the end of one gesture merged into the beginning of a successive gesture. This problem was probably caused by an inadequate number of samples fed into the sensor during the teaching mode.

Also problematic was occlusion (moment when the hand being obscured) and gestures being performed at too great a speed. Because the hand is a complicated object with more than twenty DOF, it can have many different shapes and positions. Recognising them is particularly difficult, this is due to the mutual shadowing of the palm and fingers which hinders segmentation/extraction and consequently, the recognition of precise gestures. In addition, the speed of a moving hand is high and can reach up to 5 m/s in linear displacement and 300 deg/s in wrist rotation [38]. These factors necessitate either the use of high-speed cameras or advanced algorithms for slower cameras. This is related to delays during gesture recognition due to the problem of processing large amounts of data in real time. Average error rate for stop, start and confirm gestures was 10% for false positive errors and 5% for false negatives. Worse results were obtained for change precision gestures – 22 % (false positives) and 9% (false negatives). Another problem was the control of the machine when the operator was close to the worktable. This is a very difficult situation for the system, because the operator is, to a large degree, obscured by the machine and is thus not visible to the sensor. Solving this problem requires further research and will probably enforce the use of an additional sensor for tracking the operator when working close to the machine. Due to the possibility of a collision of the machine with the operator, incorrect gesture recognition or uncontrolled movement of the operator, programmed axis limits were applied. In addition, the operator was equipped with a safety button which triggers the immediate activation of the brake motors. At the start



of the study, a hand switch was used; however, it was decided to use a foot switch during the tests for ergonomic reasons. After all, the problem of incorrect movements or gesture classification is much more complex and requires deeper analysis and more comprehensive solutions than simply having a safety button. Thus, the safety system for the gesture control system will be the subject of further research.

Since the main aim of the research was to test the possibility of using gesture control to operate CNC machines, at this stage of the project, the precision with which an operator can control the machine tool was not measured. However, based on research on Kinect v2 [39], it is possible to estimate the Kinect distance sensor accuracy. Depth accuracy in the middle of the sensor operating range (1000-2000 mm) is marked as: ± 2 mm, while the at the ends of the range is marked as: ± 4 mm. Additionally, changing movement resolution (feed override - a gesture change), the operator can obtain leverage effects. Hand movements in the range of 1-1000 mm, will move an appropriately selected axis by 0.001-1 mm. A similar solution is used in CAD software. Very precise parts are designed using an inaccurate computer mouse.

Given the purposes of the system, the values are sufficient for tasks such as pre-positioning the machine, tool setting and referencing the. It is also possible to apply additional support functions.

5. Summary

The paper presents and discusses the concept and the constituent elements of the proposed system and test results. The advantages and potential problems of the system are shown. Based on the conducted research and test results, a conclusion can be drawn that effective gesture control requires the careful and thoughtful selection of gestures. Gestures should be able to convey precise control signals. Therefore, gestures should be simple and intuitive for the operator. They should not cause problems with their recognition as this may lead to undesirable delays and latency in the system. The tests demonstrated that the Microsoft Kinect v2 sensor, which is dedicated to computer games, was quite efficient at communicating with the control system of the CNC machine tool and properly detected simple gesture commands. The presented results only reveal the potential of gesture control. An industrial implementation of a gesture control system would involve the use of professional tracking systems and advanced program functions that support various technological aspects of the programming CNC machine tools. To eliminate control panels, future systems would also feature voice control and augmented reality systems.

In this paper, a new approach to control basic function of CNC machines which uses the technology vision and gesture recognition has been proposed. The proposed core set of gestures do not allow the use of all possibilities and options of advanced CNC control systems, because the aim of the research is not to replace complex operator panels but add new capabilities to the CNC system and facilitating simple control movements. The G-code programming procedure is complicated, especially without access to the keyboard. Therefore,

the complete replacement of the operator panel with the control system of gestures still requires a lot of research.

The proposed system extension represents a new approach with regard to CNC machine control. It uses vision and gesture recognition technologies. Although it has some drawbacks, it undoubtedly has a significant advantage – it is very intuitive and does not require CNC operators and programmers to undergo difficult and long-lasting training. It also allows the simultaneous use of a traditional control panel.

This project is financed by the National Centre for Research and Development, Poland (NCBiR), under the Applied Research Programme – Grant agreement No. PBS3/A6/28/2015 and Dean's Grant for Young Scientists No. 517-06-011-5404/17

References

- [1] Shuo-Peng L., Hsu J., Chun-Chien T., *Development of the common human-machine interface for multi-axis machine tools*, IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), Kaohsiung, Taiwan, 2012, 650–653.
- [2] Saków M., Miądlicki K., Parus A., *Self-sensing teleoperation system based on 1-dof pneumatic manipulator*, Journal of Automation Mobile Robotics and Intelligent Systems, Vol. 11, 64–76.
- [3] Saków M., Parus A., Pajor M., Miądlicki K., *Nonlinear inverse modeling with signal prediction in bilateral teleoperation with force-feedback*, Methods and Models in Automation and Robotics (MMAR), Międzyzdroje, 2017, 141–146.
- [4] Kopp S., Wachsmuth I., *Gesture in Embodied Communication and Human Computer Interaction*, 1 ed., Springer-Verlag, Berlin 2010.
- [5] Majewski M., Kacalak W., *Smart Control of Lifting Devices Using Patterns and Antipatterns*, [in:] *Artificial Intelligence Trends in Intelligent Systems: Proceedings of the 6th Computer Science On-line Conference*, R. Silhavy, R. Senkerik, Z. Kominkova-Oplatkova, Z. Prokopova, P. Silhavy (Eds.), Springer International Publishing, Cham 2017.
- [6] Majewski M., Kacalak W., *Innovative Intelligent Interaction Systems of Loader Cranes and Their Human Operators*, [in:] *Artificial Intelligence Trends in Intelligent Systems: Proceedings of the 6th Computer Science On-line Conference*, R. Silhavy, R. Senkerik, Z. Kominkova-Oplatkova, Z. Prokopova, P. Silhavy (Eds.), Springer International Publishing, Cham 2017.
- [7] Corcoran P., *To Gaze with Undimmed Eyes on All Darkness [IP Corner]*, IEEE Consumer Electronics Magazine, Vol. 4, 99–103.
- [8] Jing Y. et al., *A novel hand gesture input device based on inertial sensing technique*, 30th Annual Conference of IEEE Industrial Electronics Society (IECON), Busan, South Korea 2004, Vol. 3, 2786–2791.
- [9] Almetwally I., Mallem M., *Real-time tele-operation and tele-walking of humanoid Robot Nao using Kinect Depth Camera*, 10th IEEE International Conference on Networking, Sensing and Control (ICNSC), Evry, France 2013, 463–466.

- [10] Sanna A., Lamberti F., Paravati G., Manuri F., *A Kinect-based natural interface for quadrotor control*, (in en), *Entertainment Computing*, Vol. 4, 179–186.
- [11] Pajor M., Miądlicki K., Saków M., *Kinect sensor implementation in FANUC robot manipulation*, *Archives of mechanical technology and automation*, Vol. 34, 35–44.
- [12] Gośliński J., Owczarek P., Rybarczyk D., *The use of Kinect sensor to control manipulator with electrohydraulic servodrives*, *Pomiary, Automatyka, Robotyka*, Vol. 17, 481–486.
- [13] Pietruszewicz K., Miądlicki K., *Gestures can control cranes*, *Control Engineering*, Vol. 61, 14.
- [14] Majewski M., Kacalak W., Budniak Z., Pajor M., *Interactive Control Systems for Mobile Cranes*, [in:] *Advances in Intelligent Systems and Computing*, Vol. 661, ed, 2018, 10–19.
- [15] Majewski M., Kacalak W., *Human-Machine Speech-Based Interfaces with Augmented Reality and Interactive Systems for Controlling Mobile Cranes*, *Interactive Collaborative Robotics: First International Conference, Budapest 2016*, 89–98.
- [16] Liu W., Ren H., Zhang W., Song S., *Cognitive tracking of surgical instruments based on stereo vision and depth sensing*, *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Shenzhen, China 2013, 316–321.
- [17] Gallo L., Placitelli A., Ciampi M., *Controller-free exploration of medical image data: Experiencing the Kinect*, *24th International Symposium on Computer-Based Medical Systems (CBMS)*, Bristol, United Kingdom 2011, 1–6.
- [18] Stateczny K., Pajor M., *Project of a manipulation system for manual movement of CNC machine tool body units*, *Advances in Manufacturing Science*, Vol. 35, 33–41.
- [19] Berman S., Stern H., *Sensors for Gesture Recognition Systems*, *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, Vol. 42, 277–290.
- [20] Kalgaonkar K., Raj B., *One-handed gesture recognition using ultrasonic Doppler sonar*, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Taipei, Taiwan 2009, 1889–1892.
- [21] Ketabdard H., Ali K., Roshandel M., *MagiTact: interaction with mobile devices based on compass (magnetic) sensor*, *15th international conference on Intelligent user interfaces*, Hong Kong, China 2010, 413–414.
- [22] Pajor M., Stateczny K., Pietruszewicz K., *Virtual reality applied for programming CNC machine tools*, *Control Engineering*, Vol. 60, 50.
- [23] Miądlicki K., Pajor M., Saków M., *Ground plane estimation from sparse LIDAR data for loader crane sensor fusion system*, *Methods and Models in Automation and Robotics (MMAR)*, Międzyzdroje 2017, 717–722.
- [24] Miądlicki K., Pajor M., Saków M., *Real-time ground filtration method for a loader crane environment monitoring system using sparse LIDAR data*, *Innovations in Intelligent Systems and Applications (INISTA)*, Gdynia 2017, 207–212.
- [25] Shuying Z., Li S., Wenjun T., *A method of dynamic hand gesture detection based on local background updates and skin color model*, *International Conference on Computer Application and System Modeling (ICCASM)*, Taiyuan, China 2010, Vol. 5, V5-657-V5-660.

- [26] Mohan P., Srivastava S., Tiwari G., Kala R., *Background and skin colour independent hand region extraction and static gesture recognition*, Eighth International Conference on Contemporary Computing (IC), Noida, India 2015, 144–149.
- [27] Haiting Z., Xiaojuan W., Hui H., *Research of a Real-time Hand Tracking Algorithm*, International Conference on Neural Networks and Brain (ICNNB), Beijing, China 2005, Vol. 2, 1233–1235.
- [28] Ionescu D., Suse V., Gadea C., Solomon B., Ionescu B., Islam S., *An infrared-based depth camera for gesture-based control of virtual environments*, IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), Milan, Italy 2013, 13–18.
- [29] Feng Z., Du P., Song X., Chen Z., Xu T., Zhu D., *Research on Features Extraction from Frame Image Sequences*, International Symposium on Computer Science and Computational Technology (ISCSCT), Shanghai, China 2008, Vol. 2, 762–766.
- [30] Nianjun L., Lovell B. C., Kootsookos P. J., Davis R. I. A., *Model structure selection & training algorithms for an HMM gesture recognition system*, Ninth International Workshop on Frontiers in Handwriting Recognition (IWFHR), Tokyo, Japan 2004, 100–105.
- [31] Liu K., Chen C., Jafari R., Kehtarnavaz N., *Multi-HMM classification for hand gesture recognition using two differing modality sensors*, Circuits and Systems Conference (DCAS), Dallas, USA 2014, 1–4.
- [32] Fu X., Lu J., Zhang T., Bonair C., Coats M. L., *Wavelet Enhanced Image Preprocessing and Neural Networks for Hand Gesture Recognition*, IEEE International Conference on Smart City/SocialCom/SustainCom, Chengdu, China 2015, 838–843.
- [33] Watanabe T., Yachida M., *Real time gesture recognition using eigenspace from multi-input image sequences*, Third IEEE International Conference on Automatic Face and Gesture Recognition, Nara, Japan 1998, 428–433.
- [34] Shehu V., Dika A., *Curve similarity measurement algorithms for automatic gesture detection systems*, 35th International Convention on Information and Communication Technology, Electronics and Microelectronics, Opatija, Croatia 2012, 973–976.
- [35] Chen M., AlRegib G., Juang B.H., *A new 6D motion gesture database and the benchmark results of feature-based statistical recognition*, IEEE International Conference on Emerging Signal Processing Applications (ESPA), Las Vegas, USA 2012, 131–134.
- [36] Van Nieuwenhove D., Van der Tempel W., Grootjans R., Kuijk M., *Time-of-flight Optical Ranging Sensor Based on a Current Assisted Photonic Demodulator*, Annual Symposium of the IEEE Photonics Benelux Chapter, Ghent, Belgium 2006, 209–212.
- [37] Domek S., Pajor M., Pietruszewicz K., Urbański Ł., *Experimental open control system OCEAN for linear drives*, Inżynieria Maszyn, Vol. 16, 40–49.
- [38] Erol A., Bebis G., Nicolescu M., Boyle R.D., Twombly X., *Vision-based hand pose estimation: A review*, Computer Vision and Image Understanding, Vol. 108, 52–73.
- [39] Wang Q., Kurillo G., Ofli F., Bajcsy R., *Evaluation of Pose Tracking Accuracy in the First and Second Generations of Microsoft Kinect*, International Conference on Healthcare Informatics (ICHI), Dallas, USA 2015, 380–389.

