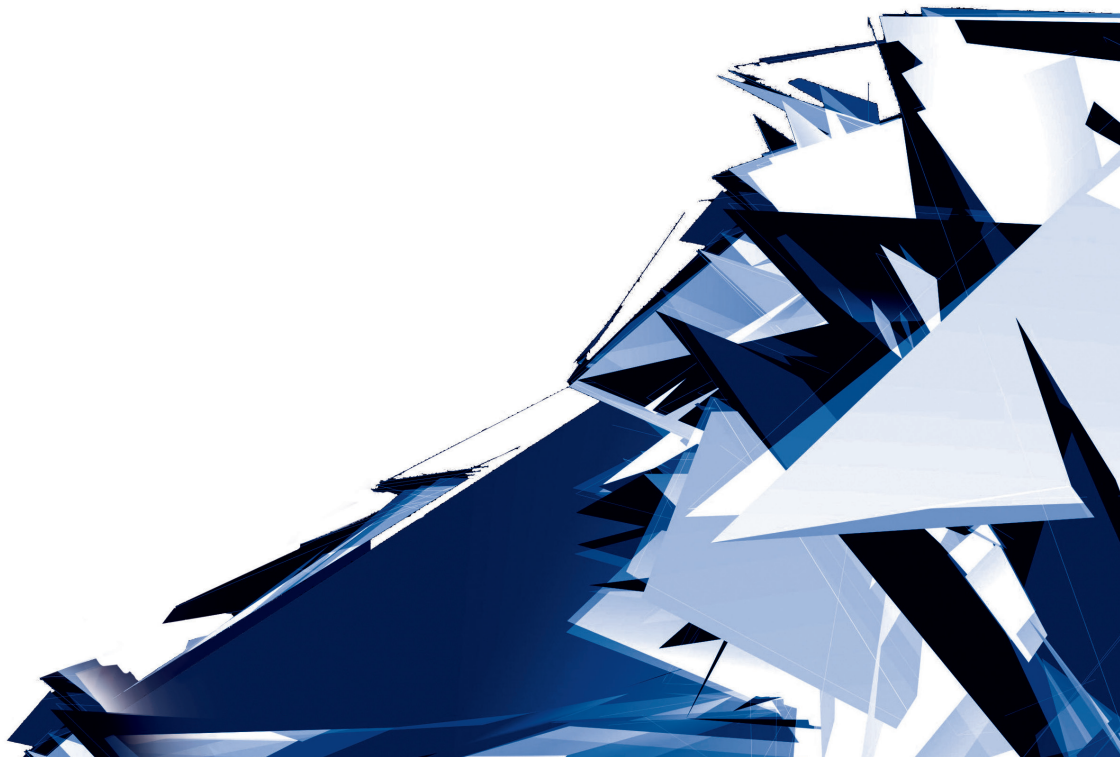
Technical Transactions

Czasopismo Techniczne

Issue 3
Volume 2018 (115)



ISSN 0011-4561

Basic version of each Technical Transactions magazine is its online version
Pierwotną wersją każdego zeszytu Czasopisma Technicznego jest jego wersja online

Contents

MECHANICS

TECHNICAL TRANSACTIONS 3/2018
CZASOPISMO TECHNICZNE 3/2018
ARCHITECTURE AND URBAN PLANNING
DOI: 10.4467/2353737XCT.18.033.8295
SUBMISSION OF THE FINAL VERSION: 26/2/2018

Tingting Chen
Department of Housing and Urban – Rural, School of Urban Design, Wuhan University
Justyna Kobylarczyk
Michał Krupa
Dominika Kuśnierz-Krupa
Faculty of Architecture, Cracow University of Technology

The complexity of a small town

Wielowątkowość małego miasta

Abstract

The article presents deliberations about a small city as a desirable living environment of a multi-layered character focusing on sociological issues that take into account the lifestyle of the residents and their preferences. Environmental studies, which reflected the expectations of the residents regarding their living environment, proved to be very important for the matter of the paper. Another issue related to small cities is the architecture responsible for the vitality of the place, its character and intimacy. It often determines the attractiveness of the centre, its history and culture. The considerations provided by the authors were supported by urban analyses, interviews, literature studies and own observations regarding selected urban centres located in the Podkarpackie Voivodeship.

Keywords: small cities, architecture, spatial arrangement

Streszczenie

W artykule przedstawiono rozważania na temat małego miasta jako pożądanego środowiska zamieszkania o wielowątkowym charakterze skupiającym zagadnienia socjologiczne uwzględniające styl życia mieszkańców oraz ich preferencje. Dla tej części rozważań istotne okazały się badania środowiskowe, których wyniki odzwierciedliły oczekiwania mieszkańców względem ich środowiska zamieszkania. Drugi wątek towarzyszący małym miastom to architektura odpowiadająca za żywotność miejsca, jego nastrój oraz kameralność. Niejednokrotnie decyduje o atrakcyjności ośrodka, jego historii i kulturze. Wskazane rozważania podparte zostały analizami urbanistycznymi, wywiadami, studiami literatury oraz własnymi spostrzeżeniami dotyczącymi wybranych miast zlokalizowanych w województwie podkarpackim.

Słowa kluczowe: małe miasto, architektura, układ przestrzenny

1. Introduction

Nowadays, small towns are becoming more and more popular. While some people find a desirable peace and cosy atmosphere in such places, others still cannot imagine living there. The extreme emotions that are evoked by such a town result from its multi-level character and specificity that reveals contradictory demands of people directed towards the environment they inhabit.

Modern man feels lost in the world filled with technological novelties, in which anonymity, the pace of life, demands and expectations increase daily, and are so mercilessly exposed in a large urban centre. Despite the feeling of constant threat resulting from the rapid development of cities, their inhabitants frequently feel dependent on them. In such moments, people need some respite, intimacy, tranquillity, a possibility of communing with nature, which can be provided by a small town. Unfortunately, such small urban centres are frequently not able to meet the expectations of their residents, usually those on the basic level, such as availability of employment.

Small towns can be perceived in various ways because of their multi-level character co-creating their complete image. And so, among others, a sociological level can be distinguished – the most significant, since it is focused around the image of a town created by man – his lifestyle, preferences, ways of communicating, interacting with others. Man shapes the culture of a place, strengthens its identity, and builds up the awareness of a place. Because of those values, a town has an opportunity to develop and be perceived as a unique place.

The second level is the morphology of a town with its important elements – architectonic objects and spaces, which can create a picture of a degraded, forgotten town [1], or a place that is alive particularly by means of public spaces, such as squares, streets, lanes, pathways, parks, quiet or loud, atmospheric or typical. The architecture of the place and its interior can be pleasant and friendly, or can evoke the feeling of fear or being lost as well as uncertainty.

The third level includes the morphological network of the town – its pathways, roads and passages connecting essential points and larger spaces inscribed within the town structure.

The efficiency and safety of those connections frequently become a priority because of the growing intensity of traffic, slowed down commuting and increasing air pollution caused, among other things, by excessive use of individual means of transport (car).

The last level seems to be a compilation of all the one mentioned above, yet perceived from a different perspective. It allows for interpreting a town as a space in which the existing architecture becomes an event for us – its viewers. The social dimension is forcibly present on this level. Architecture, urban space and man constitute a complete image of a town and are its inextricable elements. The townscape is the background for architecture, and the latter is a static point, which can be put into action by man. From this perspective, the town's atmosphere, dynamics of changes, size, scale and proportions – i.e. elements shaping the composition and physiognomy of the urban space - are not without significance [2].

The sociological level that allows for seeing the town as a set of places where people with their faces, names and identity meet, and spend their leisure time together, seems provided to a greater extent in a small town than in a large city, despite activities of cultural character offered by the latter. Usually, such activities are available in public spaces, mostly squares, including

main market squares – thus open interiors accessible for everybody. Considering the number of residents which is higher in cities than towns, it should be emphasised that participants of such events in larger urban centres remain anonymous which unfortunately limits the possibilities of strengthening social relations. Therefore, the function is largely taken over by social spaces used by a definite group of residents, frequently concentrated around a housing estate. The size and the manner in which such spaces are furnished should be adjusted to the number of users and their preferences. It seems that a city and its fragments are human-friendly if they meet social requirements and follow their expectations, which is influenced by the morphology of a town with its spaces and architecture. Partially, it determines possibilities concerning the availability of specific places and services which belong to the primary group of the so-called living ones, and the secondary some of which can be treated as urban attractions.

The network of morphological conditionings addresses the issue of mobility and accessibility developed by the idea of a compact town. The idea focuses on the issue of healthy public transport generated by pedestrians and cyclists. Although the awareness of the society sensitised to the problems caused by the growing intensity of car traffic is increasing considerably, yet the manner of arranging public spaces in Polish cities still does not sufficiently ensure the safety of fans of bicycle transport. A town, even though they less frequently possess bicycle paths or footpaths on both sides of streets, is a much safer place to live in this respect, which is caused by restricted traffic.

Urban space and its architecture create a complete image of a town with all its assets and shortcomings. Assets of towns usually result from characteristic elements which make up the town atmosphere. It is created by a user-friendly scale, cosy ambience, the sense of safety, the impression of peace and privacy. Those elements mutually influence one another strengthening the desired impressions shaping urban space and its mood. The small scale of objects creates the already mentioned cosy atmosphere that enhances the sense of balance and peace.

Why can towns be an important subject of scientific research? The issues associated with small towns have been addressed relatively rarely.

Because of the frequently uncertain economic situation of a small urban centre, it is much more difficult to try to promote it. It results from limited possibilities offered by towns which are unable to provide such convenient access to various services, and primarily places of employment, as big cities. Thus. it is more difficult to introduce changes in their spatial structure. For this reason, they also often remain overshadowed by a neighbouring, thriving city. It is not a rule, however, since one can notice small towns strong enough not to need support from a larger urban centre – they are usually towns specialising in one predominant function (e.g. tourist function – Zakopane, or health resort function – Iwonicz Zdrój). Small towns also frequently meet some of the requirements of a high standard living environment. In small towns, one can find a lower intensity of building development, greater intimacy, restricted car traffic in the centre, more greenery, more low buildings, but more difficult access to services especially those more exclusive.

An indisputable asset of small towns is their often individual character built up by their indigenous culture, regional architecture, or valuable architectonic objects and unique landscape. Not always are those unique assets of small towns commonly known. The residents themselves are often unable to notice and appreciate the value of particular architectonic

objects, especially those whose origins date back to the distant past. The lack of social awareness concerning the architectonic value of a place does not help to promote it.

Moreover, popularising such places requires efficient and suitably developed tourist accommodation. In the case when towns are in financial straits, the process is much harder. It requires charting and indicating directions of development, supported by specialist functional-spatial analyses that would allow for identifying the main problems of a town, and on the other hand possibilities of using those elements which could contribute to promoting the place [3].

It also ought to be noticed, that the number of residents of small urban centres is still growing. They are inhabited by approximately 50% of the population (according to the United Nations Organisation). At the time, when we observe increasing threats posed by the development of civilisation, such as growing air pollution, a small town with its pro-ecological solutions and conditions seems to be an important and fundamental place of residence. Traditional solutions, only slightly disturbing the environment built by natural elements, shape the so important nowadays health-oriented conditions, which are indubitably enhanced by biologically-active areas, water reservoirs, and rich and diverse vegetation. Climatic changes partially connected with the manner of spatial development in the densely inhabited areas are not without significance. We affect only a part of climatic changes. To a much greater degree, we aggravate environment degradation and upset the appropriate relations with our surroundings. The process is generally visible in large cities. Expansively developing cities and their smaller structures, such as residential districts, create a new micro-climate changing the existing local factors which, subsequently, contribute to changing the temperature, humidity, sun exposure and direction of the wind. In large agglomerations there occurs an increased emission of heat loss and a rise in the external temperature resulting from the heat loss generated by buildings. The development of urban tissue occurring in cities results in upsetting the balance between green areas and built-up areas, which may lead to the appearance of urban heat islands. Another problem in densely built-up areas is poor air circulation, and consequently poor airing, and continuing intensive air pollution. Despite many difficulties related to living in a small town, in light of health-oriented conditions a town can be regarded as an attractive place to live, both because of climatic and environmental conditions, and those resulting from its structure: namely its scale, proportions shaping the small-town, peaceful ambience. In this work, social conditions partially determined by the morphology of a town – its spatial layout and architecture were regarded as a priority. The analytical part of the work presents the significance of those issues in shaping a high-quality residential environment in small towns.

2. Analysis of selected examples

2.1. General comments

Three small towns located in the Podkarpackie Voivodeship – in south-eastern Poland – were selected for this analysis. On the one hand, the Podkarpacie area is a poorly developing region of Poland, but on the other is has a high tourist value because of a variety of natural

elements, clean rivers, low air pollution and architectonic treasures. The main problem is fragmented agriculture, which is the fundamental branch of economy in this region of Poland, as well as hidden unemployment.

The inventorying, urban planning analyses and sociological research were carried out in Kołaczyce, Sieniawa and Zagórz, which allowed for evaluating elements constituting their multi-level character.

The sociological theme developed in the work by means of interviewing residents of the towns selected for the research reflects the needs of man addressed to the inhabited area: their perception of the town, expectations concerning it and opinions about changes that seem necessary to is development.

The morphology of a town – its components – was evaluated on the basis of field research. It represents spatial relationships linked by significant architectonic objects of historic importance. Results of analyses allow for obtaining a complete image of a town – a complex living environment of various social groups with their needs and expectations.

Kołaczyce is located in the western part of the discussed voivodeship, by the national road no 73. The population of Kołaczyce (1,437 people) places it among small urban units. The main functions of the town are agriculture and tourism which develops in Kołaczyce because of historic monuments located there. Among the most valuable are: the Main Square with the *Bartek* fountain (Fig. 1), the neo-Gothic parish church of St. Anna with a stoup from 1632, a shrine from the 1st half of the 19th century, a statue of Our Lady from 1803, a statue of St. Francis from 1885 at the exit from the Jasło – Pilzno road, and three houses from the 18th century.

Sieniawa – the second town selected for this analysis – is located in the eastern part of the voivodeship, in the Przeworsk district, in the Lower San Valley, at the junction of regional roads: 867, 835 and 870. The town numbers 2182 inhabitants and covers the area of 6.76 km^2. The economic development of Sieniawa is based on industry and tourism. Worth noticing is the urban layout of the town from the time of its foundation dating back to the 17th century, with the main square (Fig. 2), a park-and-palace complex of the Czartoryski family, historic churches, and a fragment of masonry-earthwork fortifications of the no-longer-existing castle of the Sieniawski family.

Another town subjected to numerous analyses is Zagórz located in the southern part of Podkarpacie, in Bukowskie Upland. The town is inhabited by over 5 thousand residents and covers the area of 22.39 km^2. Tourism is the leading function of Zagórz. Among the most valuable spaces and objects of historic character there are: the main square, the parish church of the Assumption of the Blessed Virgin from the 18th century with a Gothic-Renaissance painting of the Annunciation to the Blessed Virgin, the orthodox church of St. Michael from 1836 (Fig. 3), an orthodox church in Wielopole from 1865, a burial chapel (1840), and a wooden church in Dolina from 1836.

Fig. 1. Historic market square in Kołaczyce (photo by J. Kobylarczyk)

Fig. 2. Market square in Sieniawa (photo by J. Kobylarczyk)

Fig. 3. Orthodox church in Zagórz (photo by J. Kobylarczyk)

2.2. Survey results

The survey allowed for obtaining information about the residents' views on the living conditions in the examined towns, their expectations regarding their living environment, the elements they particularly appreciate within their place of residence. The work presents

respondents' answers to six basic questions all of which referred to assessing the comfort of living (Fig. 4). Within the topic the survey participants evaluated the following elements and factors of their living environment: the feeling of intimacy and tranquillity (IIS), a view from the window overlooking greenery (WZ), proximity to the place of employment (BMP), proximity to basic services, the size of backyard (Podw.), and proximity of leisure grounds – up to 3 minutes (TS).

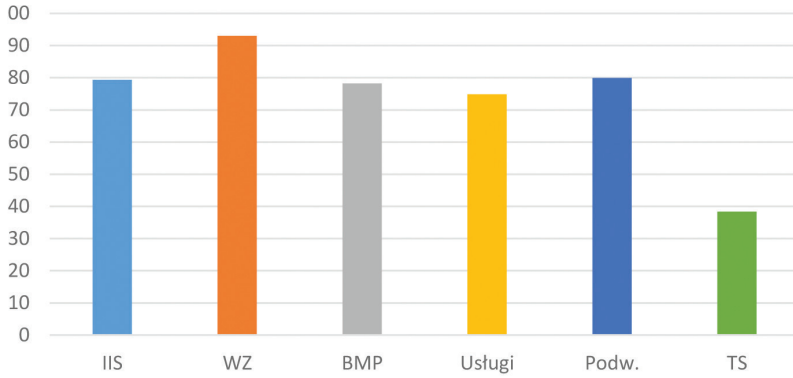


Fig. 4. Survey results evaluating comfort of the place of residence (by J. Kobylarczyk)

The presented research results indicate that 80% of respondents believe that their living environment ensures the feeling of intimacy and tranquillity. Over 90% of the survey participants find a view from the window onto greenery highly attractive, and 80% have decided that their place of work is located relatively close to their place of residence. Moreover, in the respondents' opinion, their living environment ensures proximity to basic services (approximately 75% of answers) and a sufficient size of backyard space (80%). Only 60% of the polled have indicated that leisure grounds are further than within 3-minute walking distance.

On the basis of sociological research, it can be assumed that in many respects small towns can be perceived by their inhabitants as an attractive living environment. The sociological dimension of a small town created by its residents is an immense asset, particularly considering the importance of man – the community forming the image of a town – its identity and culture. The priority among the town functions is whether it satisfies the expectations of its residents [3].

2.3. Spatial conditions in selected areas, considering culturally valuable spaces and objects

Besides the above described multi-level aspects affecting the quality of life in a small town, one should also mention the urban layout and valuable architectonic objects. The spaces and architecture shaped in the past create the cultural landscape characteristic for particular towns. That landscape offers the residents a feeling of "familiarity", being at home, but is also a significant asset in promotion and development of those towns. Skilfully used cultural values can make a small, forgotten and lethargic town flourish again, and become an important local centre, a tourist attraction and a credit to its inhabitants.

The analysed towns exemplify how elements of the cultural landscape, such as a historic urban layout or objects that improve the quality of the residents' lives, very well, and they have a positive impact on the development of those centres. Therefore, the above-mentioned elements of the cultural landscape of selected towns ought to be more thoroughly analysed in order to understand their uniqueness and features that make them attractive for both inhabitants and visitors.

Kołaczyce is a town whose origins date back to the medieval period. It was founded around the year 1339 by the Benedictine monastery in Tyniec [4, p. 57]. In the 1st half of the 20th century, Kołaczyce lost its town rights, but regained them, thanks to the efforts of local authorities, in 2010.

The urban layout of the town dates back to the time of its foundation and is enclosed within a 9-square model, based on the module of 'sznur duży' (app. 45 m). It has the shape of a rectangle measuring 3.5 × 4 'sznur' (157.5 × 180 m). The centre of the layout was the main square measuring 2 × 1.5 'sznur' (90 × 67.5 m) and covering the area of 0.6 ha. This interior was enclosed on 4 sides with single blocks of buildings 1 'sznur' deep, divided into settlement plots [5]. The described layout has not radically changed and has survived until the present time. Its regularity, with a precisely marked centre that is also the heart of the town, is one of several features which make it human-friendly. Another one is connected with its historic buildings creating the character of the place. Buildings fit well into the urban tissue, and boast native details characteristic for the region. Among the most valuable examples in Kołaczyce, there are relics of timber building constructions, including the former post office located by the market square, the church of St. Anna with the 17th-century furnishings, as well as the already mentioned culturally valuable architectonic element in the form of a fountain, statues or shrines.

The current parish church of St. Anna was built on the site of a previous shrine at the beginning of the 20th century. It is a magnificent edifice, erected in the neo-Gothic style, according to the design of an architect from Przemyśl, Stanisław Majerski. The church was built of brick and sandstone. On the plan, its shape resembles the Latin cross. Its silhouette is dominated by the lofty tower situated in the front elevation. The church is located in the vicinity of the main market square. Its height makes it predominant in the town panorama, as well as a town landmark and an important architectonic element.

The historic building of the former post office was erected in the 1st half of the 19th century and shows features of traditional buildings of the region. It was built from wood using the log-house construction. It is an important accent among the present-day buildings around the market square, unfortunately mostly masonry and largely transformed.

In Kołaczyce, one can find many examples of historic architectonic features of which the most valuable is the 19th-century masonry road shrine located in the main square. It was built in the eclectic style. Inside, an 18th-century Baroque statue of Our Lady stands on the stylobate and column. The shrine has recently been renovated which makes it a particularly attractive ornament of the main square in Kołaczyce.

The described objects of historic provenance, designed or built by indigenous artists and craftsmen, occupy a special place in the inhabitants' consciousness. They make people feel safe in their surroundings and identify with the cultural landscape of their town.

The second analysed town is Sieniawa, which was granted its town rights in 1676, as a result of the urbanisation of private estates of the Sieniawski family. The town was founded as a local market centre and a powerbase for a port. An earthwork-and-brick fortress linked to the town was built next to it, with a wooden manor and a parterre garden. This led to the creation of a Renaissance urban layout known in the literature of the subject as a „town-fortress", complemented by a river port as the main city-forming factor, the Observant Dominican monastery and the Tarnogrodzkie suburb. The urban layout of the town covered the area of approximately 12 hectares. It was enclosed within an outline of Renaissance defensive ramparts which no longer exist. In the centre, the main market square was measured out which was surrounded by a row of single building blocks [6, pp. 102n].

The cultural landscape of the town, besides the already mentioned main square and relics of the fortress, consists also of the park-and-palace complex, as well as the town hall, the former monastery of Observant Dominicans and the edifice of the Gymnastic Association "Sokół".

The park-and-palace complex was founded in the 18th by Adam Mikołaj Sieniawski. Its current appearance is due to the rebuilding carried out in the years 1881–1883 by the Czartoryski family. Nowadays, the complex consists of 6 manorial buildings, a historic park with an Italian and an English garden, and valuable architectonic features such as fountains and bowers [6, pp. 102n].

The town hall in Sieniawa was built in the 2nd half of the 17th century, soon after the town foundation. It was located in the centre of the main square as a predominant feature, and has remained so until today. Originally it was a masonry object, partially cellared, with a gable roof. With time, it was expanded, more rooms were added some of which served as shops, while others were used by the town administration. Currently, it is a 2-storey building, maintained in the classicist style. In recent years the object has been revalorised. A tower was also added, which is supposed to resemble the one that once crowned the object [6, pp. 102n].

The former monastery of the Observant Dominicans serves today as the parish church for the town and is dedicated to the Assumption of the Blessed Virgin. The church was erected in the years 1719–1749 according to the design by G. Spazzia in the Baroque style. In 1754, the monastery was added, which ceased to function as such in 1778 as a result of the dissolution of monasteries implemented by the Austrian oppressor [1]. It was then that the church began to serve as the parish church for Sieniawa, which has continued to the present day. Revalorised a few years ago, the object is one of the town landmarks, and together with the edifice of the former Gymnastic Association "Sokół" creates a specific gateway to the town from the south.

The last among the above mentioned historic buildings in Sieniawa is the edifice of the Gymnastic Association "Sokół", which was erected around 1920, on the site adjacent to the parish of Sieniawa [7, p. 2680]. The building is an ideologically important object in the town. It was here that the town residents met since the 1st quarter of the 20th century. At first, it was connected with the functioning of the organisation i.e. propagating sports, a healthy life style, maintaining and developing national awareness. Nowadays, the object serves to promote the town, as a venue for organising events, performances and concerts, as well as to integrate residents. Consequently, it is significant in the awareness of Sieniawa inhabitants as it combines the past and the present, and revives their sense of community.

Fig. 5. Urban layout of Kołaczyce on an aerial photo (soruce: Archive of Dep. of History of Architecture, Urban Design and Art, Faculty of Architecture, Cracow University of Technology)

Fig. 6. Fragment of the main square in Kołaczyce nowadays. View from the north-west. (photo by D. Kuśnierz-Krupa, M. Krupa)

Fig. 7. View of the former post office building on the east frontage of the main square (photo by D. Kuśnierz-Krupa)

Fig. 8. View of the palace in Sieniawa (photo by D. Kuśnierz-Krupa, M. Krupa)

Finally, it should be emphasised that Sieniawa has developed in recent years. Culturally valuable spaces and objects have been revalorised. It positively affected the living standards of residents, their feeling of familiarity and perception of the so-called community symbols i.e. spatial forms to which inhabitants themselves give a symbolic meaning [8, pp. 99–111].

Zagórz was founded on the site of a former village established in the 14th century. In the 15th century, the settlement was the property of Mikołaj from Tarnawa, later of Piotr Kmita from Wiśnicz, and then returned to the hands of the Tarnawski family. The most valuable elements of the cultural landscape of Zagórz, with which the local community particularly identifies, are undoubtedly the ruins of the Discalced Carmelites monastery, the orthodox church of Archangel Michael, and the church of the Assumption of the Blessed Virgin. The mentioned church objects are also local landmarks, continuously present in the memory of Zagórz residents, as well as attractive for visitors.

A particularly picturesque accent in the Zagórz townscape is the ruin of the Discalced Carmelites monastery. The complex was erected in the 18th century, in the late-Baroque style. It was located on the Mariemont Hill (345 m AMSL), in the bend of the Osława River. The former monastery complex consisted of the church of the Assumption of the Blessed Virgin, the monastery and utility buildings. The golden era of the monastery lasted until the first Partition of Poland. Later the monastery fell into decline, having been a site of many battles and fires [9, pp. 29–31]. Despite numerous attempts, it has never been rebuilt. Nowadays, it is a valuable historical monument in Zagórz, important for residents and fascinating for visitors.

Fig. 9. View of the western elevation of the town hall (photo by M. Krupa)

Fig. 10. Ruins of the Discalced Carmelites monastery in Zagórz (photo by D. Kuśnierz-Krupa)

Fig. 11. Orthodox church of Archangel Michael in Zagórz (photo by D. Kuśnierz-Krupa, M. Krupa)

The orthodox church of Archangel Michael is yet another significant object in Zagórz, especially for the Orthodox members of the community, since it reminds them of their roots. The object is also a vital part of the cultural heritage in Zagórz. It was erected in 1836 as an oriented masonry building with an impressive 2-storey tower.

The church of the Assumption of the Blessed Virgin was erected in the 18th century in the late-Baroque style. It is a single-nave edifice, with a vestibule and a sacristy. A Baroque altar, and a late-Gothic painting depiction the Annunciation to the Blessed Virgin, which attracts numerous pilgrims, can be seen in the church interior. The building is another landmark of considerable significance for residents of Zagórz, with the so-called tradition of place [10, pp. 23–32].

The described elements of the cultural landscape of Zagórz, like the previously mentioned historic buildings in Kołaczyce and Sieniawa, have their own inimitable climate palpable mainly for the inhabitants of the town. Those objects, shaped throughout centuries, existing within the town space for decades, are important evidence of the community heritage, which has to be taken care of and should be preserved for future generations.

3. Conclusion

A complete image of a town is shaped by its spatial conditions with architectonic objects, as well as its culture and identity, building an individual value and character of the place. An inseparable element of an urban centre and, at the same time, its most important link is the man – a user of the urban space, its inhabitant. He is the reason why a town develops in order to satisfy his requirements. The presented results of the sociological research indicate that respondents highly appraised small towns in Podkarpacie (Kołaczyce, Sieniawa, Zagórz) because of the presence of elements selected for research and phenomena determining the comfort of living. According to the survey participants, those towns largely ensure the feeling of intimacy and tranquillity. Because of the proximity to green areas, residents can enjoy attractive views from their windows and, more importantly, the majority of respondents stressed the significance of their place of residence being close to their place of employment. For the elderly or the disabled, it is vital that the basic services are within reach, or that the social spaces enhancing neighbourly relations are of suitable size.