Agnieszka Żyra

Rafał Bogucki

Sebastian Skoczypiec (skoczypiec@mech.pk.edu.pl)

Institute of Production Engineering, Faculty of Mechanical Engineering,
Cracow University of Technology

AUSTENITIC STEEL SURFACE INTEGRITY AFTER EDM IN DIFFERENT DIELECTRIC LIQUIDS

STAN POWIERZCHNI STALI AUSTENITYCZNEJ PO OBRÓBCE EDM W RÓŻNYCH CIEKŁYCH DIELEKTRYKACH

Abstract

Electrodischarge machining (EDM) can be used as an alternative machining method compared to conventional ones, especially while very good surface integrity and high machining accuracy during machining of difficult-to-cut materials is needed. In EDM, the dielectric liquid has a crucial influence on the technological material surface integrity since it ensures the occurrence of controlled electrical discharges between the tool and the workpiece, cooling and solidification of gaseous EDM debris, removing erosion products and also dispersing heat generated during the process. In the paper, the influence of a carbon-based and a water-based dielectric liquid on the selected structural and morphological characteristics of 304 stainless steel after EDM sinking was investigated. The surface roughness, micro hardness and quality as well as the chemical changes after the corrosion test of machined surfaces were analysed.

Keywords: electrodischarge machining, dielectric, surface layer, austenitic steel

Streszczenie

Obróbka elektroerozyjna (EDM) stanowi alternatywę dla konwencjonalnych metod obróbkowych, szczególnie przy kształtowaniu materiałów trudnoskrawalnych, gdy konieczne jest otrzymanie powierzchni o bardzo dobrej jakości. W EDM ciekle dielektryk ma kluczowy wpływ na właściwości technologicznej warstwy wierzchniej materiału obrabianego, umożliwiając zachodzenie kontrolowanych wyładowań elektrycznych pomiędzy elektrodą roboczą a przedmiotem obrabianym, schładzanie i usuwanie produktów powstałych w wyniku obróbki oraz odprowadzanie wygenerowanego w trakcie procesu ciepła. W artykule przedstawiono wpływ dielektryka węglowodorowego i wodnego na wybrane cechy strukturalne i morfologiczne stali nierdzewnej 304 po procesie drążenia elektroerozyjnego. Analizie poddano chropowatość powierzchni, mikrotwardość oraz zmiany składu chemicznego obrabianych powierzchni po przeprowadzonej próbie korozyjnej.

Słowa kluczowe: obróbka elektroerozyjna, dielektryk, warstwa wierzchnia, stal austenityczna

1. Introduction

Electrodischarge machining is one of the most commonly used non-conventional machining methods in all branches of the industry [15]. Materials with electrical conductivity above 10^{-2} S/cm can be machined by the EDM process [7], irrespective of the machined material chemical structure and mechanical properties (such as hardness). EDM is also characterised by very high machining accuracy and achievable surface roughness [1, 11]. Another advantage of using the EDM machining method is the possibility of getting complicated shapes of elements, also with thin walls, 3-D macro, micro and nano features, which would be impossible to obtain with conventional machining methods [11, 15].

In EDM-sinking, material removal takes place as a result of the occurrence of a plasma channel (erosion effect) between the tool electrode and the workpiece surface, separated by dielectric liquid [2, 12]. In each pulse, the machining material is melted and evaporated during each discharge occurring at a single location [19]. The debris in the molten phase is cooled, resolidified and then evacuated by dielectric liquid from the machining gap. The thermal character of the material removal mechanism causes the occurrence of the heat-affected zone in the workpiece surface layer [2]. The heat-affected zone consists of an upper layer of the recast and is characterised by a solidified material, the occurrence of microcracks, high porosity, increased dimensions of grains arising from other alloys, which come either from the tool electrode or the dielectric liquid [5]. Recently, many experiments [9, 13, 14, 16] have been conducted to show how complicated the mechanism of material removal during single electrodischarge machining is. However, it is still challenging to observe the complex material removal in real EDM conditions, where a huge number of discharges occur.

The character of the EDM-machined surface integrity changes is connected with machining conditions, such as pulse time, current intensity, tool electrode material and the applied dielectric liquid [23]. Dąbrowski [6, 17] points out that the main factors which influence the machined element's surface integrity are current intensity and pulse time, while the other machined parameters are defined. The current intensity increases the results in a higher amount of material removal during a single discharge. The current intensity growth causes the surface quality deterioration [22, 25]. The extended pulse duration reduces the energy density (which improves surface quality by reducing surface roughness) of the discharge spots by expanding the plasma channel and, as a consequence, the material removal rate decreases and the tool electrode wear is reduced [12, 22, 23]. Tool wear, which is determined by tool electrode material properties, has a significant influence on the EDM machining accuracy. The tool electrode is removed during the EDM process simultaneously by generated sparks, causing a characteristic corner and end (radial and axial) working electrode wear [12, 20]. It is also worth mentioning that the melted electrode material can diffuse to the workpiece surface layer. Świercz and Oniszczyk [18, 21] attempt in their works to solve an interesting problem of the current pulses' character influence on the machined surface integrity, as basic factors determining the EDM machining results. It was found that for short pulse times and low current values, the surfaces' geometrical structure is characterised by a huge number of rough peaks. The current intensity increase results in the increase of the discharge channel's

diameter and power increase, which definitely brings higher roughness, characterised by a higher height as well as the distance between particular rough peaks.

The type of dielectric liquid is another important factor, which determines the machined surface integrity [23]. Proper dielectric circulation in the machining gap is required to maintain needed peak performance and to control the electric spark [4]. There are several functions of dielectrics: insulation, ionisation, cooling and waste removal [4]. Due to the fact that each dielectric has a specific composition and physical properties, it is very important to choose the one that would be optimal, ensuring the best surface integrity and accuracy in a particular situation. The type of a dielectric influencing the technological surface properties is still a complex problem, which will be described and investigated in more details further in this paper.

2. The role of a dielectric in EDM

Many authors [4, 17, 26] point to the crucial role of the dielectric type in electrodischarge machining. Dielectrics ensure the possibility of the occurrence of controlled electrical discharges between the tool electrode and the workpiece, cooling and solidifying of gaseous EDM debris from the discharge, removing erosion products from the machining gap and also dispersing heat generated during the process from the tool electrode and workpiece material [4]. Currently used dielectrics can be divided into three groups: hydrocarbon-based dielectric liquids, water-based dielectric liquids and gaseous dielectrics [4].

Table 1. Physical parameters of chosen dielectrics with examples of commercial product names [26–28]

Physical parameter	Dielectric type		
	kerosene	De-ionised water	air
Thermal conductivity λ (W/m·K)	0.14	0.62	0.016
Melting point δ (°C)	–	0	–
Specific heat C (J/kg·°C)	2100	4200	1.005
Viscosity η (Pa·s)	2.3×10^{-6}	1×10^{-6}	
Density ρ (kg/m ³)	860	1000	1.29
Composition (vol%)	100% hydrocarbon	100% H ₂ O	78% N ₂ , 21% O ₂ , 1% other gases

Physical parameter	Dielectric type		
	kerosene	De-ionised water	air
Type of EDM with used dielectric	EDM-sinking	EDM-wire, Micro-hole drilling	Dry EDM
Examples of dielectric liquids	Carbon-based: 30 NEUTRAL, OIL VITOL-2, VITOL-2-S, VITOL-KO (manufacturer: Sodick Co., Ltd), EDM 3033 (manufacturer: EDM Zap)	Water-based: VITOL-KS, VITOL-KN (manufacturer: Sodick Co., Ltd)	

Carbon-based dielectric liquids were applied at the earliest and are still the most commonly used, especially during EDM-sinking, because of their low viscosity, which improves flushing of the machining area [26]. On the other hand, there are a few drawbacks connected with a low flash point and evaporation temperature (quick transformation to the volatile state). What is important, fumes arising during machining (carbon monoxide and methane) are toxic. Fumes generated by electric discharge diffuse deeply into the white layer, which results in generating carbide separations [4]. That affects further technological operations – during heat treatment, the carbonisation of the surface layer increases, and along the grain's boundaries, the microcracks expansion occurs [4, 26] These kinds of liquids are not environmentally friendly, they also have a negative influence on human health. On the other hand, using carbon-based dielectric liquids makes it possible to get higher machining accuracy compared to water-based dielectric liquids [4].

Water-based dielectric liquids (de-ionised water, tap water, pure water with additives such as ethylene glycol, dextrose, sucrose, glycerine) are more thermally stable compared to carbon-based dielectric liquids [4] because they have eight times higher vaporisation energy and lower boiling-point. Greater thermal stability facilitates receiving higher discharge energy and as a consequence results in obtaining higher machining efficiency and lower working electrode wear [26]. The surface roughness is better, the number of microcracks is reduced compared to machining with carbon-based liquids. Further heat treatment does not affect microcracks expansion; however, material oxidisation is observed [4]. On the other hand, properties of water-based liquids are hard to stabilise (in the course of time the isolation features are lost) and the machining accuracy is definitely lower compared to carbon-based dielectric liquids [4]. The main area of application of such a kind of dielectrics is electrodischarge wire cutting and high ration hole drilling. An important advantage of using water-based dielectrics is lower environment contamination. What is more, they do not affect the human body [4].

Dry electrodischarge machining is a method, which uses gas (e.g. air or oxygen) as a dielectric. The melted material is removed from machining gap by the high-pressure gas flow, supported by a thin-walled pipe electrode [4]. The role of gas is to cool, solidify and prevent molten material from welding into the tool electrode and machined surface. Furthermore,

high-pressure gas jet extinguishes the plasma channel between sequent discharges, which facilitates dielectric strength recovering in the machining gap [4]. It should be emphasised that tool electrode wear ratio during dry EDM is very low and, what is more, it is independent of the single pulse duration [4]. Dry EDM is not commercially used in industrial applications yet [26], even though it is economically and environmentally friendly, compared to other dielectrics, because of huge problems with heat dissipation during electrodischarge machining.

Zhang et al. [26] investigated the external and sectional appearance of the craters after a single pulse discharge, to characterise material removal in different dielectrics. Significant differences were found, comparing geometrical craters' shapes formed while using different dielectric types (oxygen, air, de-ionised water, kerosene, W/O emulsion) in the same machining conditions. It was stated that the machining efficiency was lower in the gaseous dielectric. However, observed melted material volume in kerosene liquid was much lower compared to other dielectrics. It shows that there is a huge difference between the volume of the melted material and the removed material (it is known that a higher removal material rate is observed in the carbon-based dielectric liquid than in water-based or gaseous dielectrics) [26]. Basing on the experiments and simulations, the authors [26] suggested that a higher efficiency of material removal rate is connected with a higher pressure above the discharge point [26]. On the basis of a literature review, it was stated that there are huge differences between dielectrics, which have the crucial influence on machined surface integrity.

The technological properties of EDM machined elements determine the possibility of their applications in specific environments, where high machining accuracy and very good surface integrity are needed. The main aim of this work was to identify the influence of carbon-based and water-based dielectric liquids on the selected structural and morphological characteristics of 304 stainless steel after electrodischarge sinking. The investigated material is commonly used in all the branches of industry (automotive, chemical, food-processing, construction, turbomachinery industry and an energy sector) due to the very good thermal, physical, chemical and mechanical properties. One of the most important features of austenitic steels is high corrosion resistance and resistance to adverse working environments (high oxidation, acidity, temperature) [3]. Austenitic stainless steels are also commonly used in production of medical devices and implants because of corrosion resistance, formability, reasonable fatigue resistance, due to their biocompatibility ("tolerant of life", can be understood as "being able to cope with the full variety of conditions prevailing during its lifetime") [8]. In this aspect, EDM can be used for efficient production of specific medical devices, such as tools for insertion and extraction or recovery of implants, surgical cathodes and syringe components, splints and supports for orthotic and prosthetic devices, bone and jam reamers for dental implants, tooling and dies for manufacturing and stamping medical equipment and tools [7]. In reasonable cases (implants customisation) [10], EDM can also be used for shaping surfaces in surgical screws and bolts or knee, shoulder and hip joints implants [7]. It is important to underline that manufacturing technology has a significant influence on biocompatibility, and especially when machining is carried out in the final stage of the production process one can state that that biocompatibility is determined by manufacturing technology.

3. Materials and methods

Sinking electrodischarge machining of 304 austenitic stainless steel was conducted on a research post equipped with electrodischarge generator BP 95, manufactured by a Polish company ZAP B.P. Końskie-Kutno (Fig. 1). The influence of two kinds of dielectric liquids on the surface layer quality was investigated: distilled water and carbon-based dielectric liquid – Exxsol D80, at the three values of current intensity (1, 5 and 10 A). These dielectrics were chosen as representatives of the most commonly used kinds of dielectric liquids in industry. Such current intensity values were selected to have stable EDM machining conditions with the generator used. Machining in distilled water was conducted with the $\varnothing 10.84$ mm cylindrical copper working electrode (cathode), machining in Exxsol D80 with the $\varnothing 4.44$ mm cylindrical copper working electrode. Simple polarity of the working electrode and machined element was used (working electrode was the cathode (-), workpiece was the anode (+)). Other machining conditions were constant in both dielectric liquids: pulse time 100 μs and off time 10 μs .

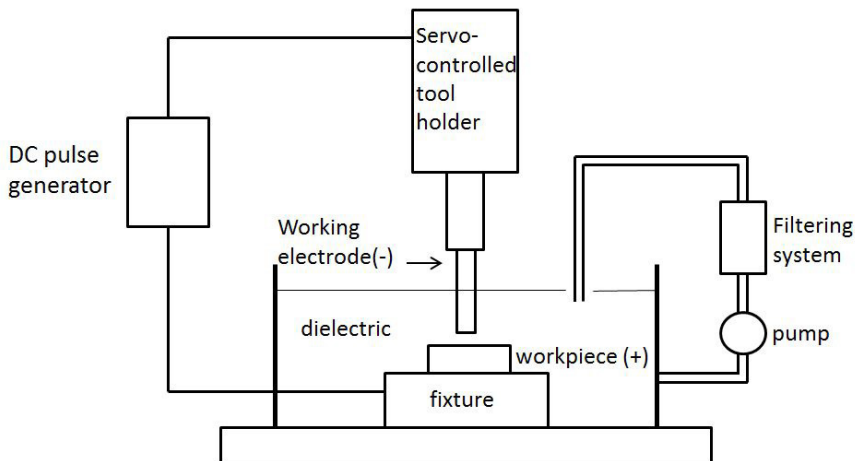


Fig. 1. Scheme of the research test stand

After machining, the following tests were carried out on the machined surfaces. Surface roughness was measured using Taylor Hobson Surtronic 25 with the sampling length of 0.8 mm. The Vickers hardness of each machined surface was determined with the INNOVATEST machine (loading level 0.5 N, at 10 s time). The measurements were repeated five times at each machined surface. Then, with the Scanning Electron Microscope, SEM photos were taken to identify structural and morphological changes of machined surfaces. The determination of changes in the chemical composition after EDM was investigated by the EDS method (Energy Dispersive X-Ray Spectroscopy). This qualitative method was chosen to find out what kind of chemicals appear in the surface layer, after using different dielectric liquids. The corrosion resistance of each sample was investigated by placing the samples in 65% nitric acid in 22 °C for 14 days. After this time, SEM photos and EDS analysis were repeated.

4. Discussion of results

On the basis of the roughness average (R_a) (Fig. 2) and mean roughness depth (R_z) (Fig. 3) measurements on the EDM machined surfaces, it is possible to say that with the increase of current intensity machining causes a proportional increase of R_a and R_z surface roughness. A proportional increase of R_a was also observed at machining in the carbon-based dielectric; however, machining with a current intensity value of 10 A in the carbon-based dielectric liquid caused an R_z roughness decrease (Fig. 3). It is possible that the increase of discharge energy in the machining gap was high enough to remove the whole molten material (exploding pressure) away from the crater. In these machining conditions solidified molten material may not build-up on the craters' edges.

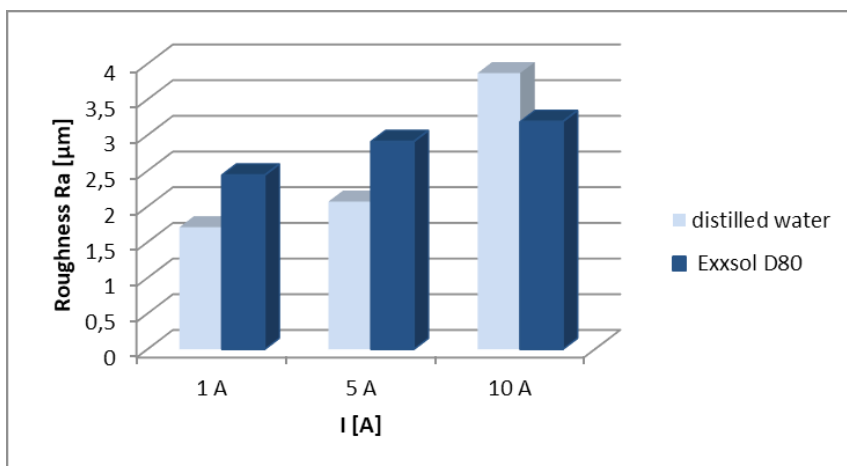


Fig. 2. The influence of current intensity on the value of R_a surface roughness during EDM with different dielectrics

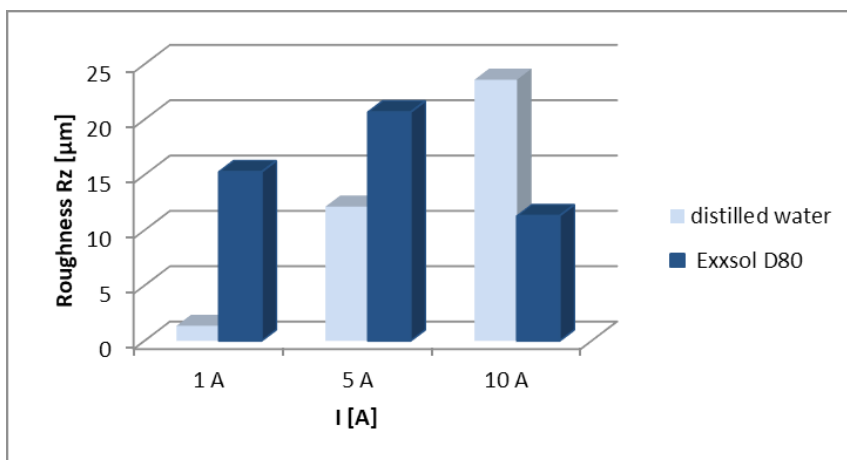


Fig. 3. The influence of current intensity on the value of R_z surface roughness during EDM with different dielectrics

Using distilled water as a dielectric determines the arising density of a single discharge and as a result removing higher material amount during single discharge; therefore, Rz increases. It was also noticed that machining in distilled water at $I = 10\text{ A}$ was more stable than in the case of Exxsol D80. According to Zhang [26] results of electrodischarge machining of 8407 mould steel, the first decrease of crater removal depth was observed either while using distilled water or kerosene as a dielectric liquid, then increase with the increase of pulse duration.

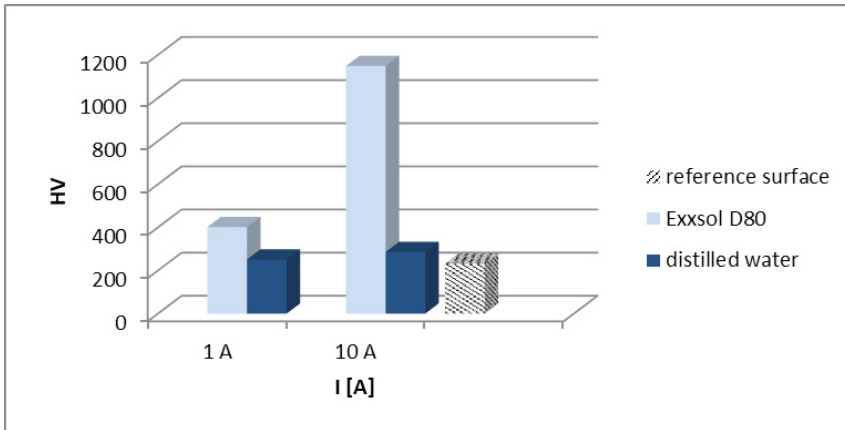


Fig. 4. The influence of current intensity and dielectric liquid type on the micro-hardness in EDM

An influence of the working liquid type on the micro-hardness of the surface layer was observed (Fig. 4). In comparison to the reference sample (surface before EDM) and surface machined in distilled water, machining in carbon-based dielectric causes significant increase of surface hardness. In this case, hardness increases several times with the current intensity increase, starting with the value of HV 400 for $I = 1\text{ A}$, up to HV 1150 for $I = 10\text{ A}$. Hardness after machining in distilled water was in the range of HV 250-300, regardless of current intensity. The hardness of reference sample surface was HV 231; therefore, it should be stated that the impact of the machining in distilled water on the surface hardness can be neglected.

In tables 3 and 4, the SEM photos of electrodischarge machined surfaces in distilled water and Exsol D80 with various current intensities were presented. The SEM photos of these surfaces after the corrosion test were included as well. As it was mentioned in the introduction, a very high temperature is generated during EDM; that is why the material has melted and evaporated and, as a consequence, craters occur on the machined surface [8]. It was noticed that the current intensity increase causes the craters' diameter increase and boundary grain growth in both dielectrics. The increase of the current intensity while EDM machining was carried out in a carbon-based dielectric resulted in a great number of deep microcracks on the surface. This effect was not observed during EDM machining in distilled water, where a small number of microcracks was observed only when machining with $I = 10\text{ A}$. Also, no additional molten spherical debris was observed, which can mean that proper flushing during machining was provided.

Table 2. SEM photos of electrodischarge machined surfaces in distilled water and kerosene; current intensity $I = 1$ A

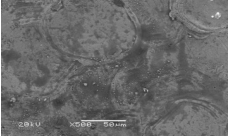
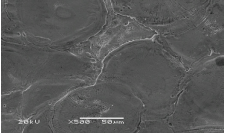

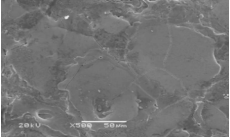
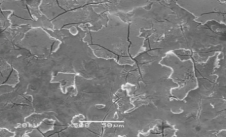
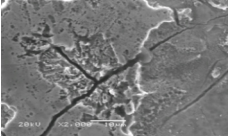
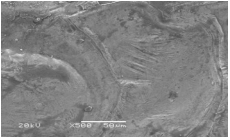
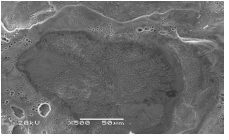
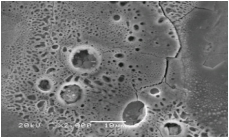
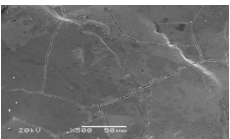
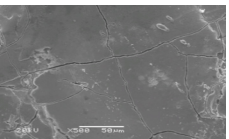
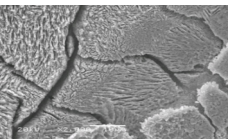
EDM	EDMed surface after corrosion resistance test	
	Zoom x 500	Zoom x 2000
$I = 1$ A Distilled water		
		
$I = 1$ A Exsoll D80		
		

Table 3. SEM photos of electrodischarge machined surfaces in distilled water and kerosene; current intensity $I = 10$ A

EDM	EDMed surface after corrosion resistance test	
	Zoom x 500	Zoom x 2000
$I = 10$ A Distilled water		
		
$I = 10$ A Exsoll D80		
		

After the corrosion test, the increase of oxygen separation on the grain boundaries and near the craters was observed. The pit propagation and its intensity increase with the current amplitude were also noticed. On the sample surface machined in the carbon-based dielectric liquid, the intensity of corrosion processes in the area of crevices and microcracks was significantly higher than in distilled water.