The results of the survey have shown that the towns selected for research meet the demands of their inhabitants. Despite the difficult economic situation of the region in which the towns are located, residents view them positively because of the quality of the living environment they offer. The social aspect accompanying each town is in the case of Kołaczyce, Sieniawa and Zagórze an asset, the more so since it has a strong positive impact on shaping the towns' identity, which allows for preserving the culture and uniqueness constituting the value of each urban centre.

To sum up, considering the issues associated with the culturally valuable space and objects and their impact on the life of residents of a small town, one has to conclude that the cultural landscape offers local communities a sense of safety, familiarity and mental comfort. Historic spaces and objects also give a chance for development to small towns, frequently being their

main assets. Revalorised market squares and renovated architectonic objects with valuable details can become targets for cultural tourism, which might stimulate the economy, provide workplaces, and consequently improve the quality of the life of inhabitants.

References

[1] *Miejsca, nie-miejsca Marca Auge'a,* J. Kobylarczyk, K. Paprzyca (eds.), Politechnika Krakowska, Kraków 2015.
[2] Rewers E., *Post- polis. Wstęp do filozofii ponowoczesnego miasta,* Universitas, Kraków 2010.
[3] Kobylarczyk J., *Ocena jakosci środowiska zamieszkania w wybranych miastach województwa podkarpackiego po okresie "transformacji" w pierwszej dekadzie XXI wieku,* Politechnika Krakowska, Kraków 2013.
[4] Kiryk F., *Urbanizacja Małopolski. Województwo sandomierskie XIII–XVI wiek,* ROSOŚK, Kielce 1994.
[5] Kuśnierz-Krupa D., *Fundacje miejskie benedyktynów tynieckich w Małopolsce południowej w okresie średniowiecza/Town foundations of the Benedictine monastery in Tyniec in southern Lesser Poland during the Middle Ages,* Politechnika Krakowska, Kraków 2014.
[6] Kuśnierz K., *Sieniawa. Historia rozwoju przestrzennego,* Politechnika Krakowska, Kraków 2001.
[7] Kuśnierz K., Kuśnierz-Krupa D., *Revalorization of the town hall in Sieniawa,* [in:] *Structural Analysis of Historical Constructions,* J. Jasieńko (ed.), Vol. 3., DWE, Wrocław 2012.
[8] Kuśnierz-Krupa D., *Rewaloryzacja zabytków dziedzictwa kulturowego małych miast podkarpacia na przykładzie Sieniawy,* JCEEA, Vol. XXX, No. 60 (4/13), 2013.
[9] Rykiel Z., *Przemiany struktury społeczno-przestrzennej miasta polskiego a świadomość terytorialna jego mieszkańców,* "Prace geograficzne" No. 170, PAN IGiPZ, Wrocław 1999.
[10] Osenkowski Z., *Zagórz nad Osławą. Z dziejów miasta i gminy,* Oficyna Wydawnicza Miejskiej Biblioteki Publicznej im. Grzegorza z Sanoka w Sanoku, Sanok 2006.
[11] Pawłowska K., *Idea swojskości w urbanistyce i architekturze miejskiej,* Politechnika Krakowska, Kraków 1996.

TECHNICAL TRANSACTIONS 3/2018
CZASOPISMO TECHNICZNE 3/2018
ARCHITECTURE AND URBAN PLANNING
DOI: 10.4467/2353737XCT.18.034.8296
SUBMISSION OF THE FINAL VERSION: 20/2/2018

Elżbieta Kaczmarska (kaczmarskaelzbieta@poczta.onet.pl)
Małgorzata Kaczmarska
Department of Architecture and Fine Arts, Andrzej Frycz Modrzewski Kraków University

New developments in the Liverpool Waterfront[1]

Nowe realizacje w nadbrzeżnej strefie Liverpoolu

Abstract

In the present article, we focus on the changing face of the revitalised parts of the city of Liverpool. We present the successful, consistent implementation of a comprehensive vision for the renewal of the city's port area – neglected and dangerous still at the end of the last century. We indicate the importance of determining the elements of the space that crystallise the urban layout of the city, and skilfully applying those in future concepts. We refer to the main issues related to the renewal process, including the economic and social aspects of the city, as well as to the use of art in the humanisation of the place and the importance of social participation in the approval of design solutions.

Keywords: Liverpool, docklands, revitalisation, vision, concept, participation, public space

Streszczenie

Artykuł koncentruje się na zmianie oblicza rewitalizowanych części miasta Liverpool. Prezentuje sukces konsekwentnej realizacji całościowej wizji odnowy zaniedbanej i niebezpiecznej, jeszcze pod koniec ubiegłego wieku, strefy portowej miasta. Wskazuje na znaczenie ustalenia elementów przestrzeni krystalizujących układ urbanistyczny miejscowości i umiejętne ich wykorzystanie w koncepcjach przyszłościowych. Nawiązuje do głównych zagadnień związanych z procesem odnowy, obejmujących stronę gospodarczo-ekonomiczną i społeczną miasta, odnosi się do wykorzystania sztuki w humanizacji miejsca oraz znaczenia partycypacji społecznej w akceptacji rozwiązań projektowych.

Słowa kluczowe: Liverpool, doki, rewitalizacja, wizja, koncepcja, partycypacja, przestrzeń publiczna

[1] The article was written in reference to the monograph (in preparation for publication) *Przemiany przestrzeni miast na wybranych przykładach* [Transformations of urban space, illustrated with selected examples] by Elżbieta Kaczmarska and Małgorzata Kaczmarska at the Andrzej Frycz Modrzewski Kraków University (research task WAiSP/DS/5/2017). All photographs included herein were taken by the authors during their field visits in 2017.

1. Introduction

Many towns in the north-west and central provinces of England, close to Liverpool, such as Manchester, Chester, Blackwood, Leeds, and Stafford, were associated with industrial development in the late eighteenth century and throughout the nineteenth century, and its gradual decline since the 1950s. During this period, the search for new functions began, albeit without rejecting the tradition of the place. The aforementioned cities, especially Liverpool, are often referred to as "textile cities". In the past, they have grown rich in sea trade, including cotton and wool, as evidenced by the port buildings, and the number of docks on the right bank of the Mersey River Valley; and later on, a revolutionary, industrial-scale production of textiles, which provided Britain with wealth, and the local population and the largely immigrant community from the neighbouring Ireland, with jobs. Due to the declining significance of this industry in the twentieth century, the need to improve environmental conditions, and the promotion of new social ideas about a better and more comfortable life, these cities have gradually evolved into tourist resorts and "people-friendly places". The English city of Liverpool, selected as the European Capital of Culture in 2008, has been undergoing its revitalisation process since 2004.[2] Nowadays, the results of this renewal can be traced in the spatial changes within the Mersey river embankment. Renovated historical dock facilities with service functions such as shops, bars, restaurants, galleries, and museums, already in 2004 were entered on the UNESCO World Cultural and Natural Heritage List, which became an impulse for ordering the entire space. A spectacular example of the transformation is the functioning and the image of Albert Docks, so far the key investment and historical testimony to tradition, and to the spatial metamorphosis occurring throughout the city and its social history[3].

In this process, a particular role is played by the urban and architectural concepts, developed by the outstanding personalities of this world of "creation and magic", producing the suitably attractive atmosphere of novelty and modernity at the highest level. One cannot fail to mention the significant role of the adopted spatial strategy, the entire process of management, implementation and maintenance of the transformed public spaces and their icons, as well as the participation of the community in their approval. Each of these factors contributes to the creation of a beautiful space.

---

[2] The article omits the mention of the definitions of the "revitalisation" concept being in circulation and the models of this process, referring to the author's publications, including: [5, 6]. It is worth emphasising, however, that in the case of the city of Liverpool, the native, British model of renewal was applied in the revitalisation process, which means significant financial involvement of private partners. This model assumes the creation of ventures, which include the public sector and private investors selected in the tender procedure – the public-private partnership (PPP). The PPP model has also been successfully applied in Manchester, Belfast and London.

[3] In 2014, Liverpool counted about 467,000 inhabitants as was the ninth largest city in the United Kingdom; the Liverpool region with the population of Warrington, Wirral, Flintshire, Chester was estimated at 2 million inhabitants; the metropolitan zone of Liverpool with several cities in Cheshire and Lancashire has reached a population of 1.38 million inhabitants [10] In 2014, the number of temporary users spending at least one night in the city was 608 thousand; Liverpool being the sixth UK city most visited by tourists in that year [11].

The article focuses on the visual side of the revitalised space on the River Mersey, turned into a lively walking boulevard, a place offering an array of attractions, music events, sports and social activities. The presentation of the revitalisation of the Mersey waterfront between Prince's Dock and Queen's Dock does not omit references to changes in the structure and image of other districts of the city, because they are the result of the principle of spreading the good practice in the process of renewal.

Transformations in the urban space of Liverpool, regarding the adaptation of the degraded historic docks to new functions, initiated the process of creating the image of the city and its new investment projects, which are noticeable and significant – and which go beyond the public spaces associated with many promenades in the waterfront part of the city.

Other parts of the city have also been transformed, including:

- a shopping district with predominantly commercial function,
- a cultural area with museums (UNESCO World Heritage Area) and central library, the quarter of taverns and music, related to rock-and-roll and The Beatles (The Cavern Club),
- a University district on the east side of the city, and
- a zone between sacred buildings: the Metropolitan Cathedral and the Liverpool Cathedral, which brings together adapted objects of the historical industry for artistic purposes, transformed into clubs, cafés and places for performance activities.

In its vicinity, the planned and heavily promoted China Town brings new services and residential facilities with avant-garde architectural features.

In the central part of the city, with predominantly residential and production functions in the south-western zone from the very centre, we find industrial transformations (the Baltic Triangle) and a significant number of altitude objects, housing residential and office space. The scale of the objects is clearly visible in the panorama and the silhouette of the right bank of the Mersey River.

2. Key investment projects in Liverpool waterfront

The advantageous location of Liverpool city at the Mersey estuary to the Irish Sea, the expansion of the ocean port and its operation, have ensured in the past a favourable, versatile connection between the city and the continent of the Americas. Nowadays the accessibility of the centre of Liverpool and its coastal zone is provided via road connections and new facilities of the John Lennon airport about 15 km from the centre, as well as the proximity to the city of Manchester – much larger than Liverpool, but alike in character – and Manchester airport. The metro line, and the bus station located along the Albert Dock walking route, provide the internal connections.

The city has maintained the status of the UNESCO Maritime Mercantile City; yet currently, due to the spatial changes, it has been associated mostly with new urban and architectural creations in the waterfront, the place of art and artistic activity, the place of sports – thanks to its football teams – and, above all, of "rock" music, thanks to The Beatles. The history of the famous four, the souvenirs and the memorabilia associated with the band are exhibited in every zone of

the city, and penetrate it in various ways; also, a new building at the revitalised boulevards (The Beatles Story) is devoted to this tradition.

In the city system, the waterfront boulevard is a linear element along the river. In relation to the city, the line runs from north to south, slightly twisted towards the east, and it forms the western border of the central area. It is well connected with the rest of the city by the already functioning East-West walking routes. Spatial development of the boulevards area southwards, through the preserved historic urban layout, indicates the possibility of providing good walking connections with new districts and revitalisation in the area of sacred buildings and music-club facilities.[4]

In the boulevard zone, the existing facilities are highly significant in the organisation of the space, due to the functions and representative style of architecture of buildings from the beginning of the last century. The monumental character is found in the Liver Building, Cunard Building, and Port of Liverpool Building. Their presence is of particular importance, as it indicates the continuity of the tradition of the place, its character and the undeniable riches of the past.

The zone of the main – and the most important – connection between the centre and the area of the boulevards is determined by the first revitalisation in that place: Albert Dock buildings, signalled by two port basins of Canning Dock and Salthouse Dock, and preceding tourist information and modern facilities of the Hilton hotel centre. The objects, which – by themselves – constitute a modern solution of a hotel function in the form of a complex of two buildings on a plan of a curve, and of a significant height, highlight the entrance to the much lower Albert Dock buildings.

The zone north of Canning Dock, functionally connected with music and art, boasts numerous buildings with a remarkable, modern architecture, and the arrangement of "magical places", the accumulation of objects with an individual expression, aspiring to the works of art, as well as an experimental combination of various shapes, directions and designers' aspirations for individual artistic expression. These are the Museum of Liverpool (with an interesting interior and stairway solution), the RIBA, the Open Eye Gallery and the Beatles Story, the latter slightly further to the north[5].

It is also worth pointing out the promoted symbols of the city – both those traditional ones, such as the mythical bird crowning the Liver Building tower[6], also found on sports trophies and appearing in the form of tourist gadgets, and the more contemporary "eye candy", promoted as souvenirs. Superlambanana belongs to the latter category. It is a combination of a lamb and a banana – in an artistic association – into one yellow "creature". It is omnipresent, in various sizes, in souvenir stores, in museums, in open public places, including residential areas.

[4] The plan obtained from the City of Liverpool Municipal Information is the reference for the location of the discussed city zone, and the major objects mentioned in the article. Other public access cartographic and information materials were also used.

[5] The Museum of Liverpool was realized by 3XN office (main architect: Kim Herforth Nielsen); the "Black Granite" (as nicknamed by the authors of the text) – the headquarters of RIBA North and Open Eye Gallery, among others – was the Broadway Malyan winning competition project from 2015 [12–14].

[6] Currently the headquarters of the British Music Experience – an institution presenting contemporary British popular music – for more information, see: reference [15].

A special role in the entire complex of new realisations belongs to the RIBA facility, with an elegant, tranquil shape and a functional, modern interior, in its part devoted to educating children in architecture and urban-planning. Workshops accompanying this task are aimed at developing aesthetic sensitivity from an early age, and promoting spatial imagination, understanding and openness of the youngest generations to art. The RIBA object was accompanied by a particularly beautiful artistic installation. The passage to the RIBA building, from the side of the Canning Dock, encloses the square, covered with rolls of delicate muslin falling from above, forming a rectangular prism. One could wander in it, wrap oneself around it, enjoy the delicacy of the fabric and the breeze, pose for photographs, and appreciate the subtlety of this idea.