On the basis of EDS analysis (Tab. 5,6), it can be stated that the chemical composition of the surface layer changed, comparing to the raw material. On the machined surfaces, a trace amount of copper was found, which indicates that the material melted from the copper electrode transfers to the surface layer. Additionally, in a corrosive environment, the increase of oxygen amount is observed, which indicates that the surface layer oxidation occurs. It is

worth emphasising that Energy Dispersive X-Ray Spectroscopy, which was chosen as the test method, is just a qualitative method, which enables qualitative analysis only.

Table 4. Example of EDS analysis after EDM in distilled water with current intensity $I = 10A$ and on reference surface

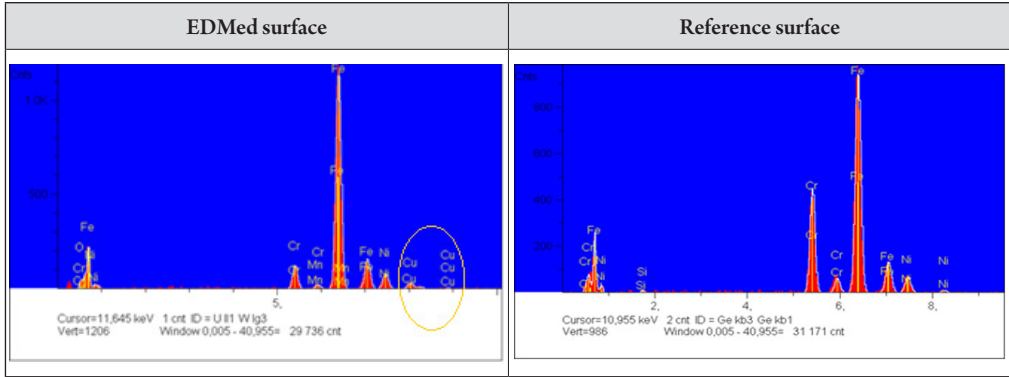


Table 5. Results of EDS analysis on a reference surface, EDMed surface ($I = 10A$, deionised water) and EDMed surface after the corrosion resistance test, [%]

Element	Reference sample	EDM	EDMed surface after corrosion resistance test
O	2,2	10,5	30,1
Cr	19,3	15,5	31,3
Fe	70,3	65,3	33,0
Ni	7,8	6,9	2,9
Cu	0	1,3	0,7

5. Conclusions

In this paper, the attempt of the influence identification of a carbon-based and a water-based dielectric liquid on the selected structural and morphological characteristics of 304 stainless steel after electrodischarge sinking was undertaken. It is important to emphasise a possible influence of austenitic steel surface integrity in the aspects of medical devices, taking into account an increasing demand for implants as well as surgical features. It was stated that the workpiece surface layer is affected by heat; however, an appropriate selection of machining conditions (in the investigated case - current intensity and dielectric type) can reduce this effect. During the research, lower roughness and hardness was obtained by using distilled water as a dielectric liquid (see below in Table 7), which is also connected with fewer microcracks.

Table 6. Surface roughness and hardness obtained by EDM machining in different dielectric liquids

Dielectric type	Current amplitude I [A]	Ra [μm]	Rz [μm]	Current amplitude I [A]	HV	HV standard deviation
Distilled water	1	1.71	1.36	1	251.4	8.55
	5	2.07	12.13	10	286.2	8.03
Exsoll D80	1	2.44	15.27	1	401.5	62.5
	5	2.91	20.67	10	1148.8	141.6

Also, higher corrosion resistance occurred while the material was machined in distilled water. It was also noticed that the melted material removal process is very complicated and hard to define, thus more research in this area is needed. The type of dielectric liquid has a crucial influence on the technological material surface integrity during EDM machining, thus determining the operational properties of the machined element.

References

- [1] Ayesta I., Flaño O., Izquierdo B., Sanchez J.A., Plaza S., *Experimental Study on Debris Evacuation during Slot EDMing*, Procedia CIRP, Vol. 42, 2016, 6–11.
- [2] Banu A., Ali M. Y., *Electrical Discharge Machining (EDM)*, International Journal of Engineering Materials and Manufacture, Vol. 1, 2016, 3–10.
- [3] Boillot P., Peultier J., *Use of stainless steels in the industry: Recent and future developments*, Procedia Engineering, Vol. 83, 2014, 309–321.
- [4] Chakraborty S., Dey V., Ghosh S. K., *A review on the use of dielectric fluids and their effects in electrical discharge machining characteristics*, Precision Engineering, Vol. 40 2015, 1–6.
- [5] Coteață M., Floca A., Dodun O., Ionescu N., Nagiț G., Slătineanu L., *Pulse Generator for Obtaining Surfaces of Small Dimensions by Electrical Discharge Machining*, Procedia CIRP, Vol. 42, 2016, 715–720.
- [6] Dąbrowski L., Świercz R., Zawora J., *Struktura geometryczna powierzchni po obróbce elektroerozyjnej elektrodą grafitową i miedzianą – porównanie*, Inżynieria Maszyn 16, 2011, 32–39.
- [7] Ferraris E., Vleugels J., Guo Y., Bourell D., Kruth J. P., Lauwers B., *Shaping of engineering ceramics by electro, chemical and physical processes*, CIRP Annals – Manufacturing Technology, Vol. 65, 2016, 761–784.
- [8] Goriainov V., Cook R., Latham J.M., Dunlop D.G., Oreffo R.O.C., *Bone and metal: an orthopaedic perspective on osseointegration of metals*, Acta biomaterialia, Vol. 10, 2014, 4043–4057.
- [9] Hayakawa S., Kusafuka Y., Itoigawa F., Nakamura T., *Observation of Material Removal from Discharge Spot in Electrical Discharge Machining*, Procedia CIRP, Vol. 42 2016, 12–17.
- [10] Hollinger J.O., *An Introduction to Biomaterials Second Edition*, Biomedical Engineering, CRC Press, 2011.

- [11] Hourmand M., Sarhan A.A.D., Sayuti M., *Micro-electrode fabrication processes for micro-EDM drilling and milling: a state-of-the-art review*, The International Journal of Advanced Manufacturing Technology, Vol. 2016, 1–34.
- [12] Jha B., Ram K., Rao M., *An overview of technology and research in electrode design and manufacturing in sinking electrical discharge machining*, Journal of Engineering Science and Technology Review, Vol. 2, 2011, 118–130.
- [13] Kitamura T., Kunieda M., Abe K., *High-speed imaging of EDM gap phenomena using transparent electrodes*, Procedia CIRP, Vol. 6, 2013, 314–319.
- [14] Kitamura T., Kunieda M., *Clarification of EDM gap phenomena using transparent electrodes*, CIRP Annals – Manufacturing Technology, Vol. 63, 2014, 213–216.
- [15] Klink A., *Process Signatures of EDM and ECM Processes – Overview from Part Functionality and Surface Modification Point of View*, Procedia CIRP, Vol. 42, 2016, 240–245.
- [16] Koyano T., Hosokawa A., Suzuki S., Ueda T., *Influence of external hydrostatic pressure on machining characteristics of electrical discharge machining*, CIRP Annals – Manufacturing Technology, Vol. 64, 2015, 229–232.
- [17] Kozak J., Rozenek M., Dabrowski L., *Study of electrical discharge machining using powder-suspended working media*, Journal of Engineering Manufacture 217 (11), 2003, 1597–1602.
- [18] Oniszczyk D., Świercz R., *An investigation into the impact of electrical pulse character on surface texture in the EDM and WEDM process*, Advances in Manufacturing Science and Technology, 36.3, 2012, 43–53.
- [19] Rajurkar K.P., Sundaram M.M., Malshe A.P., *Review of Electrochemical and Electrodischarge Machining*, CIRP, Vol. 6, 2013, 13–26.
- [20] Risto M., Haas R., Munz M., *Optimization of the EDM Drilling Process to Increase the Productivity and Geometrical Accuracy*, Procedia CIRP, Vol. 42, 2016, 537–542.
- [21] Świercz R., *Wpływ charakteru impulsów natężenia prądu i napięcia elektrycznego na strukturę warstwy wierzchniej po obróbce EDM*, Politechnika Warszawska, Warszawa 2013.
- [22] Torres A., Puertas I., Luis C.J., *Modelling of surface finish, electrode wear and material removal rate in electrical discharge machining of hard-to-machine alloys*, Precision Engineering, Vol. 40, 2015, 33–45.
- [23] Tripathy S., Tripathy D.K., *Optimization of Process Parameters and Investigation on Surface Characteristics During EDM and Powder Mixed EDM*, Springer Singapore, Vol. 2017, 385–391.
- [24] Wang X., Xu S., Zhou S., Xu W., Leary M., Choong P., Qian M., Brandt M., Xie Y.M., *Topological design and additive manufacturing of porous metals for bone scaffolds and orthopaedic implants: A review*, Biomaterials, Vol. 83, 2016, 127–141.
- [25] Zhang M., Zhang Q., Zhu G., Liu Q., Zhang J., *Effects of Some Process Parameters on the Impulse Force in Single Pulsed EDM*, Procedia CIRP, Vol. 42, 2016, 627–631.
- [26] Zhang Y., Liu Y., Shen Y., Ji R., Li Z., Zheng C., *Investigation on the influence of the dielectrics on the material removal characteristics of EDM*, Journal of Materials Processing Technology, Vol. 214, 2014, 1052–1061.
- [27] http://www.sodick.jp/product/consumables/fluids/p01_01.html (access: 17.11.2016).
- [28] <http://www.holepop.com/dielectric-fluid> (access: 17.11.2016).

Marcin Majka (marcin.majka@ifj.edu.pl)

The Henryk Niewodniczański Institute of Nuclear Physics Polish Academy of Sciences

Giacomo Gadda

Angelo Taibi

Department of Physics and Earth Sciences, University of Ferrara

Piotr Zieliński

Institute of Physics, Cracow University of Technology

EFFECTS OF AORTIC VALVE DISEASES ON PRESSURE PROFILES IN SELECTED LOCATIONS OF THE HUMAN ARTERIAL SYSTEM

WŁYW CHOROÓB ZASTAWKI AORTALNEJ NA PRZEBIEGI CIŚNIENIA W WYBRANYCH MIEJSCACH UKŁADU TĘTNICZEGO CZŁOWIEKA

Abstract

Aortic valve diseases such as aortic stenosis and aortic regurgitation, are the most frequent valvular heart diseases. The lesions in the valves affect circulation in the whole arterial system. We study the effects with the use of a 1-D model in which an arterial segment transmits a single mode of pulse waves. The appropriate reflection coefficient and the form of the stroke pressure are devised to simulate the function of the healthy and morbid aortic valve. The time dependence of the arterial pressure is predicted at the most important locations of the arterial tree. A remarkable result is that little variations of the reflection coefficient of the valve due to the modelled diseases cause significant changes of the pressure profiles, especially at the ascending aorta, the left brachial artery and in the anterior communicating artery.

Keywords: arterial model, aortic valve diseases, aortic stenosis, aortic regurgitation, transmission line, pressure profiles, Circle of Willis

Streszczenie

Zwężenie zastawki aortalnej (*stenosis ostii arteriosae sinistri*) oraz niedomykalność zastawki aortalnej (*insufficiencia valvulae aortae*) są najczęstszymi chorobami zastawek serca. Zmiany chorobowe zastawki niewątpliwie wpływają na obieg krwi w całym układzie tętniczym. Wpływ tych zmian zbadano przy użyciu jednowymiarowego modelu, w którym każdy segment tętnicy transmituje tylko jeden mod fali tętna. Opracowane odpowiednie wartości współczynnika odbicia zastawki aortalnej oraz kształt ciśnienia generowanego przy skurczu serca symulują realistycznie działanie zastawki zdrowej i chorej. Otrzymane wyniki pozwalają przewidywać przebiegi ciśnienia w najistotniejszych miejscach drzewa tętniczego. Zaskakujący jest fakt, że nawet najmniejsze różnice w współczynniku odbicia od zastawki powodują znaczące zmiany w profilach ciśnienia, a w szczególności w aorcie wstępującej, lewej tętnicy ramiennej oraz w tętnicy łączącej przedniej.