An undeniable advantage of the revitalised space is water. Its presence, the direct access, the light, the breeze, the slight wave, the reflections at different times of day and night, all create a mood, which is an added value, difficult to measure. This opportunity is certainly not wasted in the discussed city. The docks on the Mersey River, allowing the safe mooring of river and ocean vessels, as well as their repairs, have been fully utilised in the revitalised area, to enhance the aesthetic values of places; friendly to leisure activities and sports competitions. In addition, due to the number of docks, their different dimensions and different shapes of port basins, the impressions, positive feelings and moments of surprise have been multiplied – attracting both the local community and the tourists. Involvement in the organised mass events also constitutes "soft" participation[7].

We should not fail to mention that the promenades are equipped with very comfortable, aesthetically pleasing and warm, wood-lined seating places stretching on several levels, places to unbuckle bicycles along the route, and sculptural accents – for example, The Beatles and local hobos made of bronze, and metal musical instruments.

The accumulation of so many objects, each with their individual, sometimes aggressive expression, can be debatable, but it can be seen as an attempt to find new forms of artistic creativity and innovative principles in urban planning and architecture.

The area south of the Albert Docks is definitely different in character from the ones previously discussed. Wapping Dock and the – not yet fully fitted out – Queen's Dock have gained modern sports and hotel facilities, loud entertainment venues, conference facilities, as well as performance and concert halls. These functions are located in glass buildings with large cubic capacities. The buildings of ACC Liverpool, Liverpool Exhibition Centre and Pullman Hotel are restrained in their expression. It can be assumed that the next two basins in the southern zone will also receive proper function and arrangement. The latter are closing the area of new housing developments, mainly apartment buildings, office buildings, and transformed industrial facilities, mostly factory halls (the Baltic Triangle district).

In Liverpool, we have met with a new model of employment, and a method for promoting significant places in the functioning of the city; we believe it may potentially serve as an example to follow. In the areas of high tourist traffic, and the accumulation of public buildings, the confused newcomer is met by senior citizens – of the charming "type" we know from English

[7] By this wording, the authors understand such actions by the authorities and the designers, which will allow – in an unobtrusive way – to perceive the acceptance of the places by their users.

films and theatre plays – who are ready to provide comprehensive information. Wearing reflective vests, they give basic information about the city. They are friendly, reliable, and very knowledgeable. This model has many benefits. It employs people of retirement age and allows them to be useful for society; through familiarity with their own social environment, it meets basic educational functions at various levels of knowledge. The seniors assist visitors from both Europe and Asia – in fact, tourists from China and Japan constitute a significant group of users of this information service.

The place becomes a theatre of a kind, where spectators and actors alternately create their roles. This only confirms that the deindustrialisation processes in Liverpool were carried out well.

3. Symbols of the city and their significance in creating Liverpool's image

Under the heading of "Liverpool", it is not easy to visualise one single building or engineering object. There is no equivalent to instantly recognisable and place-associated building, such as the Eiffel Tower, the Empire State Building or the Burj Dubai. It seems that the city is most often associated with The Beatles music band and the Liverpool Football Club. Both these city symbols serve for community building, and they are also used to a certain degree in creating the image of the city. All places in Liverpool related to the creation and success of The Beatles are advertised, indicated on the maps and signposted, although they lack a common, designed visual identity. There is also a monument to the musicians – in the most representative space of the city, on the boulevard along the Mersey River. Just like the others, also this statue does not surprise. It is a realistic representation of the four heroes, twice the life size – a perfect background for a selfie.

The Liverpool F. C. coat of arms boasts a curious origin [16]. In its central part, we find the Liver Bird – recognised throughout the world as a symbol of the Liverpool club – a mythical bird, a combination of a cormorant and an eagle. It was placed there, as it had been a symbol of the city for over 800 years. The symbol goes back to the first half of the thirteenth century, as a characteristic sign on a unique seal used to authenticate documents of the local trade guilds. It appeared in the city crest in the second half of the seventeenth century, and it still remains there. In addition to the omnipresent logo of the football club, it can be found at the top of the representative Royal Liver Building. In fact, the building is crowned with two birds, facing opposite directions, and seated on two opposite clock towers.[8] Liver Bird's silhouette was used as the logo for the 800th anniversary of the city in 2007, and it continues to play an important role in identifying the place – marking numerous buildings throughout the city. In a variety of

[8] 18 feet tall each, they were sculpted by Carl Bernard Bartles. One is directed towards the river, the other towards the city. Their names are Bella and Bertie. Legend has it that one protects the inhabitants of the city, and the other protects those who are coming into the port. The building was established in 1911 as the headquarters of the Royal Liver Assurance group. It is known as one of the "Three Graces" of the Liverpool waterfront, together with the main port building (Port of Liverpool Building) and the seat of the maritime carrier (Cunard Building) [17].

materials, colour versions and forms, it draws attention to itself in every point of sale of souvenirs – as a typical "gadget", keychain, and a graphic element on anything else. Recently, however, it has met with strong competition – from the Superlambanana.

Superlambanana has its own history and message, and it is also an art object in its own right. However, even before you get to know its history, you will have seen it all over Liverpool, in countless versions – it undoubtedly has a high image-creation potential. It is a yellow sculpture, weighing over 8-tons, 17 feet (5.2 m) high, showing the cross between a lamb and a banana, as its name suggests. It is a conscious reference to the port traditions of the city, and more specifically to the export of English wool, and the import of bananas. At the same time, the form of the sculpture is to be a warning against hasty manipulations of genetic engineering and GMOs. The main Superlambanana is currently standing in front of Liverpool John Moores University, and its numerous miniature versions and free interpretations can be found throughout the city. Superlambanana was created in connection with the ArtTransPennine[9] exhibition in 1998. It was designed by Taro Chiezo[10] – a Japanese artist living in Manhattan – and presented to the public during the opening of the Liverpool branch of Tate Gallery. Its 125 mini version (2 meters high each) were created in connection with the celebration of Liverpool as a European Capital of Culture in 2008. For 10 weeks (from June to August 2008), the mini Superlambananas, painted by local artists or groups representing local communities according to their designs, were exhibited around the centre of Liverpool, one of them in London, and another in Wales – intended as the showcase of the city. Afterwards, most of them were sold at charity auctions, obtaining high prices, far exceeding the initial estimates. Finally, in 2010, eight new mini versions were created[11] as permanent fixtures for Liverpool's public space. Four of them can now be seen in front of the Liverpool Museum, that is, in the most representative part of the waterfront. They serve the users of the promenade – as objects of photographs, and for children to climb. Here they are painted in different colours and patterns, thus diversifying and cheering up the place. Superlambananas focus attention. The mini versions – as they are closer to the human scale – have become a popular "selfie material". In the gift shops, Superlambananas boldly compete with the Liver Birds, even with those in the "club" version; and they outshine The Beatles – mainly in terms of their bright colour, and their compact, yet characteristic form. The ceramic souvenir Superlambananas constitute the "supermini" version. They are sold in a box of 16.5 × 20 × 8.5 centimetres – and they have the status – if not of art pieces (being a reproduced form), then certainly craft pieces. The sellers place the history of the sculpture on display, thereby justifying the high status of the souvenir (covered by copyright), as well as its price. For less sophisticated customers, there are more price-friendly Superlambananas – as stuffed animals, pendants or key rings.

---

9 ArtTransPennine Exhibition was part of the initiative to create the "corridor of art" through the North of England. Superlambanana was Liverpool's contribution to the 2008 exhibition. The festival was first organised in 1998 as the largest regional arts event, covering the area from Manchester to Liverpool and the Pennines region with Leeds, Sheffield, and Hull [18].

10 Chiezo created a small-scale model, while local sculptors: Andy Small, Julian Taylor, Tommy Reason and Ray Stokes, made the final version. A wire mesh frame was used to support the concrete and glass fibre shell [19].

11 Known as "Eight for 08".

The importance of objects such as Superlambanana, and symbols such as Liver Bird for the local community is enormous. All you have to do is look at any of the few tapestries that have been created in connection with the celebration of Liverpool as the European Capital of Culture. They were created as a result of embroidery workshops, open to the public, with the task of capturing the spirit of the city – its inhabitants, places and passions. Therefore, amongst the 34 embroidery pieces in one of the tapestries, as many as eight show Superlambanana; two – both Liver Birdies; one – the Liver Building; one – The Beatles; one – Tate Liverpool; one – each of the cathedrals (Anglican and Catholic); two represent ships; and the remaining ones show other objects, places or figures.[12] This example shows that Liverpudlians turn towards art and architecture – they notice various architectural works, note the existence of a local branch of the prestigious Tate Gallery, and finally – they notice the artworks located in the urban space. It is also clearly visible that in 2008 the city lacked a building, which by creating the city's image and improving its status would become a kind of landmark, for the locals to identify with, and for the tourists to appreciate. It should also be added that despite the existence of numerous sculptures in the urban space, it is only the Superlambanana that draws everyone's attention.

Without reflection on the pleasant Superlambanana, the city could now try to commission new art objects, the quality of which should at least match the quality of the most recently created architecture. It is possible that ambitious artistic projects have a chance to arise in connection with the new architecture. The first presentation at the Open Eye Gallery at the newly established RIBA North headquarters is a good example. The Open Eye Gallery is a space between buildings – a passage called the Winter Garden, parallel to the Mersey River, connecting the space in front of the Liverpool Museum and the "Black Granite" with the space next to the Liver Building. The installation, created by members of RIBA North – Karsten Huneck and Bernd Truempler – was meant to interpret visually and spatially the goals and tasks of the new national centre for architecture. It was built of numerous curtains of fibre mesh in red, hanging freely in the gallery space. The curtains did not touch the ground, but their distance from the ground was graded, by which a rather large "breathing space" was created inside. The whole was in the form of a cuboid, with the height of one storey in relation to the buildings between which it was located. A clearly defined block of the installation, both from the outside and inside, and its decisive colour, combined with the lightness of the mesh moving freely under the influence of wind or people walking, created the effect of floating above the ground and thus gave a weight-balanced effect, interesting aesthetically, and having a multisensory impact.

4. Conclusions

The revitalisation process of Liverpool waterfront, initiated in the first decade of the twenty-first century, was clearly conducive to turning the neglected and degraded areas of the industrial city into places saturated with a modern lifestyle, tranquillity, respite and beauty. In the clearly

[12] 73 persons of varying ages contributed to the embroidery; 338 panels were thus created, combined into several tapestries. The event was organised by the local embroiders' guild. End result is on display at St. George's Hall.

defined functional areas of the waterfront, the following aspects were not overlooked: physical activity, sports competitions, as well as loud music events with crowds of participants. The mass nature of such events has been programmed in the spatial concept for the designated areas. Each of these zones has a clearly defined character, psychological climate, a distinct sense of place and unique architecture, arrangement of interiors, and external spaces.

The compositional values of the layout of the new boulevard refer to the planning concept of the entire city. Its most important axes connect the centre with the most spectacular places, with intriguing architecture, and with the main walking routes in the area of seafront boulevards.

The image of the place – perceived through its visual setting – clearly impacts the changing face of the entire centre and its neighbouring districts, as evidenced by the investment projects commenced in further downtown areas.

The new buildings – designed by outstanding architects, forming project teams selected in international competitions – play a special role in these activities. Modern forms of architecture, design ideas, construction materials used, despite their diversity and perhaps their experimental compositional intent, create attractive spaces that link the interiors with the external context.

The applied spatial details enrich individual spaces and define the symbols of the place, both traditional and new ones, contemporarily created.

The attractiveness of the place is evidenced by crowds of its users, benefitting in different ways from the modern spaces available to them: walking, resting, participating in various sports competitions, listening to music, looking at works of art, and learning to take advantage of numerous organised events, which is facilitated by the logical compositional principle of the whole planning scheme.

There is no doubt that the significant role in such an interesting space is played by historical objects and their traditions, as well as well-known personalities of the art world associated with the place. By virtue of the dialogue with tradition and place, as well as the contemporary outlook, all the components mentioned here fit in well with modernity.

References

[1] *Anglia*, eds. E. Kosiba, Mediaprofit Sp. z o.o., Warszawa 2007, 291–292.

[2] *Time, Place and Identity, Planning, Urban Design, Landscape Architecture*, eds. L. Rueb, S. Khandekar, BDP, United Kingdom 2009, 5–19.

[3] *Visit Liverpool. The official Guide to the Liverpool City Region*. Marketing Liverpool: Liverpool 2017.

[4] *Wielka Brytania*, eds. M. Leapman, Wiedza i Życie, Warszawa 1996, 362–365.

[5] Kaczmarska E. *Procesy rewitalizacyjne we współczesnym rozwoju miast Polski*, [in:] *Wybrane problemy funkcjonowania administracji publicznej*, eds. H. Franaszek, Wydawnictwo KSW im. A. F. Modrzewskiego, Kraków 2008, 87–92.

[6] Kaczmarska E., *Nowy zawód – koordynator procesów rewitalizacji miast*, [in:] *Społeczne, marketingowe i innowacyjne aspekty kreowania zmian organizacyjnych*, eds. D. Fatuła, Oficyna Wydawnicza AFM, Kraków 2011, 109–115.

[7] Kaczmarska M. *Instalacja artystyczna w przestrzeni otwartej*, submitted for publication in Oficyna Wydawnicza AFM, Kraków 2017.
[8] Kaczmarska M. *O potrzebie tworzenia*, [in:] *Inspiracje twórcze a dzieło stworzone*, eds. S. Hryń, Oficyna Wydawnicza AFM, Kraków 2009, 57–62.
[9] *Liverpool pop city*. Quickmap, Luton 2016.
[10] http://worldpopulationreview.com/world-cities/liverpool-population (access: 26.01.2018).
[11] https://www.liverpoolecho.co.uk/news/business/liverpool-visitor-economy-now-worth-9581949 (access: 26.01.2018).
[12] https://www.archdaily.com/788905/riba-to-open-new-national-architecture-centre-in-liverpool (access: 26.01.2018).
[13] https://www.archdaily.com/151034/museum-of-liverpool-3xn (access: 26.01.2018).
[14] http://www.broadwaymalyan.com/projects/mann-island/ (access: 26.01.2018).
[15] http://www.britishmusicexperience.com (access: 05.01.2018).
[16] http://www.lfc.pl/HistoriaHerbu (access: 05.01.2018).
[17] http://www.royalliverbuildingvenue.co.uk/liver-bird (access: 05.01.2018).
[18] http://www.culture24.org.uk/art/art58357 (access: 26.01.2018).
[19] https://www.liverpoolecho.co.uk/news/liverpool-news/iconic-superlambanana-inspired-liverpools-trading-3388595 (access: 26.01.2018).

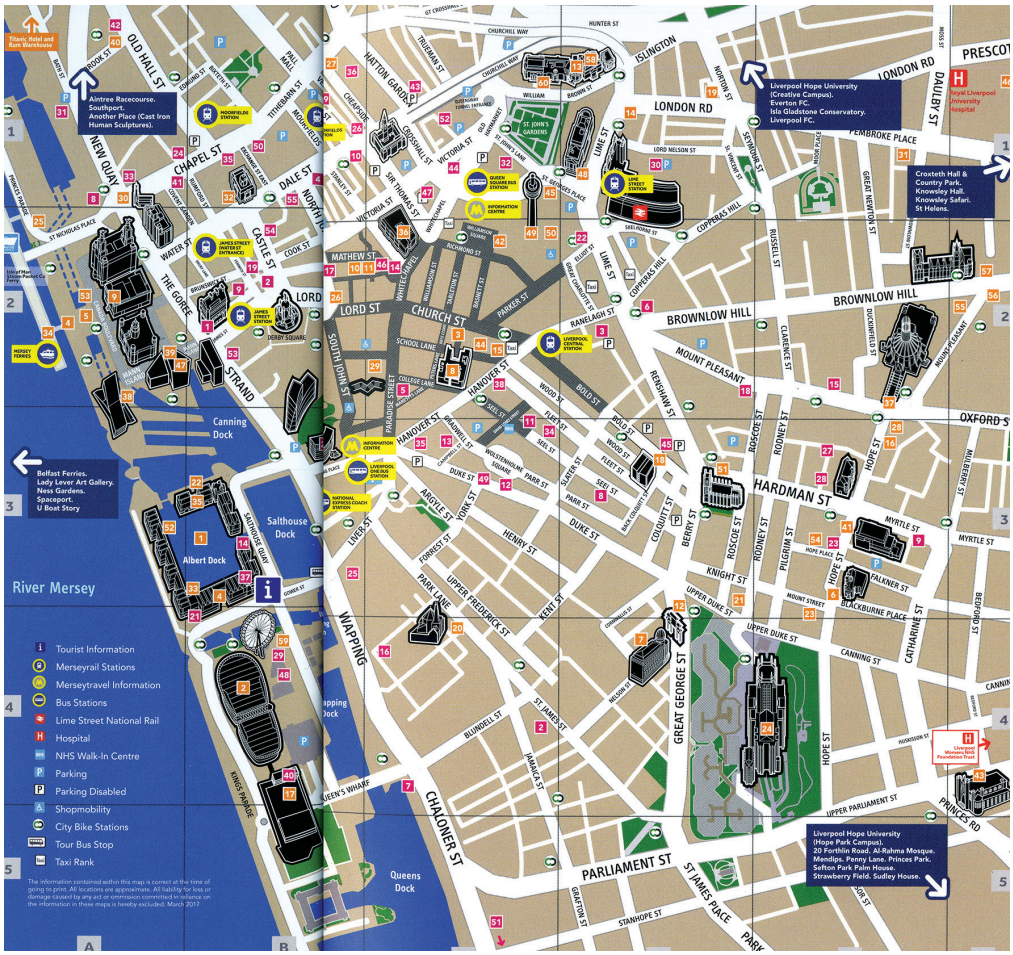


Fig. 1. Drawing of the centre of Liverpool, with clearly marked waterfront zone of the River Mersey, and distinct features assigned to the historic docks. Urban layout of the waterfront zone is a continuation of the main axes of the city centre, its historical part, and new urban developments [3]

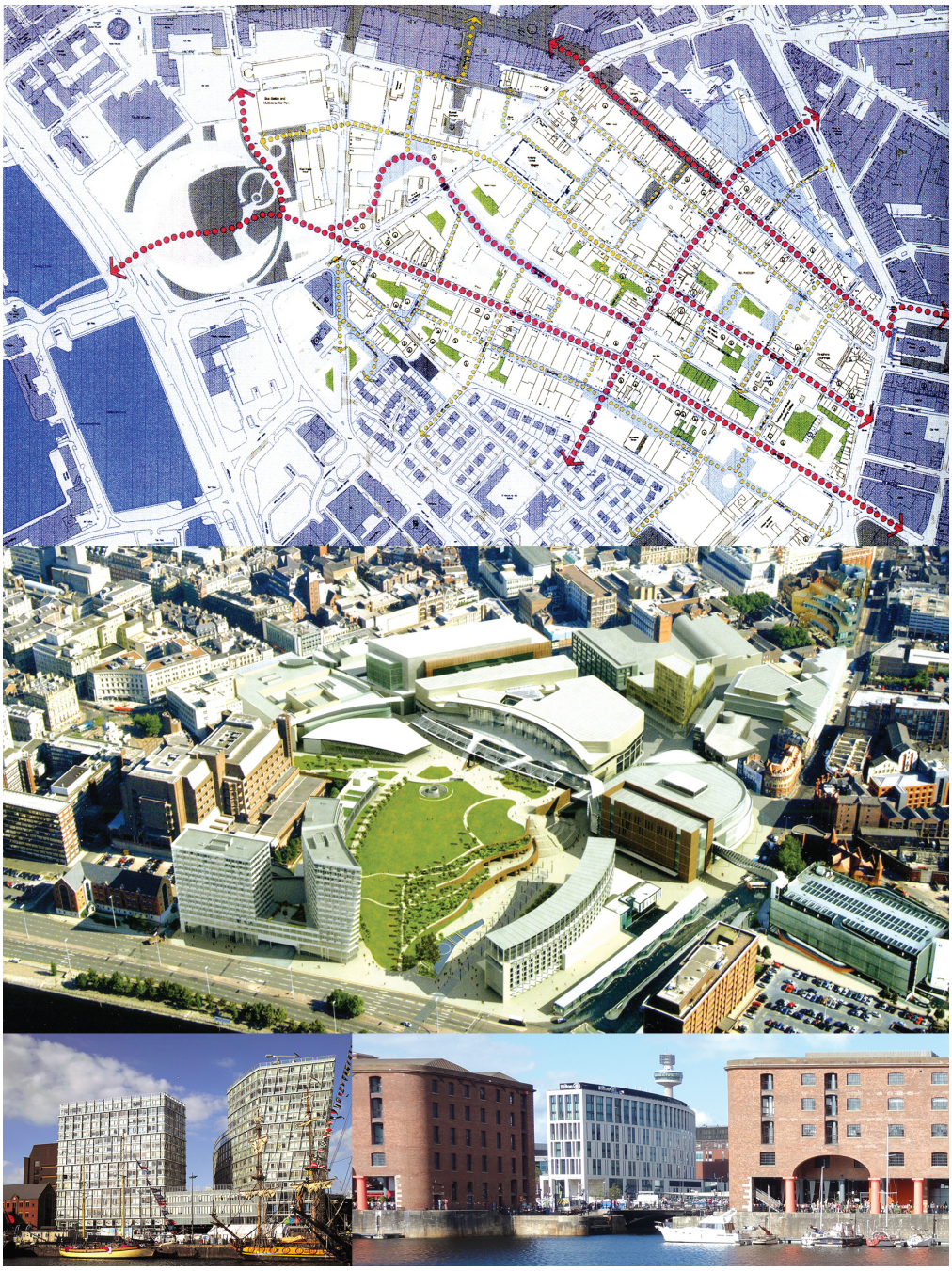


Fig. 2. The concept and visualisation for the reconstruction of the centre of Liverpool in the Canning Dock area, by the London branch of BDP. The company has prepared a development plan for the following areas of the city: Ropewalks, Liverpool One, and Northshore (plan and bird's eye view – [2])

Fig. 3. Public utility buildings from the beginning of the twentieth century: the "Three Graces" of the Liverpool Waterfront, and the later development of the city centre. The northern waterfront zone connects cultural objects, including the Liverpool Museum; objects presenting musical heritage of England, including The Beatles Story in Mersey Ferries, and the British Music Experience in the Liver Building; and objects with leisure and sports functions

Fig. 4. New cultural objects in the Canning Dock area – RIBA North, the "Black Granite" buildings, and galleries – make up the area of art. Due to their expressive, simple shapes, saturated black colouring, and smooth facades, they dominate in this part of the boulevards, constituting the most attractive and modern part of the waterfront area

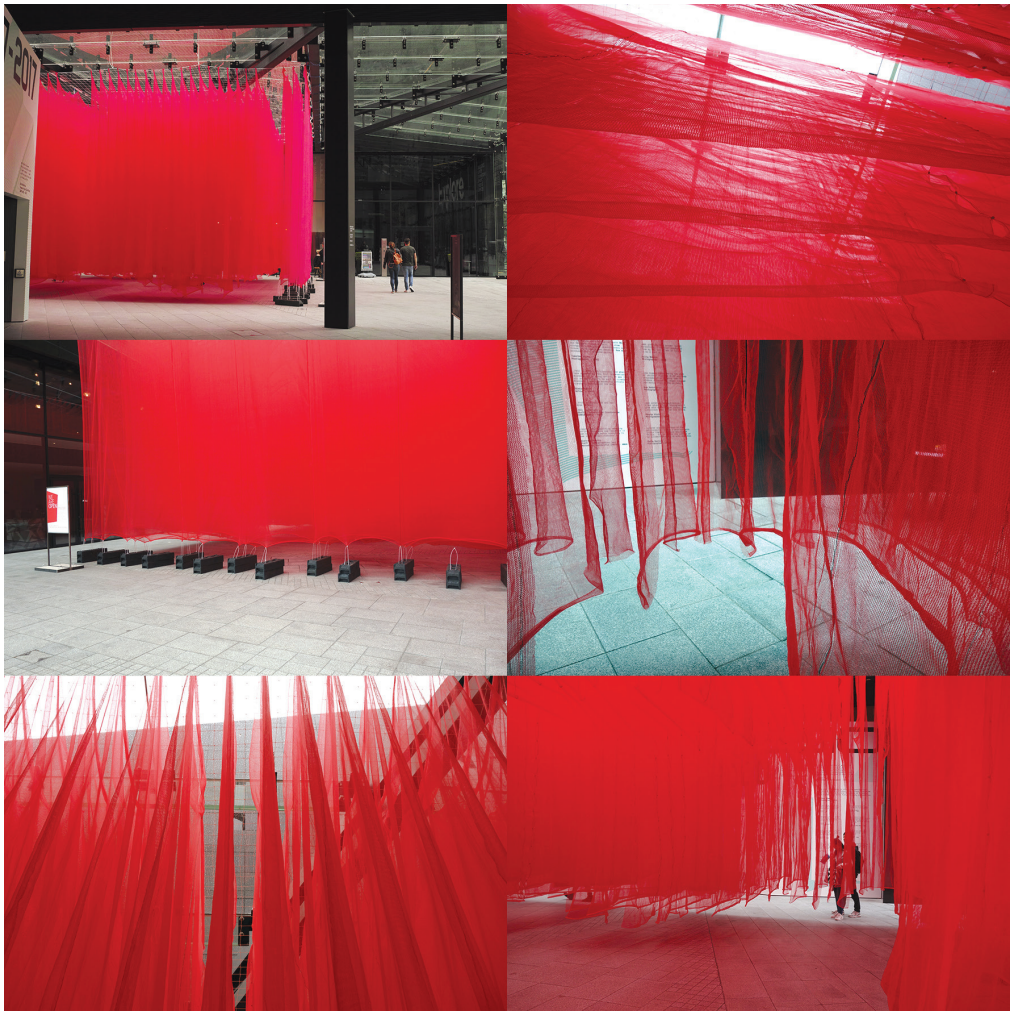


Fig. 5. The installation outside the quarters of RIBA North – expressive, intriguing, and beautiful. It emphasises the elegance of the place, and its ephemeral quality. It is light and airy; and it promotes the space in the best possible, original style

Fig. 6. Albert Dock is interesting at any time and any season – also due to its historical heritage and traditions. Simple forms of the buildings, which enclose the water basin, were built in 1846 according to Jesse Hartley's design; they went into decline at the beginning of the twentieth century, and closed down in 1972. After the revitalisation of all five surviving warehouses, they combine residential, entertainment and service functions: bars, restaurants, pubs, clubs, and boutiques. The promenade by the colonnade provides access to the service part, and is a fashionable place for socialising; you can watch yachts swaying in the dock, feel the breeze and sunlight, walk and sit in a pleasant, safe environment

Fig. 7. The area of sports, mass events and loud music concerts in the southern part of the waterfront, brings together large-sized venues and the style of architecture characteristic for its function. ACC Liverpool houses the Echo Arena and the Conference Centre, while the Liverpool Exhibition Centre serves the International Trade Fairs and hotel functions

Fig. 8. Post-industrial buildings and structures, adapted to new functions in the Baltic Triangle zone, located in the Queen's Dock area of the southern part of the waterfront. The modern office and residential district neighbours the old warehouses of historic Liverpool, transformed into avant-garde places of culture and entertainment. These include clubs, cafés, restaurants and the seat of the Contemporary Urban Centre

Fig. 9. Liver Bird and Superlambanana are the best known and well-liked Liverpool symbols. The mythical bird is a respected symbol; also acting as a trophy or a souvenir. Superlambananas, however, fit well into virtually anywhere in the city – be it a promenade, a square in front of the building, a souvenir kiosk, or a museum. They are of various sizes, yellow, painted in many colours, and performing different functions; they are watched, photographed, and collected

TECHNICAL TRANSACTIONS 3/2018
CZASOPISMO TECHNICZNE 3/2018
ARCHITECTURE AND URBAN PLANNING
DOI: 10.4467/2353737XCT.18.035.8297
SUBMISSION OF THE FINAL VERSION: 26/2/2018

Krystyna Paprzyca (paprzyca@pk.edu.pl)
Chair of the Shaping of the Housing Environment, Institute of Urban Design,
Faculty of Architecture of the Cracow University of Technology

Concept of the development of Oświęcim through an offering of increasing cultural and utilitarian qualities

Koncepcja rozwoju miasta Oświęcim poprzez ofertę wzrostu walorów kulturowych, użytkowych

Abstract

The author discusses the city of Oświęcim in terms of the abovementioned aspects, trying to justify that the city of Oświęcim is following the path of development. There are many issues which cover urban renewal through offering an increase of quality: – an improvement of the quality of the public space within a city; the regeneration and modernisation of buildings and technical infrastructure within its Old Town area; – increasing the rest, recreation and tourism offering of the city in areas along the Soła River [16]; – uphold the identity and culture of the area; – use the creative potential that supports economic growth, including the creation of new employment opportunities, creating conditions for new economic investment. Reinforcing the intellectual capital of the city; – use the potential that is Oświęcim's location at the crossing of transit routes of national and voivodship-level significance.