Słowa kluczowe: model tętniczy, chory zastawki aortalnej, zwężenie zastawki aortalnej, niedomykalność zastawki aortalnej, linia transmisyjna, przebiegi ciśnienia, koło Willis'a

1. Introduction

Aortic stenosis (AS) and aortic regurgitation (AR) are the most common valvular heart diseases in developed countries [1-6]. AS is often a gradually progressive degenerative disease which results in the calcification of the aortic valve leaflets [3]. This causes an obstruction to the blood flow from the left ventricle to the aorta which leads to an increase in the left ventricular afterload. In contrast to that AR allows for a diastolic flow of blood from the aorta into the left ventricle. The incidence of clinically significant AR increases with age, typically peaking in the fourth to sixth decade of life [6]. AR may be caused by malfunction of the valve leaflets themselves, by dilatation of the aortic root and annulus.

Hypertension can modify the physical examination findings of AS, particularly in the elderly patient [7]. The characteristic exploratory findings of AS are all influenced by peripheral aortic impedance, wave reflections, post stenotic aortic dilatation, and vessel stiffness. If AS coexists with hypertension, the valvular obstruction and the increased vascular resistance complementarity raises the systolic stress of the doubly loaded left ventricle [8]. Moreover,

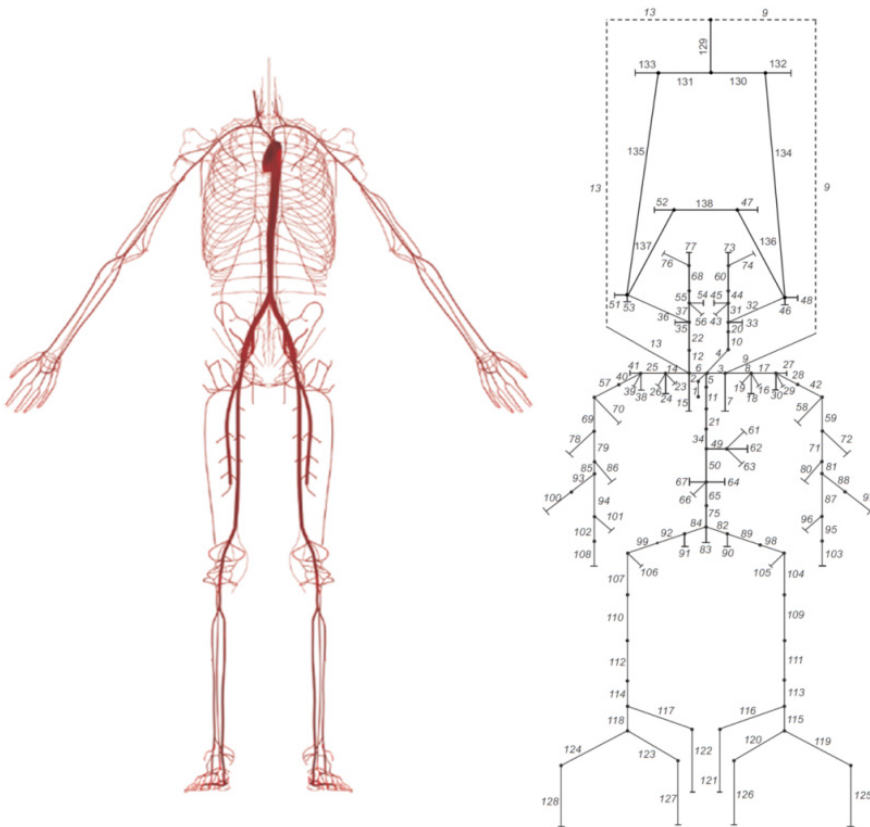


Fig. 1. The arterial tree of the human body. Form right – the topology of model with 138 segments available in [10], From left – real topology

these two sources of afterload interact following a characteristic behaviour. The intrinsic systolic load is competitively modified by the extrinsic load (the systemic blood pressure) in an opposite and parallel fashion [3, 8]. This mechanism explains why vascular biomechanics may also modify the quantification of AS severity. However, the magnitude, clinical relevance, and diagnostic consequences of this vascular-valvar interaction are incompletely understood.

AR causes volume overload of the left ventricle [6]. The total stroke volume ejected by the left ventricle (sum of effective stroke volume plus regurgitant volume) is increased; in severe AR the regurgitant volume may equal or even exceed the effective stroke volume [4, 5]. An increase in the left ventricle end-diastolic volume is the main compensatory mechanism needed to maintain a normal effective stroke volume. The left ventricular ejection fraction is initially normal, however, the left ventricle end-diastolic pressure rises [6]. In time the left ventricle end-diastolic volume continues to increase further and the ejection fraction drops. In AR there is not only a volume overload but also an increase in the afterload and therefore of systolic wall stress. This distinguishes AR from mitral regurgitation [9], where a left ventricle volume overload is also present, but where the systolic wall stress is normal or even lower, since the regurgitant blood is ejected into the low pressure left atrium.

In the present paper we use a 1-D model of the arterial tree to study the effects of AS and AR on the pressure time dependence in peripheral vessels. The model consists of 138 arterial segments transmitting pulse waves.

The topology of the models is presented in Fig. 1.

The purpose of the present study is to give a quantitative account of the effect of AS and AR on the temporal pressure in the arterial system, i.e. to find out how the corresponding disturbances of pressure have an effect on pressure profiles in the ascending aorta, in the left ulnar artery and in the anterior communicating artery. The problem is interesting from the fundamental research point of view, namely to what extent passive models, i.e. those without voltage and/or current sources are capable of reproducing the physiology of the arterial system (for models of neural heart rate control via baroreceptors and reflex arc see e.g. [11–13]). The adequacy of the present computations can be partly verified with the measurement results from patients with AS and AR available in the literature [1–9].

2. Materials and methods

2.1. Propagation of pulse waves

The elasticity of vessel walls is at the origin of the wave-like character of blood flow in arteries. In principle the number of propagation modes is infinite [15, 16], but the most significant effect belongs to the Young mode easily palpable at the wrist. Restricting the attention to this mode allows one to treat the vessels as monomode 1-D waveguides.

The method of obtaining a 1-D equation of motion in a distensible vessel of varying cross section has been developed in [14, 17, 18]. The governing equations involve conservation of mass and the momentum balance in a control volume of the 1-D vessel [17, 18]. The



volumetric flow, as a function of space x and time t , $Q(x,t) = AU$ relates the cross section area A and the average axial velocity U tethered in longitudinal direction. The pressure P is assumed constant across the section, whereas the radial and azimuthal components of velocity are neglected. Gravitational effects are ignored because all the simulations performed in this work refer to a supine subject. The system of equations is:

$$\begin{cases} \frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = 0 \\ \frac{\partial Q}{\partial t} + \frac{\partial}{\partial x} \left(\alpha \frac{Q^2}{A} \right) + \frac{A}{\rho} \frac{\partial P}{\partial x} = \frac{f}{\rho}, \end{cases} \quad (1)$$

where

ρ – the blood density,
 $f(x,t)$ – the frictional force per unit length,

$\alpha(x,t) = \frac{1}{AU^2} \int_A u^2 d\sigma$ – the non-dimensional profile shape factor (also called the Coriolis coefficient) that relates the nonlinear convection acceleration term to the velocity profile $\mathbf{u} = (\mathbf{x}, t)$ [19].

The blood density ρ and viscosity μ are assumed to be constant at 37°C, so that $\rho = 1050 \text{ kg m}^{-3}$ and $\mu = 4.0 \text{ mPa s}$ [14]. Eq. (1) can be derived by integrating the incompressible Navier-Stokes equations over a generic cross section of a cylindrical domain [19–24].

In 1-D modelling the velocity profile is commonly assumed to be constant in shape and axisymmetric. A typical profile satisfying the no-slip condition ($u|_{r=R} = 0$) is [19, 20]:

$$u(x, r, t) = U \frac{\xi + 2}{\xi} \left[1 - \left(\frac{r}{R} \right)^\xi \right] \quad (2)$$

where

r – the radial coordinate,
 $R(x,t)$ – the radius of the lumen (assumed to be circular),

$\xi = \frac{2-\alpha}{\alpha-1}$ – a constant.

The velocity profile is required to close the system of equations (1), since it directly affects convective accelerations and the frictional term $f(x, t)$. For $\xi = 2$ the Coriolis coefficient is equal to 4/3 and it leads to the Poiseuille parabolic velocity profile and the corresponding frictional force equal to $f = -8\mu\pi U$. This limit is valid for the narrowest resistive vessels. On the other hand, when considering the largest vessels one should consider flattened velocity profile as has been shown by Womersley [25]. For this reason we use the value $\xi = 9$ and, appropriately, $\alpha = 1.1$.

An explicit algebraic relationship between P and A (the tube law) is also required to close Eq. (1). The tube law is determined by the viscoelastic properties of the vessel walls [26]. Voigt-type visco-elastic laws reproduce, in the first approximation, the main effects of the

walls' properties on the blood flow in large arteries, including hysteresis and creep [27–32]. An example of this type of law that neglects the effects of wall inertia and longitudinal pre-stress [36] is given by [32]:

$$P = P_e(A; x) + \frac{\Gamma(x)}{A_0(x)\sqrt{A}} \frac{\partial A}{\partial t} \quad (3)$$

$$P_e(A; x) = P_{ext} + \frac{\beta(x)}{A_0(x)} \left(\sqrt{A} - \sqrt{A_0(x)} \right),$$

$$\beta(x) = \frac{4}{3} \sqrt{\pi} E(x) h(x), \quad \Gamma(x) = \frac{2}{3} \sqrt{\pi} \phi h(x), \quad (4)$$

where

- P_e – the elastic component of pressure,
- $h(x)$ – the wall thickness,
- $\beta(x)$ – a value related to the wall elasticity, independent of the transmural pressure,
- $\Gamma(x)$ – a value related to the wall viscosity, independent of the transmural pressure,
- $E(x)$ – the Young's modulus of the vessel wall,
- $\phi(x)$ – the dynamic viscosity of the vessel wall.

The reference area $A_0(x)$ is the vessel area when $P = P_{ext}$ and $\partial A / \partial t = 0$, which are typical initial conditions for numerical analysis. Therefore, the local cross-sectional area $A(x, t)$ will depend on the shape of the artery given by $A_0(x)$ and the mechanical properties of the wall, which may change with x . For example, the arterial wall becomes stiffer as the distance from the heart increases.

In the 1-D formulation the nodes connecting the arterial segments are treated as discontinuities, which is consistent with the long-wavelength approximation. Detailed 3-D calculation of flow at arterial bifurcations show that the flow is generally very complex with the possibility of transient separation and the development of secondary flows [14]. Most of these flow features are confined to the region near the bifurcation and the long wave approximation allows one to neglect their effects on the pulse wave in the 1-D formulation.

The linearization of the governing equations yields an analogy with a transmission line: where resistance R , inductance L , and capacitance C are calculated per unit length of vessel as follows:

$$R = \frac{2(\xi + 2)\pi\mu}{A_0^2}, \quad L = \frac{\rho}{A_0}, \quad C = \frac{2A_0^{\frac{3}{2}}}{\beta}. \quad (5)$$

Eq. (5) allows us to determine the wave speed $c = \sqrt{1/CL}$ and the characteristic impedance $Z = \rho c / A_0$ for every segment of a given artery [14]. These quantities determine the transmission and reflection coefficients at every bifurcation of the arterial tree.

2.2. Models of the arterial tree

We solve the nonlinear 1-D equations (1) and (3) using finite element methods, such as Galerkin [33] and Taylor-Galerkin (combined with operator splitting techniques) [14] schemes. The linearized system of governing equations yields an analytical solution for wave reflection and transmission where the physical properties of the arteries change. At a splitting and merging junction, the reflection coefficients R_f^a, R_f^b, R_f^c for wave propagating in the parent a and two daughter, vessels b and c respectively can be defined as the ratio of the pressure amplitude in the reflected wave to the pressure amplitude in the incident wave. They can be expressed as a function of the characteristic impedance of each segment:

$$R_f^a = \frac{(Z_0^a)^{-1} - (Z_0^b)^{-1} - (Z_0^c)^{-1}}{(Z_0^a)^{-1} + (Z_0^b)^{-1} + (Z_0^c)^{-1}}. \quad (6)$$

The transmission coefficients T_a, T_b, T_c for waves propagating in the parent, a , and two daughter, b and c , vessels can be defined as the ratio of the pressure perturbation transmitted to the other two vessels to the pressure perturbation in the vessel where the initial wave is propagated:

$$T^j = 1 + R_f^j, \quad j = a, b, c. \quad (7)$$

In any system, the reflection and transmission coefficients are related with a mismatch of characteristic impedance. When applied to the triple junctions the Kirchhoffs laws for a 1D conductor provide Eqs. (6) and (7).