Keywords: town development, identity and attractiveness of places, offering of an increase in cultural and utilitarian qualities, functional and utilitarian diversity

Streszczenie

Autorka omawia miasto Oświęcim pod kątem aspektów, starając się uzasadnić, iż Oświęcim podąża drogą rozwoju. Jest wiele zagadnień, które obejmują odnowę miejską poprzez ofertę wzrostu walorów: – poprawa jakości przestrzeni publicznej w mieście; rewitalizacja i modernizacja budynków i infrastruktury technicznej, na terenie Starówki; – zwiększenie oferty rekreacyjno-wypoczynkowej i turystycznej miasta na terenach wzdłuż rzeki Soły; – podtrzymywanie tożsamości i kultury miejsca; – wykorzystanie potencjału kreatywnego, który sprzyja rozwojowi gospodarczemu, w tym doprowadza do tworzenia miejsc pracy, tworzenia warunków dla nowych inwestycji gospodarczych, wzmocnienia kapitału intelektualnego miasta; – wykorzystanie potencjału, jakim jest położenie Oświęcimia na skrzyżowaniu tranzytowych ciągów komunikacyjnych o znaczeniu krajowym i wojewódzkim.

Słowa kluczowe: rozwój miasteczek, tożsamość i atrakcyjność miejsc, oferta wzrostu walorów kulturowo-użytkowych, różnorodność funkcjonalno-użytkowa

1. Introduction

The spatial structure of the city of Oświęcim[1] has been determined by the development of the Oświęcim Chemical Plant, around which multi-family buildings were being built en-masse. Funds were not being allocated towards the Old Town area during this time.

The Old Town remained poorly funded, the results of which are visible to this day. This uneven development has caused a series of problems of a social, economic and environmental nature. These include: the process of the ageing of the population, a high level of migration, an insufficient development of businesses.

The tragic heritage of the Auschwitz-Birkenau former German concentration camp has a very significant impact on the image of the city. The historical heritage of the city remains forgotten in its shadow.

The uneven urban development of the city and the phenomenon of its depopulation and sprawl have been caused by a crisis of urban space and by a crisis of public space as well [12]. It is associated with numerous processes, as a result of which the city and its fragments undergo cultural, economic, social, technical and environmental degradation. Significant issues focused around the improvement of the city's image include: eliminating areas suffering from urban decay, with a low quality of life and housing, as well as a broad communisation of this process, directed towards residents, entrepreneurs and tourists.

The qualities of a city that is well-suited for habitation and that teems with life that are of key importance include interventions in urban space which facilitate an increase of cultural and utilitarian qualities.

According to the assumptions of the Heritage and Leisure Industries Strategic Programme, a part of the Lesser Poland Voivodship development policy, one of the key elements of social and economic renewal in Oświecim is the implementation of the Oświęcimska Przestrzeń Spotkań project (Oświęcim Meeting Space in English - transl. note, a project indicated as a part of Action 1.3 Comprehensive social and economic renewal programmes and the shaping of space)[2]. A city that teems with life is a city that is filled with diversity and life, thanks to projects that lead to an increase in the attractiveness of places that are essential to

[1] Oświęcim is a small city with a population of 38,972, of which 52,5% are women and 47,5% are men. In the years 2002–2016 the number of residents fell by 6,7%. The average age of the residents is 43,3 and is higher than the average age of the residents of the Lesser Poland Voivodship, in addition to being slightly higher than the average age of the inhabitants of Poland overall.
The population dynamics index, which is the relationship between the number of live births and the number of deaths is 0,98 and is considerably lower than the average for the Voivodship and comparable to the population dynamics index for the entire country. 58,9% of the residents of Oświęcim are in working age, 16% are in pre-working age and 25,1% are in post-working age. The average age of residents is 43,4, including 45,5 for women and 41,0 for men. In Oświęcim there are 314 employed persons per 1000 inhabitants. This is a significantly higher number than the value for the Lesser Poland Voivodship and considerably more than in Poland. 52,0% of all employed are women, and 48,0% are men. See [17] "Oświęcim » mapy, nieruchomości, GUS, szkoły, kody pocztowe..."

[2] This project is understood as a tool making it possible to overcome the key development problems of the city, among which the following are pointed to: the spatial structure of the city determined by the development of large communities, Programme of the Regenertion of the City of Oświęcim for the years 2015–2025. See [18, p. 11].

a city. Improving the quality of life and a diverse use of space ensure residents and visitors different experiences, impressions and sensations. The climate of a city is also conducive to all manners of technological and economic innovations. It promotes artistic and cultural activity.

The project that will improve the image of the city will affect the increase of its attractiveness and covers the carrying out of a project through developing urban space and establishing an infrastructure that creates bridges that lead from the tragic history of this place to the contemporary life of the city. The project is located along the axis from the Place of Remembrance and along the Soła River to the Old Town and centre of Oświęcim. The leading functions of the structure will include meeting spaces: for people, cultures, experiences and thoughts, which are to be implemented by such functions like: education through history, tourism traffic services and opening it up to the city, cultural and educational projects as well as contact with nature.

The planned project features comprehensive urban renewal, encompassing the urban renewal of the historical part of the city with a focus on the city's diversity in tourism-related aspects.

Its main goals are:

- improving the quality of the public space, particularly that of the city centre and that of a friendly urban space which makes it possible to improve the quality of life of residents and visitors - both tourists and entrepreneurs;
- making use of the potential of the areas along the Soła River;
- establishing a meeting space of a symbolic, educational, cultural and recreational character along the axis between the Place of Remembrance and the Old Town, whose goal will be to link the past with the future;
- making use of the potential that is Oświęcim's location at the crossing of transit circulation routes of national and voivodship-level significance.

The quality of public spaces is a problem that is key in towns that are on the path to development. Their high quality causes residents, as well as other current and future users, including tourists and entrepreneurs, to become more interested in them.

Fig. 1. Oświęcim – the Old Town as seen from a birds-eye view: Oświęcim City Hall | Oświęcim – a City of Peace, Oświęcim City Hall | Oświęcim – a City of Peace 1000 x 666 Image search (source: [19])

2. Improving the quality of public spaces. The urban regeneration of the old town and riverside areas[3]

Fig. 2. Oświęcim – Old Town – panorama (source: [20])

Oświęcim is an old town, with a tradition going back 800 years. Its urban layout and the historical sites of its Old Town section are a monument to history (Fig. 1, 2). One consequence of this situation are difficulties with access to individual public spaces, public services, as well as the necessity of significant costs associated with the development of public infrastructure. The Old Town is characterised by a poor technical condition of buildings and was characterised by a poor state of technical infrastructure up until the time of the Old Town becomes renovated. This state of affairs is the result of an uneven development of the city, as well as the age and construction technology of the buildings.

In the last couple of years the Old Town was being subjected to a gradual modernisation. Renovation work was performed on the Piast Castle and the Castle Hill, where a museum was established. Kościelna Street along with ks. Skarbek Square in the area of the Jewish Centre and the Synagogue were redeveloped and modernised, Bulwary Street was modernised as well, along with the parking lot adjacent to this street. An urban regeneration of the Main Market Square along with its accompanying infrastructure, which was preceded by ownership disputes regarding the surface of the Main Market Square, was performed in recent years (Fig. 3, 4).

All of the interventions that were performed in the area of the Old Town affect the quality of life and housing within it, as well as the quality of its public space, so that it can teem with life[4]. The appearance of public spaces: streets and squares, tells much about the city. People feel safe in friendly and well-kept places[5] and are eager to spend their time there, returning to them often.

---

[3] The area of the Old Town constitutes the historical city centre, which is located on the right bank of the Soła River. The area includes the Oświęcim historical old town along with the surrounding small streets. It is an area of around 68,27 ha, which includes the areas od the Old Town, the Boulevards, the Fika housing estate near Górnickiego Street, as well as the zone between Kracickiego and Fryderyka Chopina streets. The area of the Old Town is inhabited by 2018 people, see [18, p. 61].

[4] Attractiveness – the quality of something, which through its merits, causes the interest of others. Great Polish language dictionary [21].

[5] It is estimated that 768 crimes were perpetrated in Oświęcim in the year 2016. This means that for every 1000 inhabitants 19,68 crimes were registered. This is a value that is lower than the one for the Lesser Poland Voivodship and comparable to the average for Poland overall.

The effective management of "urban resources" by municipal governments counteracts and limits the development of negative phenomena: urban decay, the decapitalisation of buildings, depopulation and deurbanisation. The upholding of cultural values and identity has the purpose of engraining a feeling of belonging to, as well as of ties with the urban community in the conscious of the residents. It is a form of a counterweight to the processes of globalisation that are visible in cities.

Fig. 3, 4. A view of the modernised Main Market Square in Oświęcium (photo by K. Paprzyca, 2016)

Thanks to the investment meant to improve the quality of public spaces that were made, the Old Town of Oświęcim has become a hallmark of the city, a place where the most important cultural events are cyclically held (Fig. 5–8).

Fig. 5–8. Cultural events on the surface of the Main Market Square of Oświęcim (source: [22])

The area of the Old Town features mixed-use residential and commercial forms of use. Among residential buildings, as many as 41 were built before the year 1945, while 12 were built in the years 1946–1989. The average number of people inhabiting an apartment in the years 2010–2015 was characterised by a downward trend [18, p. 67].

The age of the buildings of the Old Town leads to the necessity of carrying out renovation, redevelopment, modernisation and replacement works associated with the technical condition of the buildings and their installations (Fig. 9–14).

Fig. 9. Oświęcim. A vandalised corner townhouse near Klasztorna Street, as well as an abandoned building near the Small Market Square (photo by K. Paprzyca, 2013)

Fig.10. Oświęcim. Renovated corner townhouse on the corner of Klasztorna Street and the Small Market Square (photo by K. Paprzyca, 2015) (source: [23])

Fig. 11, 12. Townhouse at the Small Market Square in the year 2014 and the year 2017 (photo by K. Paprzyca, 2014, 2017)

Fig. 13, 14. Townhouse on the corner of Sienkiewicza and Mickiewicza streets in the year 2014 and 2017 (photo by K. Paprzyca, 2014, 2017)

3. Upholding the identity and culture of a place. Social activation, building social capital in the area of the Old Town and its vicinity

Oświęcim, similarly to many other small and medium-sized cities, is being affected by problems associated with depopulation and the ageing of the local community. The first of these problems involves migration, which is local and is associated with the city being affected by urban sprawl. Oświęcim has a limited amount of buildable areas, including those assigned for residential buildings, which leads to residents often searching for plots outside of its borders. Thus the process of the ageing of society within the city is linked with suburbanisation.

The municipal authorities are trying to prevent this problem through various efforts. In the area of the Old Town we can observe a tendency to develop "empty townhouses" that do not have any known owners (Fig. 9–14). The following vandalised townhouses can serve as an example: the townhouse on the corner of Klasztorna Street and the Small Market Square, as well as the one on the corner of Sienkiewicza and Mickiewicza streets. The renovation, demolition and modernisation work was carried out by Towarzystwo Budownictwa Społecznego (Fig. 9, 10, 14, 15 – townhouse on the corner of Sienkiewicza and Mickiewicza streets) [23].

The area of the Old Town is inhabited by 2018 people, which constitutes 5,12% of the overall population of the city. The population structure is characterised by a decreasing number of people in pre-working and working age. This tendency has been stable since 2010. At the same time the amount of people past working age is increasing, which is proof of the ageing of the city's residents [18, p. 63].

This area shows the highest possible exceeding of the values of indicators associated with reported crimes and offences in relation to the same data presented for the remaining parts of the city (it is over twice as high in comparison to the rest of the city). Taking into account the number of residents of the area, we can state that it is a situation that points to a very high grouping of pathological behaviour and criminal activity. The intensity of the phenomenon of unemployment is also higher than the average for Oświęcim and points to a high scale of the problem. Municipal efforts associated with the development of abandoned townhouses by TBS in the area of the Old Town will positively affect the population structure of the people who inhabit this area.

4. Expanding the rest, recreation, cultural and tourism-related offering of the city in the areas along the Soła River

Access to green areas that offer rest and recreation has considerable influence on the quality of life within a city. Oświęcim is characterised by a high amount of greenery which is present in various forms: parks, belts that accompany vehicular circulation routes, as well as a number of recreational areas, for instance the boulevards on the Soła River (Fig. 15).