The outlet of each terminal branch is assumed to be coupled to a single resistance R_1 [34], so that the corresponding reflection coefficient is:

$$R_f = \frac{R_1 - Z_0}{R_1 + Z_0}. \quad (8)$$

The coefficient relates the pressure change associated with the reflected wave to that of the incident wave.

Putting $R_1 = 0$ yields spurious wave reflections and $R_1 = Z_0$ corresponds to complete absorption of any incoming wave [34], where Z_0 is the characteristic impedance of the peripheral segment. We use the latter assumption in all the terminal branches of the model. At the aortic valve we define a time-dependent reflection coefficient, R_v . We have $R_v > 1$ (more pressure is reflected than it comes in) when the aortic valve is open and $R_v = 1$ (total reflection) when it is closed [35].

2.3. Model of aortic stenosis

We consider an AS as a change of the aortic valve reflection coefficient (R_v). The AS occurs, in our model, abruptly during normal work of the heart. In each simulation series the

occlusion starts after 10 cycles of the heart. We selected 3 different segments of the arterial tree, to present the reaction of the pressure waveforms to the AS. This model allows us to see how the pressure values evolve with time at any desired location of the arterial tree. We present the results in ascending aorta, in left brachial artery and in the anterior communicating artery.

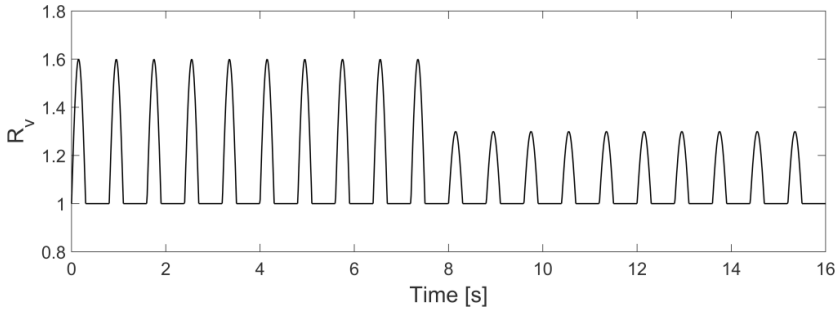


Fig. 2. Valve reflection coefficient during the AS as a function of time

In the normal work the reflection coefficient R_v of the aortic valve, i.e. at the heart point of the arterial tree, is equal to 1.6 during the systole, and 1 during the diastole [35]. A stenosis provokes two simultaneous effects: i) a decrease in the stroke volume and ii) the corresponding decrease in the effective reflection coefficient at the root of our model, i.e. heart. Our computations assume a proportional decrease in both quantities. Accordingly the reflection coefficient corresponding to systole now amounts to $R_v = 1.3$. The reflection coefficient remains equal to 1 in the stage of diastole. Fig. 2 shows the modelled reflection coefficient R_v at the heart point as a function of time; the first 10 cardiac cycles represent a healthy heart, and the stenosis (AS) occurs suddenly at the 11th stroke.

2.4. Model of aortic regurgitation

We consider an AR as a change of the aortic valve reflection coefficient (R_v). The AR occurs, in our model, during normal work of the heart. In each simulation series the occlusion starts after 10 cycles of the heart.

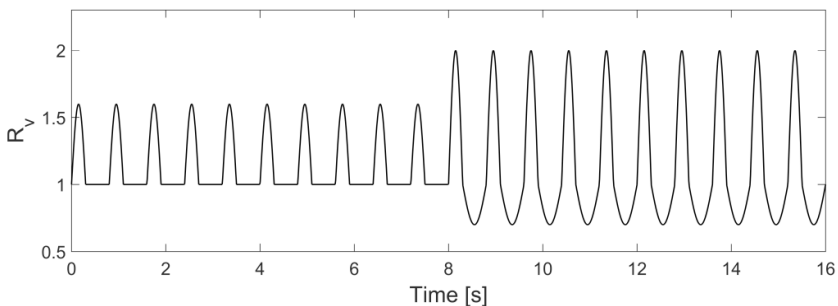


Fig. 3. Valve reflection coefficient during the AR as a function of time

In the normal work during the systole in the crucial point R_v is equal to 1.6, and during the diastole R_v is equal to 1 [35]. As far as the AR is concerned, one knows that the stroke volume increases in the systole, but the valve remains partly opened in the diastole. Accordingly, in our computation, the reflection coefficient during systole is increased to $R_v = 2$, but it drops below 1, i.e. $R_v = 0.7$ in diastole. The changes in the reflection coefficient are proportional to the changes in the stroke volume. Fig. 3 shows the R_v as a function of time. After 10 cardiac cycles the AR occurs.

3. Results

3.1. Aortic stenosis

Fig. 4a shows the pressure waveform in the segment 1 (Ascending aorta in model of 138 arteries, see Fig. 1), in response to the standard heart action given in inset of Fig. 4a.

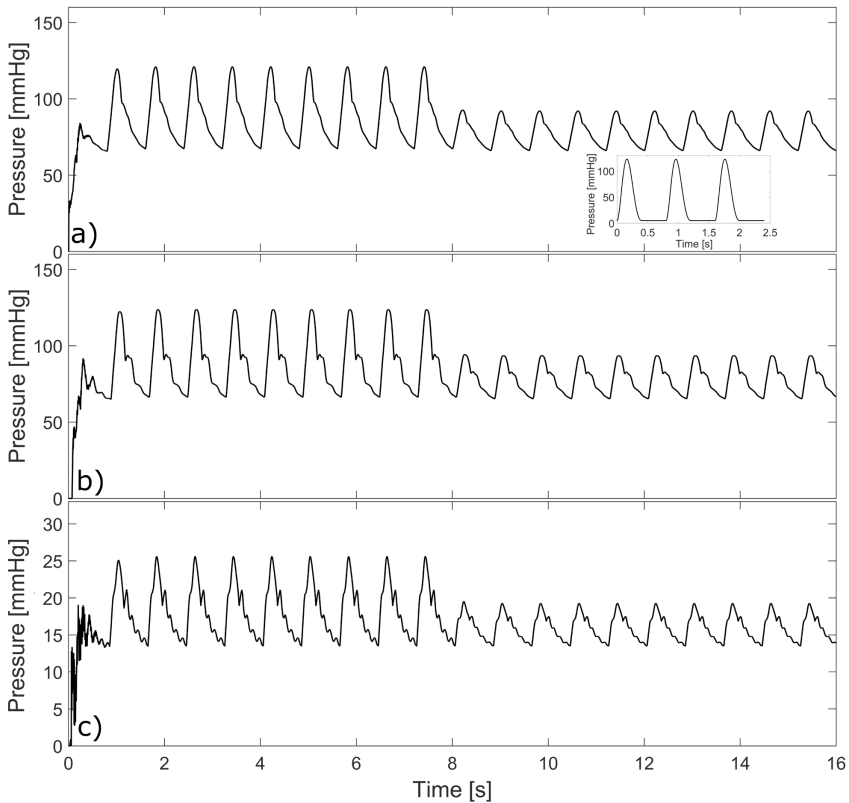


Fig. 4. The pressure waveform in a) ascending aorta, b) left brachial artery and in c) anterior communicating artery during 20 cycles of the heart at AS. The inset in Fig. 4a represents the ventricular pressure waveform in normal condition. The onset of AS takes place after 10 first cycles of heart

The pressure profile stabilizes after about 2 cardiac cycles. Fig. 4b shows the pressure waveform in the segment 59 (Left brachial artery in model of 138 arteries, see Fig. 1). Fig. 4c shows the pressure waveform in the segment 138 (Anterior communicating artery in model of 138 arteries, see Fig. 1). After 10 cardiac cycles the AS occurs.

As we can see from Fig. 4a and Fig. 4b, the AS causes a huge decrease in the pressure in ascending aorta and in left brachial artery, where the systolic pressure decreases by more than 30 mmHg. The diastolic pressure has not changed. In Fig. 4c we can see a pressure waveform in the anterior communicating artery, where AS causes also a decrease in the systolic pressure. The drop is surprisingly low, however, i.e. by 5 mmHg only. If the pressure dropped to 0 mmHg, the blood flow in the anterior communicating artery would be also equal to 0. This would amount to a cerebral hypoxia. However, because of the geometry of the arterial tree, and the rheological properties of the vessels, the pressure inside the anterior communicating artery remains still nonzero during systole. This is an example of the protective properties of the arterial tree for brain [36]. The diastolic pressure in this case has not changed similarly ascending aorta and in left brachial artery. The obtained waveforms are in agreement with experimental data [1-3, 7].

3.2. Aortic regurgitation

Fig. 5a shows the pressure waveform in the segment 1 (Ascending aorta in model of 138 arteries, see Fig. 1), in response to the standard heart action given in inset of Fig. 4a. The pressure profile stabilizes after about 2 cardiac cycles. Fig. 5b shows the pressure waveform in the segment 59 (Left brachial artery in model of 138 arteries, see Fig. 1). Fig. 5c shows the pressure waveform in the segment 138 (Anterior communicating artery in model of 138 arteries, see Fig. 1). After 10 cardiac cycles the AR occurs.

As we can see from Fig. 5a and Fig. 5b, the AR causes a huge increase in the systolic pressure and also a decrease in the diastolic pressure in the ascending aorta and the left brachial artery, where the systolic pressure increases by more than 40 mmHg and the diastolic pressure decreases by 25 mmHg. In Fig. 5c we can see a pressure waveform in the anterior communicating artery, where AR causes also an increase in systolic pressure but only by 8 mmHg and decrease in diastolic pressure only by 4 mmHg. The waveforms obtained are in agreement with experimental data [4, 6, 9].

The results show the theoretical behaviour of the pressure waveforms at the major vessels of the human arterial system in cases of AS and AR. As has been shown, all the data obtained from the model are in qualitative agreement with experimental data reported in [1-4, 6, 7, 9]. The typical values of the pressure (systole/diastole) in the ascending aorta for AS are ~110/75 mmHg [1-3], and for AR are ~150/60 mmHg [4-6]. A comparison of these clinical results to results from our models shows a really good agreement. However, our model of the arterial tree reproduces only a mechanical response of the arterial tree. The compensatory mechanisms, e.g. baroreflex or/and humoral responses are still to be studied. Therefore, our results should be treated as a plausible prediction of the behaviour of the arterial system in the very first stages after the events and, on the other hand, as a passive-mechanical reference

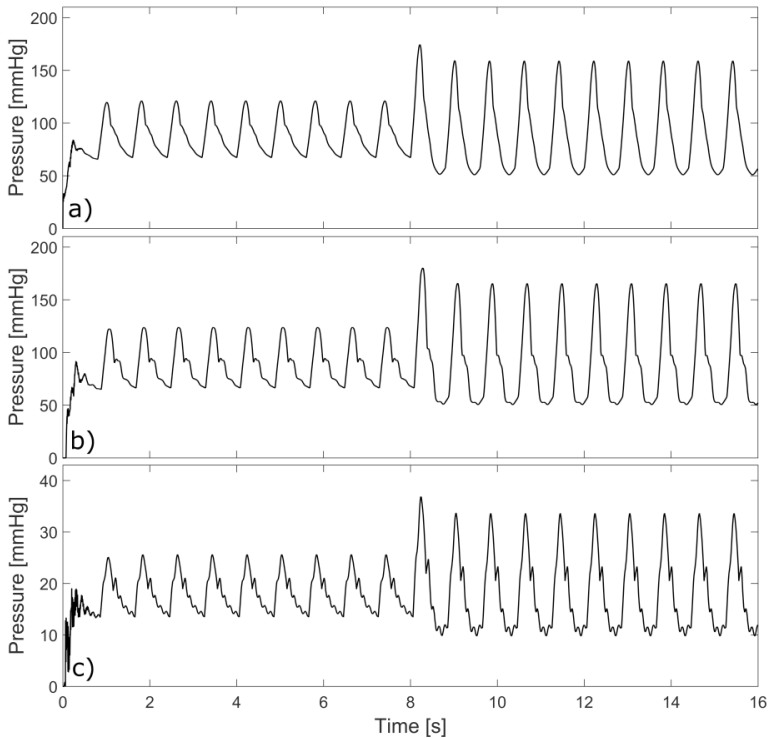


Fig. 5. The pressure waveform in a) ascending aorta, b) left brachial artery and in c) anterior communicating artery during 20 cycles of the heart at AR. The onset of AR takes place after 10 first cycles of heart

for further studies including the compensatory mechanisms that are intrinsically active and/or of a feedback nature.