Fig. 15. The development of the area along the Soła River (photo by K. Paprzyca, 2015)

The municipal government's plan of action is directed at the activation of the area of the Old Town, along with the recreational areas along the Soła River. The development of the urban space in such a manner that tourist traffic, which is currently mainly concentrated at the Auschwitz Museum, will reach the Oświęcim Old Town to a greater degree, is an important aspect. Tourists will be able to visit its historical sites and pleasantly spend their time. The plan is associated with the establishment of a space that has a symbolic, educational and cultural character along the axis that leads from the Place of Remembrance to the Old Town, so that it can affect the diversification of the attractiveness of the city in terms of tourism. The area will undergo urban regeneration to serve the needs of the public, in addition to activation through making it "attractive" to visitors – tourists. The functional and utilitarian diversity, which the Old Town is meant to be characterised by, will lead to an increase of the attractiveness of the area.

The Old Town is a key part of the city from the point of view of the possibilities of developing tourism, especially in terms of the leisure industry. The amount of tourists that visits the National Auschwitz-Birkenau Museum in Oświecim is increasing (2015 – 1,7 million), which provides the opportunity for the area around the Soła River to become an eagerly visited place. Attracting tourists to the Old Town can be performed through the Bridge of Ghosts designed by Jarosław Kozakiewicz [24], which is meant to be built in the future. It depicts the link between the Auschwitz camp, which has become a symbol of remembrance, and the city of Oświęcim, which symbolises everyday life [25] (Fig. 16).

The design is in line with the programme of the development of the areas along the Soła River in connection with the Old Town, called the Oświęcim Meeting Space [26], which is meant to serve residents and tourists visiting the National Auschwitz-Birkenau Museum.

Another project associated with dialogue above cultural and state boundaries that this meant to build peaceful relations that is highly significant to the City of Oświęcim is the multicultural Oświęcim Life Festival, which attracts crowds of people of all ages every year. Up to this point the festival attracted world-famous performers; Elton John (2016), Queen

and Adam Lambert (2016), Eric Clapton (2014), Soundgarden (2014), Jesse Ware, UB40, Sting (2013) and Peter Gabriel (2012) (Fig.17) [27].

Fig. 16. Bridge of Ghosts | Histories of Oświęcim (source: [24])

Fig. 17. Life Festival (source: [27])

Utilising the historical and cultural potential of the city to increase tourist traffic requires the allocation of resources in development. The Hampton by Hilton Hotel is being built near the entryway to the Old Town, near the Piast Castle. (Fig. 18, 19). It is to house 40 thousand guests per year. The building is being built at the site of the no-longer existing townhouse and the Jakub Haberfeld Vodka and Liquors Factory. "This is going to be the first such building in the centre – Marcin Susuł from the Oświęcim-based company Susuł&Strama Architekci, which is responsible for the design of the building, told "Gazeta Krakowska". The hotel is

Fig. 18, 19. Entryway to the Oświęcim Old Town, a view of the Hilton Hotel being built (photo by K. Paprzyca, 2018)

not simply meant to expand the accommodation offering – the developer is hoping that the building will lead to increasing the liveliness of this part of the Old Town, which has slightly decreased in attractiveness in recent years. The modern hotel is not only meant to generate greater traffic in the entire region of the city, but mainly to attract additional developers" [28].

5. Conclusions

In order for cities to develop, creating an image of a place with a rich offering of cultural, utilitarian and aesthetic qualities is necessary. This offering should be subjected to a constant valorisation. The functioning of a city is a living organism that is subjected to transformation.

The offering of the increase of the cultural and utilitarian qualities of the city of Oświęcim is a form of an intervention into urban space. It is meant to improve the image of the city, making it more attractive to residents, tourists and entrepreneurs. It positively affects transformation processes, including spatial, economic and social ones.

It includes urban renewal through:

- improving the quality of the public space within the city, as well as the regeneration and modernisation of buildings and technical infrastructure in the area of the Old Town;
- expanding the rest, recreation and tourism-related offering of the city along the Soła River;
- creating a meeting space of a symbolic, educational, cultural and recreational character along the axis between the Place of Remembrance and the Old Town, whose goal will be to connect the past with the future;
- upholding the identity and culture of the place.

The concept of the development of the city of Oświęcim through an offering of the increasing of cultural and utilitarian qualities is attractive. This can be proven by examining appropriate indices: the changing number of residents, the number of visitors and the level of the residents' income, the level of property prices and capital invested in this sector. Over the past year the city has advanced in the "Forbes" ranking, moving from eighth to third position. "The greatest number of new projects is being built on plots belonging to Synthos (the former Dwory Chemical Company), which today function as a sort of industrial cluster. Around a dozen companies have built their buildings here and two additional ones are building their factories: IPB from Italy (a metal pressing plant) and Sofir (press and production line renovation). Synthos offers industrial areas with full access to building services, located alongside access roads, with a combined surface area of 100 hectares. In Oświecim new companies are also developing in the 45 hectare Municipal Economic Acitivty Zone, as well as at the Oświęcim Entrepreneurship Incubator. Over 50 companies have decided to place their buildings there. The zone is located on the outskirts of Oświęcim, near road no. 44 (Tychy-Zator-Skawina-Kraków), which is a bypass for its centre. The further development of Oświecim will be made possible by the recently started expansion of the urban zone (its par t is covered by privileges of the special economic zone). The project will cost 16,4 million PLN, of which 9 million will be given to city from the Lesser Poland Regional Operational Programme. Most attractive cities for business – Forbes" [29].

References

[1] Baranowski A., *Projektowanie zrównoważone w architekturze*, Wydawnictwo Politechniki Gdańskiej, Gdańsk 1998.
[2] Chmielewski J.M., *Teoria urbanistyki w projektowaniu i planowaniu miast*, Warszawa 2001.
[3] Cichy-Pazder E., *Humanistyczne podstawy kompozycji miast*, Politechnika Krakowska, Kraków 1998.
[4] Floryda R., *Narodziny klasy kreatywnej*, Narodowe Centrum Kultury, Warszawa 2010.
[5] Gehl J., Życie miedzy budynkami. Użytkowanie przestrzeni publicznych, Wydawnictwo RAM, Kraków 2009.
[6] Gzell S., *Fenomen małomiejskości*, Instytut Urbanistyki i Planowania Przestrzennego Politechniki Warszawskiej, Warszawa 1987.
[7] Gyurkovich J., *Znaczenie form charakterystycznych dla kształtowania i percepcji przestrzeni*, Monografia 258, Politechnika Krakowska, Kraków 1999.
[8] Gyurkovich M., *Hybrydowe przestrzenie kultury we współczesnym mieście europejskim*, Monografia 438, Kraków, Politechnika Krakowska, Kraków 2013.
[9] Kosiński W., *Miasto i piękno miasta*, Politechnika Krakowska, Kraków 2011.
[10] Kobylarczyk J., *Ocena jakości środowiska zamieszkania w wybranych miastach województwa podkarpackiego po okresie „transformacji" w pierwszej dekadzie XXI w.*, Monografia 434, Politechnika Krakowska, Kraków 2013.
[11] Paprzyca K., *Harmonizowanie rozwoju urbanistycznego terenów miejskich – wybrane zagadnienia*, Monografia 417, Kraków 2012.
[12] Paszkowski Z., *Miasto idealne w perspektywie europejskiej i jego związki z urbanistką współczesną*, Universitas , Kraków 2011.
[13] Schneider-Skalska G., *Kształtowanie zdrowego środowiska mieszkaniowego. Wybrane zagadnienia*, Monografia 307, Politechnika Krakowska, Kraków 2004.
[14] Seruga W., *Warunki i kryteria kształtowania niskiej intensywnej zabudowy mieszkaniowej*, Politechnika Krakowska, Kraków 1984.
[15] Wallis A., *Miasto i przestrzeń*, Warszawa 1977.
[16] *Miasta na drodze do zrównoważonego rozwoju – PwC*, https://www.pwc.pl/pl/biuro-prasowe/.../miasto_szans_opisy_dobrych_praktyk.pdf (access: 15.01.2018).
[17] http://www.polskawliczbach.pl/Oswiecim (access: 15.01.2018).
[18] http://www.rpo.malopolska.pl/download/programregionalny/rewitalizacja/001_Miasto_Oswiecim.pdf (access: 15.01.2018).
[19] http://web.um.oswiecim.pl (access: 15.01.2018).
[20] http://poland24h.pl/atrakcje/Oswiecim-990.html (access: 15.01.2018).
[21] http://www.wsjp.pl/index.php?.id_hasla=3609 (access: 15.01.2018).
[22] http://www.ock.org.pl/Tydzien-Kultury-Beskidzkiej-184.html (access: 15.01.2018).
[23] http://dziennikpolski24.plwww.dziennikpolski24.pl/tag/otbs-oswiecim-ratuje-stare-miasto (access: 15.01.2018).

[24] https://oswiecimskiehistorie.wordpress.com/2012/12/13/most-duchow (access: 15.01.2018).
[25] http://www.dziennikpolski24.pl/artykul/3851337,niezwykly-most-polaczy-auschwitz-z-oswiecimiem,id,t.html (access: 15.01.2018).
[26] http://obserwatorium.miasta.pl/.../4-Oświęcimska-Przestrzeń-Spotkań-Kongres-Rewitalizacji (access: 15.01.2018).
[27] http://muzyka.onet.pl/life-festival-oswiecim (access: 15.01.2018).
[28] http://sztuka-architektury.pl/article/.../hampton-by-hilton-na-starym-miescie-w-oswiecimiu (access: 15.01.2018).
[29] https://www.forbes.pl/rankingi/miasta-najbardziej-atrakcyjne-dla-biznesu/nrlbnyg (access: 15.01.2018).

TECHNICAL TRANSACTIONS 3/2018
CZASOPISMO TECHNICZNE 3/2018
CHEMISTRY
DOI: 10.4467/2353737XCT.18.036.8298
SUBMISSION OF THE FINAL VERSION: 20/2/2018

Justyna Jamróz
Dawid Jankowski (jankowski@indy.chemia.pk.edu.pl)
Chair of Chemical and Process Engineering, Faculty of Chemical Engineering and Technology, Cracow University of Technology

Characteristics of producing ethyl alcohol

Charakterystyka wytwarzania alkoholu etylowego

Abstract

The paper presents the introduction to problems related to the distilling industry in Poland and its current situation. The theoretical issues related to the production of ethyl alcohol are discussed. The raw materials used and their impact on production efficiency, process steps and their conditions and quality requirements for alcoholic beverages based on ethanol are characterised. The main causes of the problem of distillery in Poland are seen in the import of cheap rectified and dehydrated spirits from other EU countries, which results in a fall in market prices that do not provide for reimbursement of production costs and increase of excise duty in 2014, which significantly contributed to the closure of the distillery. At present, there are around 100 of them in the country, most of which are made of distillates for bio ethanol production.

Keywords: ethyl alcohol, ethanol production, fermentation

Streszczenie

W pracy przedstawiono wprowadzanie do problematyki związanej z przemysłem gorzelniczym w Polsce oraz jego aktualną sytuację. Opisane zostały zagadnienia teoretyczne związane z procesem wytwarzania alkoholu etylowego. Scharakteryzowano stosowane surowce i ich wpływ na wydajność produkcji, etapy procesu wraz z ich warunkami oraz wymagania jakościowe napojów alkoholowych, otrzymywanych na bazie etanolu. Główne przyczyny problemów gorzelnictwa w Polsce upatruje się w imporcie taniego spirytusu rektyfikowanego i odwodnionego z innych krajów UE, co powoduje spadek cen rynkowych, które nie zapewniają zwrotu kosztów produkcji oraz podwyżki akcyzy w 2014 roku, która znacznie przyczyniła się do zamykania gorzelni. Aktualnie w kraju pozostało ich ok. 100, z których większość zajmuje się wytwarzaniem surówki do produkcji bioetanolu.

Słowa kluczowe: alkohol etylowy, produkcja etanolu, fermentacja

1. Introduction

Diluted solutions of ethyl alcohol obtained by fermentation were well-known to man in antiquity. Recent archaeological discoveries indicate that alcoholic beverages were produced during the Neolithic Age. The first ethanol distillation is attributed to the alchemist Geber of the VIII century. The description of the distillation apparatus and the way of obtaining alcohol from wine appeared in XIII century by prof. Arnold de Villeneuve of the University of Montpellier. In the XV century in Poland, herbs and flowers were extracted with alcohol, from which cosmetics and medicinal mixtures were made. An opinion on the beneficial effect of alcohol on human health has contributed to the increase of its consumption, production and development of craft distillery [1–3].

The intensive development of craft distillery in Poland took place in the XIX century, which was connected with the general development of the technique. Pistorius constructed the first distillation apparatus, a prototype of today's stripper columns, which allowed to produce high-percentage alcohol. Until 1830, the main raw material for the production of vodka in Poland was rye. In 1803, the method of producing potato alcohol was known, which allowed for about 7 times more alcohol than from rye, but only a few distilleries used this raw material. In 1866, the first excise tax was imposed, regulating the production of spirits in Polish territories. Before that, according to the statute of Jan Olbracht of 1546, every citizen of the country was free to sell and produce alcoholic beverages. In 1919, full control over the production and marketing of spirits was established in Poland. The State Spirit Monopoly was established, whose activity continued uninterrupted until 1939. It coordinated the production of spirits in the country, and also set the price of distillates for the individual provinces, as well as a production limit for distilleries. In this way, the State's Monopoly strictly regulated the production of spirits in the country, adjusting its size to the needs and possibilities of export. In the interwar period, there were about 1400 agricultural distilleries, which constituted an integral part of private property. In addition to the agricultural distilleries, 10 fruit distilleries, 5 molasses and 7 yeasts were in operation.

After World War II, in Poland, there were about 1200 agricultural distilleries good for commissioning and 15 industrial distilleries, adapted for the treatment of molasses. Nowadays, agricultural distilleries operate according to the demand reported from the fuel and spirits sector. In 1995, there were about 950 distilleries, and in 2011, their number was reduced to 173 [3–5].

2. Production of ethyl alcohol

2.1. Raw materials for the production of ethyl alcohol

The choice of raw material for ethanol production depends on the type and quantity of hydrocarbons fermented as well as their price. Their costs account for 76% of ethanol production costs. The yield of ethyl alcohol obtained from particular vegetable raw materials strongly depends on its sort (Table 1) [6].