These models are the first attempts of studies of AS and AR with the use of a 1-D model of the arterial tree. The simulation data explain the physical consequences of AS and AR. The abrupt switching of the morbid reflection coefficients illustrates the earliest reaction of the pressure profiles on the AS and AR in the selected locations of the human systemic arterial system.

4. Conclusions

We have constructed a 1-D model of the arterial system that is able to reliably reproduce the most fundamental features of the propagation of the pulse waves [36, 37]. This kind of model is particularly effective in the description of large-scale properties of the entire arterial system and of its response to some lesions and/or external hazards. Here we have studied the reaction of the arterial system to perturbations caused by AS and AR. This phenomena were simulated in Avolio's model of 138 arteries with the Circle of Willis in it [10]. This new approach allowed us, for the first time, to determine the pressure waveforms during the AS and AR.

To simulate the valvular heart diseases, the changes of the reflection coefficient of the aortic valve has been used. This simple approach allowed us to determine the earliest reactions of the arterial system to AS and AR.

During the AS we can see a significant decrease in the pressure waveform in the ascending aorta, in the left brachial artery and in the anterior communicating artery. However, the decrease is only in the systolic pressure. The diastolic pressure has not changed. In the case of AR we can see changes in amplitude of systolic and diastolic pressure. When, because of AR, the systolic pressure increases at the same time the diastolic pressure decreases. These changes match the experimental observations of the phenomena.

The model can be used to predict the response of the human arterial system in many other physiological and pathological conditions.

References

- [1] Chambers J.B., *Aortic stenosis*, European Journal of Echocardiography, Vol. 10, 2009, 11–19.
- [2] Baumgartner H., Hung J., Bermejo J., Chambers J.B., Evangelista A., Griffin B.P., Jung B., Otto C.M., Pellikka P.A., Quinones M., *Echocardiographic assessment of valve stenosis: Eae/ase recommendations for clinical practice*, European Journal of Echocardiography, Vol. 10, 2009, 1–25.
- [3] Bermejo J., *The effects of hypertension on aortic valve stenosis*, Heart, Vol. 91, 2005, 280–282.
- [4] Schade R., Andersohn F., Suissa S., Haverkamp W., Garbe E., *Dopamine agonists and the risk of cardiac-valve regurgitation*, N Engl J Med, Vol. 356, 2007, 29–38.
- [5] Zanettini R., Antonini A., Gatto G., Gentile R., Tesei S., Pezzoli G., *Valvular heart disease and the use of dopamine agonists for parkinson's disease*, N Engl J Med., Vol. 356, 2007, 39–46.
- [6] Maurer G., *Aortic regurgitation*, Heart, Vol. 92, 2006, 994–1000.
- [7] Pocock C., Chambers J., *The patient with a systolic murmur: severe aortic stenosis may be missed during cardiovascular examination*, QJM, Vol. 93, 2000, 685–688.
- [8] Pasipoularides A., *Clinical assessment of ventricular ejection dynamics with and without outflow obstruction*, J Am Coll Cardiol, Vol. 15, 1990, 859–882.
- [9] Wisenbaugh T., Spann J. F., Carabello B. A., *Differences in myocardial performance and load between patients with similar amounts of chronic aortic versus chronic mitral regurgitation*, J Am Coll Cardiol, Vol. 3, 1984, 916–923.
- [10] Schwarz M., Nguyen M., Kiencke U., Heilmann C., Klemm R., Benk C., Beyersdorf F., Busch H., *Integration of the circle of willis into avolio's model of the arterial haemodynamics*, Proceedings of the Sixth IASTED International Conference Biomedical Engineering, 2008.
- [11] Zamir M., Coverdale N., Barron C., Sawicki C., Shoemaker J., *Baroreflex variability and "resetting": A new perspective*, Journal of Biomechanics, Vol. 47, 2014, 237–244.
- [12] Kember G., Armour J., Zamir M., *Neural control hierarchy of the heart has not evolved to deal with myocardial ischemia*, Physiological Genomics, Vol. 45, 2013, 638–644.

- [13] Ogoh S., Fadel P., Nissen P., Jans O., Selmer C., Secher N., Raven P., *Baroreflex – mediated changes in cardiac output and vascular conductance in response to alterations in carotid sinus pressure during exercise in humans*, J Physiol., Vol. 550, 2003, 317–324.
- [14] Alastruey J., Parker K. H., Sherwin S. J., *Arterial pulse wave haemodynamics*, 11th International Conference on Pressure Surges, 2012, 401–443.
- [15] Jagielska K., Trzuppek D., Lepers M., Pelc A., Zieliński P., *Effect of surrounding tissue on propagation of axisymmetric waves in arteries*, Physical Review E, Vol. 76, 2007, 066304.
- [16] Drochon A., *Sinusoidal flow of blood in a cylindrical deformable vessel exposed to an external magnetic field*, The European Physical Journal Applied Physics, Vol. 73, 2016, 18.
- [17] Sherwin S.J., Peiro J., Parker K.H., *One-dimensional modelling of a vascular network in space-time variables*, J. Eng. Maths., Vol. 47, 2003, 217–250.
- [18] Peiro J., Veneziani A., *Reduced models of the cardiovascular system*, Cardiovascular Mathematics, 2009, 347–394.
- [19] Hughes T., Lubliner J., *On the one-dimensional theory of blood flow in the larger vessels*, Mathematical Biosciences, Vol. 18(1–2), 1973, 161–170.–
- [20] Smith N.P., Pullan A.J., Hunter P.J., *An anatomically based model of transient coronary blood flow in the heart*, SIAM J. Appl. Math., Vol. 62, 2001, 990–1018.
- [21] Hughes T.J.R., *A study of the one-dimensional theory of arterial pulse propagation*, PhD thesis, 1974.
- [22] Vosse F.N. van de, Stergiopoulos N., *Pulse wave propagation in the arterial tree*, Annu. Rev. Fluid Mech., Vol. 43, 2011, 467–499.
- [23] Canic S., Kim E., *Mathematical analysis of the quasilinear effects in a hyperbolic model of blood flow through compliant axi-symmetric vessels*, Math. Meth. Appl. Sci., Vol. 26, 2003, 1161–1186.
- [24] Quarteroni A., Formaggia L., *Mathematical modelling and numerical simulation of the cardiovascular system*, in Handbook of Numerical Analysis 12, Elsevier, 2004, 3–127.
- [25] Womersley J.R., *Method for the calculation of velocity, rate of flow and viscous drag in arteries when the pressure gradient is known*. J Physiol. 127 (3), 1955, 553–563.
- [26] Holzapfel G.A., Gasser T.C., Ogden R.W., *A new constitutive framework for arterial wall mechanics and a comparative study of material models*, J. Elasticity, Vol. 61, 2000, 1–48.
- [27] Armentano R., Barra J., Levenson J., Simon A., Pichel R.H., *Arterial wall mechanics in conscious dogs: Assessment of viscous, inertial and elastic moduli to characterize aortic wall behavior*, Circulation Research, Vol. 76, 1995, 468–478.
- [28] Armentano R., Megnien J.L., Simon A., Bellenfant F., Barra J., Levenson J., *Effects of hypertension on viscoelasticity of carotid and femoral arteries in humans*, Hypertension, Vol. 26, 1995, 48–54.
- [29] Craiem D., Graf S., Pessana F., Grignola J., Bia D., Gines F., Armentano R., *Cardiovascular engineering: modelization of ventricular arterial interaction in systemic and pulmonary circulation*, Latin American Applied Research, Vol. 35, 2005, 111–114.
- [30] Canic S., Tambaca J., Guidobon G., Mikelic A., Hartley C., Rosenstrauch D., *Modeling viscoelastic behavior of arterial walls and their interaction with pulsatile blood flow*, SIAM J. Appl. Math., Vol. 67, 2006, 164–193.

- [31] Saito M., Ikenaga Y., Matsukawa M., Watanabe Y., Asada T., Lagree P.-Y., *One-dimensional model for propagation of a pressure wave in a model of the human arterial network: Comparison of theoretical and experimental results*, J. Biomech. Eng., Vol. 133, 2011, 121005.
- [32] Alastruey J., Khir A., Matthys K.S., Segers P., Sherwin S. J., Verdonck P., Parker K. H., Peiro J., *Pulse wave propagation in a model human arterial network: Assessment of 1-d visco-elastic simulations against in vitro measurements*, J. Biomech., Vol. 44, 2011, 2250–2258.
- [33] Formaggia L., Lamponi D., Quarteroni A., *One-dimensional models for blood flow in arteries*, J. Eng. Math., Vol. 47, 2003, 251–276.
- [34] Alastruey J., Parker K.H., Peiro J., Sherwin S.J., *Lumped parameter outflow models for 1-d blood flow simulations: Effect on pulse waves and parameter estimation*, Commun. Comput. Phys., Vol. 4, 2008, 317–336.
- [35] Alastruey J., Parker K.H., Peiro J., Sherwin S.J., *Analysing the pattern of pulse waves in arterial networks: a time-domain study*, J. Eng. Math., Vol. 64, 2009, 331–351.
- [36] Majka M., Gadda G., Taibi A., Gałazka M., Zieliński P., *Protective properties of the arterial system against peripherally generated waves*, Mathematical Biosciences, Vol. 286, 2017, 16–21.
- [37] Majka M., Gadda G., Taibi A., Gałazka M., Zieliński P., *The earliest effects of sudden occlusions on pressure profiles in selected locations of the human systemic arterial system*, Physical Review E – in press, 2017.



Reviewers of 2017:

Akhlaque Ahmad, Iqra University, Pakistan

Robert Ast, Poznań University of Technology, Poland

Stanislav Avsec, University of Lublana, Slovenia

Krzysztof Badyda, Warsaw University of Technology, Poland

Marek Balazinski, Polytechnique Montréal, Canada

Ioan Balint, Institute of Physical Chemistry "Ilie Murgulescu", Romania

Andrzej Bargieła, Nottingham Science and Technology Park, UK

Wojciech Bartnik, University of Agriculture in Krakow, Poland

Adam Bartnicki, Wojskowa Akademia Techniczna, Poland

Renáta Bašková, Košice University of Technology, Slovakia

Stanisław Belniak, Cracow University of Economics, Poland

Ján Benčat, University of Žilina, Slovakia

Ali Cemal Benim, Duesseldorf University of Applied Sciences, Germany

Marek Berezowski, Silesian University of Technology, Poland

Nikolai Berlinsky, Odessa State Environmental University, Ukraine

Joanna Biedrońska, Silesian University of Technology, Poland

Andrzej Bielski, Cracow University of Technology, Poland

Jadwiga Bizon-Górecka, Cracow University of Economics, Poland

Elżbieta Bociąga, Czestochowa University of Technology, Poland

Monika Bogdanowska, Cracow University of Technology, Poland

Agata Bonenberg, Polytechnic University of Milan, Italy

Dariusz Butrymowicz, Bialystok University of Technology, Poland

Renzo Campostrini, University of Trento, Italy

Claudio Carvajal, National Research Institute of Science and Technology for Environment and Agriculture, France

Maria J. Chmiel, University of Agriculture in Krakow, Poland

Wang Chuang, Tsinghua University, China

Kurt Chudej, University of Bayreuth, Germany

Janusz Chwastowski, Polish Academy of Sciences, Poland

Ewa Cichy-Pazder, Poznań University of Technology, Poland

Małgorzata Cimochoiwicz-Rybicka, Cracow University of Technology, Poland

Vladimír Čablík, Technical University of Ostrava, Czech Republic

Dimitr Czechowicz, Silesian University of Technology, Poland

Marek Czernicki, Lille University of Science and Technology, France

Alfred Czogała, University of Silesia in Katowice, Poland

Marek Czyński, West Pomeranian University of Technology, Poland

Oleg M. Demchuk, Maria Curie-Skłodowska University

Bogdan Demchyna, Bronisław Markiewicz State Higher School of Technology and Economics in Jarosław, Poland

Tomasz Dłotko, University of Silesia in Katowice, Poland

Kajetan d'Obyrn, AGH University of Science and Technology, Poland

Lucyna Domagała, Cracow University of Technology, Poland

Mykhaylo Dorozhovets, Rzeszow University of Technology, Poland
Dorota Downarowicz, West Pomeranian University of Technology Szczecin, Poland
Ewa Dresler, Opole University of Technology, Poland
Agnes Drochon, University of Technology of Compiègne, France
Carlos Marmolejo Duarte, Technical University of Catalonia, Spain
Ludmila Dulebova, Technical University of Kosice, Slovakia
Zbigniew Dziopa, Kielce University of Technology, Poland
Nandor Eber, Hungarian Academy of Sciences, Hungary
Tomasz Falborski, Gdańsk University of Technology, Poland
Gabriel Fedorko, Technical University of Kosice, Slovakia
Michał Fiedler, Poznan University of Life Sciences, Poland
Sergiy Filin, West Pomeranian University of Technology, Poland
Dora Foti, Polytechnic University of Bari, Italy
Ado Franchini, Polytechnic University of Milan, Italy
George Gamov, Ivanovo State University of Chemical Technology, Russia
Krzysztof Gaska, AGH University of Science Technology, Poland
Jarosław Gawdzik, Kielce University of Technology, Poland
Justin Geistefeldt, Ruhr-University Bochum, Germany
Aneta Głuszek, Office of Technical Inspection (UDT), Poland
Grzegorz Ginda, AGH University of Science and Technology, Poland
Jacek Gołaszewski, Silesian University of Technology, Poland
Tomisław Gołębiowski, Cracow University of Technology, Poland
Andrzej Górecki, Jagiellonian University, Poland
Sławomir Grądział, Cracow University of Technology, Poland
Marek Graniczny, Polish Geological Institute, National Research Institute, Poland
Barbara Gronostajska, Wrocław University of Technology, Poland
Thomas Hanak, Esslingen University of Applied Sciences, Germany
Vlasta Harbulakova, Technical University of Kosice, Slovakia
Zdzisław Hejducki, Wrocław University of Technology, Poland
Ryszard Hejmanowski, AGH University of Science and Technology, Poland
Karin Hofert, Technical University of Catalonia, Spain
Bożena Hoła, Wrocław University of Technology, Poland
Šárka Hošková-Mayerová of University of Defence, Brno, Czech Republic
Vít Hromádka, Brno University of Technology, Czech Republic
Ivan Hyben, Technical University of Košice, Slovakia
Robert Idem, Gdańsk University of Technology, Poland
Monika Iwantysynova, Purdue University, USA
Artur Jasiński, Andrzej Frycz Modrzewski Krakow University, Poland
Piotr Jaśkowski, Lublin University of Technology, Poland
Natalia Junakova, Technical University of Kosice, Slovakia
Matevž Juvančič, University of Ljubljana, Slovenia
Elżbieta Kaczmarska, Andrzej Frycz Modrzewski Krakow University, Poland

Rajesh Kanna, Velammal College of Engineering and Technology, India
Daniel Kaplan, University of Pittsburgh, USA
Marek Karkula, AGH University of Science Technology, Poland
Joanna Karwan-Baczewska, AGH University of Science Technology, Poland
Tadeusz Kasprzewicz, Military University of Technology in Warsaw, Poland
Tomas Kauertz, University of Applied Sciences and Arts in Hildesheim, Germany
Ivan Kelnar, Institute of Macromolecular Chemistry, Czech Republic
Vladimir S. Kislik, The Hebrew University of Jerusalem, Israel
Ewa Konduracka, Jagiellonian University, Poland
Petr Konicek, Czech Academy of Sciences. Czech Republic
Jana Korytářová, Brno University of Technology, Czech Republic
Wojciech Kosiński, Andrzej Frycz Modrzewski Krakow University, Poland
Jarosław Kosiorek, University of Warmia and Mazury in Olsztyn, Poland
Magdalena Kozień-Woźniak, Cracow University of Technology, Poland
Adam Krasieński, Gdansk University of Technology, Poland
Peter Kravets, National Technical University of Ukraine, Ukraine
Victor Kravets, National Mining University, Dnipro, Ukraine
Artur Krowiak, Cracow University of Technology, Poland
Irina Kruchinina, Russian Academy of Sciences, Russia
Yuyiy Kryvoruchko, Lviv Polytechnic National University, Ukraine
Marek Książek, Cracow University of Technology, Poland
Stanisław Kuciel, Cracow University of Technology, Poland
Janusz Kulejewski, Warsaw University of Technology, Poland
Ryszard Kutylowski, Wrocław University of Technology, Poland
Krystyna Kuźniar, Pedagogical University in Cracow, Poland
Tone Lerher, University of Maribor, Slovenia
Tomasz Lipecki, Lublin University of Technology, Poland
Zbigniew Lozia, Warsaw University of Technology, Poland
Alena Luptáková, Košice University of Technology, Slovakia
Katarzyna Łakomy, Cracow University of Technology, Poland
Aneta Łukaszek-Chmielewska, Main School of Fire Service, Poland
Andrzej Machocki, Maria Curie-Skłodowska University, Poland
Maciej Maciejewski, Institute of Meteorology and Water Management, National Research Institute, Poland
Daniela Mačková, Košice University of Technology, Slovakia
Ildiko Mankova, Technical University of Kosice, Slovakia
Henryk Marczak, AGH University of Science and Technology, Poland
Roman Marcinkowski, Warsaw University of Technology, Poland
Roman Marszałek, University of Opole, Poland
Tomasz Martyński, Poznan University of Technology, Poland
Marco Menegotto, Sapienza University of Rome, Italy
Wiesław Meszek, Poznan University of Technology, Poland
Peter Meszaros, Particle and Gravitational Astrophysics, USA

Czesław Miedziałowski, Białystok University of Technology, Poland
Moghtada Mobedi, Shizuoka University, Japan
Maciej Motak, Cracow University of Technology, Poland
Jacek Mucha, Rzeszow University of Technology, Poland
Zbigniew Myczkowski, Cracow University of Technology, Poland
Pavel Neuberger, Czech University of Life Science Prague, Czech Republic
Kerim Emre Öksüz, Cumhuriyet University, Turkey
Marek Pabich, Lodz University of Technology, Poland
Marian Paluch, University of Silesia in Katowice, Poland
Jerzy Paślowski, Poznan University of Technology, Poland
Zbigniew Paszkowski, West Pomeranian University of Technology, Poland
Hartmut Pasternak, Brandenburg University of Technology Cottbus – Senftenberg, Germany
Maciej Bonawentura Pawlicki, Podhale State College of Applied Sciences in Nowy Targ, Poland
Małgorzata Pawłowska, Lublin University of Technology, Poland
Ettore Pennestri, Università di Roma, Italy
Ludek Pesek, The Czech Academy of Sciences, Czech Republic
Michael Peterek, Frankfurt University of Applied Sciences, Germany
Matthias Pietsch, Anhalt University of Applied Sciences, Germany
Stefan Plaum, RheinMain University of Applied Sciences, Germany
Edyta Plebankiewicz, Cracow University of Technology, Poland
Marianna Polak, University of Waterloo, Canada
Stanislav Pospíšil, Université de Montréal, Canada
Eduardo Poverbio, University of Messina, Italy
Marek Pronobis, Silesian University of Technology, Poland
Paweł Ptaszek, University of Agriculture in Krakow, Poland
Ladislav Pust, The Czech Academy of Sciences, Czech Republic
Jan Radoń, University of Agriculture in Krakow, Poland
Matthias Rädle, Universität Heidelberg, Germany
Maciej Radziejewski, Adam Mickiewicz University in Poznań, Poland
Elżbieta Radziszewska-Zielina, Cracow University of Technology, Poland
Maria Richert, AGH University of Science Technology, Poland
Zofia Sadecka, University of Zielona Góra, Poland
Sarah Sadykova, Eurasian National University – Astana, Kazakhstan
Ivor Samuels, Birmingham University, UK
Krzysztof Schabowicz, Wrocław University of Technology, Poland
Olga Nikolayevna Semeniuk, Eurasian National University – Astana, Kazakhstan
Oleksandr Shashenko, National Mining University, Dnipro, Ukraine
Alena Sičáková, Košice University of Technology, Slovakia
Dimitri Sidorov, California State University, Long Beach, USA
Krzysztof Skabek, Cracow University of Technology, Poland
Mirosław Skibniewski, University of Maryland, USA
Anna Sobotka, AGH University of Science and Technology, Poland

Adolf Sotoca, Technical University of Catalonia, Spain
Sławomir Spadło, Kielce University of Technology, Poland
Marcela Spišáková, Košice University of Technology, Slovakia
Michael Sprysch, University of Applied Sciences and Arts Hildesheim/Holzminen/Goettingen, Germany
Anton Sroka, Freiberg University of Mining and Technology, Germany
Krystyna Stec, Central Mining Institute, Poland
Zuzana Struková, Košice University of Technology, Slovakia
Wahyudin P. Syam, The University of Nottingham, UK
Ludwik Synoradzki, Warsaw University of Technology, Poland
Anna Szlachta, Rzeszow University of Technology, Poland
Tomasz Szmuc, AGH University of Science and Technology, Poland
Radomir Ščurek, VŠB-Technical University of Ostrava, Czech Republic
Maciej Szumigala, Poznan University of Technology, Poland
Waldemar Świdziński, Polish Academy of Sciences, Poland
Zdzisław Tabor, Cracow University of Technology, Poland
Dawid Taler, Cracow University of Technology, Poland
Jolanta Tamošaitiene, Vilnius Gediminas Technical University, Lithuania
Marian Taniowski, Silesian University of Technology, Poland
Tadeusz Tatara, Cracow University of Technology, Poland
Alena Tážiková, Košice University of Technology, Slovakia
Piotr Tomasik, Cracow School of Health Promotion, Poland
Grzegorz Tora, Cracow University of Technology, Poland
M. Uthayakumar, Kalasalingam University, Madura
Zbigniew Trzeciakiewicz, Silesian University of Technology, Poland
Ivan Vaníček, Czech Technical University Prague, Czech Republic
Andrea Vallati, Universita di Roma, Italy
Ingrida Veliute, Vilnius Academy of Arts, Lithuania
Martins Vilnitis, Riga Technical University, Latvia
Tatiana Vilutiene, Vilnius Gediminas Technical University, Lithuania
Miloslav Výskala, Brno University of Technology, Czech Republic
Andrzej Wałęga, University of Agriculture in Krakow, Poland
Lipo Wang, Nanyang Technological University, Singapore
Ewa Waś, Research Institute of Horticulture Apicultural Division in Pulawy, Poland
Mirosław Wciślik, Kielce University of Technology, Poland
Harald Wedwitschka, DBFZ Deutsches Biomasse Forschungs Zentrum, Germany
Stanisława Wehle-Strzelecka, Kielce University of Technology, Poland
Tomasz Węclawowicz, Andrzej Frycz Modrzewski Krakow University, Poland
Ewa Węclawowicz-Gyurkovich, Cracow University of Technology, Poland
Bohdan Węglowski, Cracow University of Technology, Poland
Jerzy Wiciak, AGH University of Science Technology, Poland
Andrzej Więckowski, AGH University of Science and Technology, Poland
Joanna Wilk, Rzeszow University of Technology, Poland

Adam Wysokowski, University of Zielona Góra, Poland
Andrzej Wyżykowski, Andrzej Frycz Modrzewski Krakow University, Poland
Małgorzata Wzorek, Opole University of Technology, Poland
Ersel Yilmaz, Adnan Menderes University, Turkey
Esam George Yousif, Technical Collage/Kirkuk, Iraq
Artur Zaguła, Lodz University of Technology, Poland
Bogusław Zakrzewski, West Pomeranian University of Technology, Poland
Daniel Załuski, Gdańsk University of Technology, Poland
Zbigniew Zembaty, Opole University of Technology, Poland
Łukasz Zieliński, Opole University of Technology, Poland
Wiesław Zwierzycki, Poznan University of Technology, Poland