Table 1. Efficiency of ethanol from individual vegetable raw materials [7]

Raw material	Rice	Corn	Wheat	Rye	Molasses	Oat	Potatoes	Apples
Yield, $dm^3 \cdot t^{-1}$	350–500	360–400	350–390	310–370	350	240	80–120	35–60

Raw materials for ethanol production are divided into two groups: straight -mono and -disaccharide sugars (e.g. glucose, fructose, sucrose) and polysaccharides (starch, glycogen). The first one of them undergoes direct fermentation, under the influence of some bacteria and yeast, the second are hydrolysed to fermentable sugars (-mono- and disaccharides). Simple sugars are present in vegetables, fruit, sugar cane, molasses and waste of food industry. Polysaccharides are found in potatoes, topinambur, manioc, chicory and waste of lignin-cellulose [6, 7].

Potatoes and cereals contain starches, which are their backup material for sourcing energy. This polysaccharide consists of glucose residues linked by a 1,4-α-glycosidic bond. Easy breakdown of starch has been used for centuries in the preparation of the mash for the production of ethyl alcohol. In traditional distillery craft, the source of enzymes for starch hydrolysis is malt (germinating grains, e.g. barley). In the germination process, enzymes of α-amylase and β-amylase are formed that break down the starches. The malt has now been replaced by enzyme preparations derived from such microorganisms as e.g. bacteria Bacillus subtilis or Aspergillus niger [8].

Molasses is a by-product of sugar production from sugar beet. It is a good raw material for ethanol because it contains a lot of carbohydrates, which undergo direct fermentation. Therefore, there is no need to use the evaporation stage (the process of starch release from the raw material) and the enzyme preparations in the technological process, which are necessary for the fermentation of starch raw materials [9].

2.1.1. Characteristics of yeast for distillery industry

The oldest known humanity biotechnology process is alcoholic fermentation with yeast share. Ethanol is produced by yeasts that live under anaerobic conditions and some bacteria species. The most commonly used species of yeast is Saccharomyces cerevisiae (Tab. 2). In the case of bacteria, important meaning for industry only have those that break down the sugars into ethyl alcohol such as e.g. Sarcina ventriculi [10].

Table 2. Yeast and bacterial species producing ethyl alcohol [10]

Yeast	Substrates	Species of bacteria	Substrates
Saccharomyces cerevisiae	glucose, fructose, galactose	*Zymomonas mobilis*	glucose, fructose, saccharose
Saccharomyces uvarum	saccharose, maltose, maltotriose	*Clostridium thermocellum*	glucose, cellobiose, cellulose
Saccharomyces diastaticus	glucose, maltose, starch	*Clostridium thermohydrosulphuricum*	glucose, xylose, cellobiose, starch
Saccharomyces rouxii	glucose, fructose, maltose, saccharose	*Thermoanaerobium brockii*	glucose, saccharose, cellobiose, starch
Schwanniomyces alluvis	dextrin, starch	*Thermobacterioides acetoethylicus*	glucose, saccharose, cellulose

Depending on the type of microorganism we use in the fermentation process, we can get different yields and reaction rates. The yield of ethanol produce from glucose for Z. mobilis and S. cerevisiae yeast is similar (Tab. 3). There are no significant differences between these microorganisms in the growth rate of biomass and ethanol production by yeasts and bacteria. The discrepancies between them appear with the ethanol production rate. The yeast produces ethanol at almost five times the speed, which is their main advantage. Also, in the case of glucose assimilation, the yeast is much faster than bacteria. As regards the type of fermentation process, it is best to use yeast for periodic fermentation because they are resistant to higher ethanol concentrations than bacteria. For continuous fermentation, it is better to use bacteria, because they are resistant to higher concentrations of ethanol in this type of process (Tab. 3) [10].

Table 3. Comparison of ethanol synthesis with bacteria (Z. mobilis) and yeast (S.cerevisiae) [10]

Kinetic parameter	Unit	*S. cerevisiae*	*Z. mobilis*
Max rate specific biomass growth	$1 \cdot h^{-1}$	0.41	0.43
Rate specific of assimilation of glucose	$(g_{microorg.} \cdot g_{glucose}) \cdot h^{-1}$	10.5	1.75
Rate specific ethanol production	$(g_{mikroorg.} \cdot g_{ethanol}) \cdot h^{-1}$	5.67	0.67
Yield of ethanol production on glucose	$g_{ethanol} \cdot g_{glucose}^{-1}$	0.47	0.43
Percentage yield	%	92	85
Max concentration of ethanol – in continuous fermentation – in batch fermentation	%	 do 5 do 20	 5.5–6.0 do 10

Yeasts of Saccharomyces cerevisiae are most commonly used in brewing and bakery craft. They can take different shapes depending on the age, the type of culture and the amount of food. These yeasts are larger in size than bacteria which are about 2–8 μm in width and 3–10 μm in length. Saccharomyces cerevisiae live in base, which contain simple sugars. They lead alcoholic fermentation mainly under anaerobic conditions [11]. Yeasts of this type ferment and assimilate most of the sugars. In the distillery craft a number of strains of Saccharomyces cerevisiae are used, which is associated mainly with their ability to rapidly alcohol ferment, resistance to alcohol concentration of 10–12% and high osmotic pressure. The most efficiently they fermented at 30–33°C. These yeasts have the ability to ferment glucose, fructose, sucrose, maltose, few strains are capable of fermenting raffinose, lactose and mannose [12].

Mixed cultures of yeasts are becoming popular, with individual strains having complementary properties each other or genetically modified to obtain the best possible properties. The used gamma rays for two types of Saccharomyces cerevisiae strains, Persian culture used in Tehran (PTCC 5269) and Armenian culture, strains resistant to high temperatures of 38–42°C and ethanol content up to 25% can be obtained. The thus microorganisms obtained allow the fermentation process to remain longer, without replacing yeast with new ones [12].

2.2. Preparation of ethyl alcohol

The stages of ethanol production vary, depending on the type of raw material used. In the case of cereal and potato distillery craft, it is necessary to distribute the polysaccharides present in the substrate in the evaporation and mashing process. This stage of alcohol production does not occur in the processing of fruit or molasses, which build from simple sugars. The remaining stages are the same for all raw materials and include the fermentation and distillation processes in which we receive distillate. In the final stage the distillate is purified in the process of rectification to spirit (Fig. 1) [13].

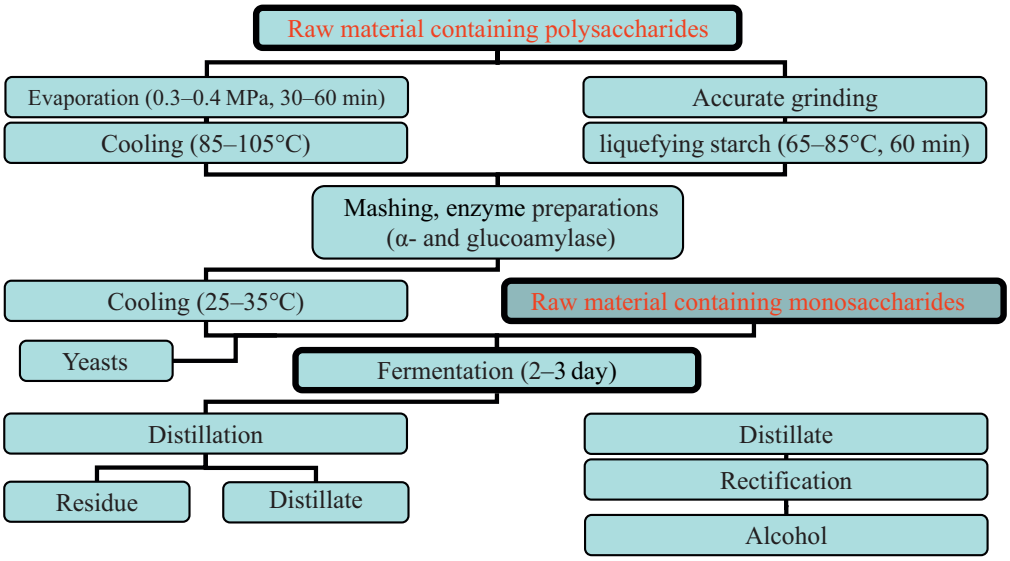


Fig. 1. Diagram of the process of making distillate and alcohol [13]

2.2.1. Preparation of the mash

Evaporation of raw material is the first step in the production of ethyl alcohol. Its purpose is to relax the intercellular substances, break the cell membranes and completely release and glue the starch. The amount of water used in this process must be approx. 4 times greater than the amount of starch that allows it to be liquefies. The evaporation process takes about an hour, at 120–140°C, at a pressure of 0.3–0.4 MPa. Liquified mash enzymatic preparations are used to reduce their viscosity and facilitate the saccharification of the starch. Enzymes break down chains of amylose, amylopectin and fermentable sugars, mainly to glucose [14].

The cold mashing process used by some distilleries involves the unpressurised release of starch. The raw material is finely crushed to about 1.7 mm, which facilitates the access of water to the starch grains. As a result, it is possible apply milder process conditions. The duration of cold mashing is approx. one hour at 65–85°C. Enzymatic preparations are also used in this process. The next step is the fermentation of the sweet mash with the use of distillery yeast [15].

The newest method of ethyl alcohol production is the simultaneous process of hydrolysis and fermentation of SSF (Simultaneous Saccharification and Fermentation), which allows for gradual release of glucose during fermentation. Starch is broken down by enzymes at temperatures below the value of the gelatinisation, significantly reducing energy consumption, the temperature throughout the process does not exceed 50–55°C [12].

2.2.2. Alcoholic fermentation

Alcoholic fermentation capacity demonstrated by the yeasts from the ascomycetes group. This process consists in the anaerobic decomposition of simple sugars into ethyl alcohol and carbon dioxide with a simultaneous release of energy. In fact, its course is more complex (Fig. 2). During fermentation, many by-products are formed, and the process itself depends on many factors: temperature, chemical composition of the base and pH of the environment. If oxygen has access to the base, yeast does not produce ethyl alcohol, only breathe oxygen and multiply. The concentration of ethyl alcohol (> 20%) produced by the fermentation process inhibits the yeast's activity [16].

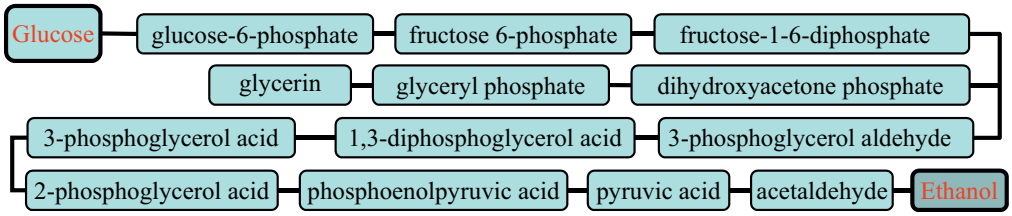


Fig. 2. Diagram of the glucose conversion to ethyl alcohol [11]

The alcoholic fermentation process can be represented by a relatively simple general equation (1). From one mole of glucose produces two moles of ethanol and carbon dioxide, the reaction proceeds with the release of energy of about 234.5 kJ. The microorganisms consume approximately 67 kJ to produce two moles of adenosine triphosphate (ATP) from adenosine diphosphate (ADP) with the participation of residuals of orthophosphoric acid (P_i) [17].

$$C_6H_{12}O_6 + 2ADP + 2P_i \rightarrow 2C_2H_5OH + 2ATP + 2CO_2 + 167.5\ kJ \quad (1)$$

In the first stage of fermentation, hydrolytic enzymes, sucrose and maltase break down sucrose and maltose (disaccharides) into simple sugars. Then the enzyme ATP (adenosine triphosphoric acid) is connect to the simple sugars and form glucose phosphate and fructose phosphate. In the next step, fructose phosphate, after the subsequent addition of the phosphate acid, is broken up into two triose: phosphoglycerol aldehyde and phosphorodihydroacetone. These sugars have the same summary formula but different structure and properties. Under the influence of the isomerase enzyme, the ketone is converted into an aldehyde which is then oxidised to phosphoglyceric acid. Part of the aldehyde, under the action of dehydrogenase, changes into glycerol phosphate, which then turns into glycerol. A large amount of aldehyde

molecules is converted to phosphoglyceric acid, which is further converted to phosphorpyruvic acid by donating part of phosphate to the enzyme adenosine diphosphate acid [4, 10, 11].

The initial fermentation step is identical like glycolysis of glucose into pyruvate (2), which could be converted into acetyl coenzyme A (acetylCoA) or oxaloacetate. In case of oxygen deficiency, the yeast metabolises pyruvate to ethanol [17].

$$C_6H_{12}O_6 + 2P_i + 2ADP + 2\,NAD^+ \rightarrow 2CH_3COCOO^- + 2ATP + 2NADH + 2H^+ \qquad (2)$$

As a result of non-oxidative decarboxylation, pyruvic acid is transformed into acetaldehyde (3). The reaction takes place under the influence of two pyruvate decarboxylase enzymes and thiamine pyrophosphate, which separate the carboxylic group in the form of CO_2 from the acid [17, 18].

$$2CH_3COCOOH \rightarrow 2CH_3\,CHO + 2CO_2 \qquad (3)$$

In the last step, the acetaldehyde is reduced to ethanol in the presence of alcohol dehydrogenase (4). During the oxidation of 3-phosphoglyceric aldehyde to 1,3-diphosphoglyceric acid is formed dinucleotide (NADH), which is a hydrogen donor. The main products of the process are ethyl alcohol and carbon dioxide [17, 18].

$$2CH_3CHO + 2\,NADH + 2H^+ \rightarrow 2CH_3CH_2OH + 2NAD^+ \qquad (4)$$

A lot of byproducts are formed in the alcoholic fermentation process, which can improve the taste of the final product or significantly worsen it. Their presence and amount in alcoholic drinks depend on the type and quality of raw materials, methods and conditions of the technology used. The storage of the obtained alcohol solution is also important for him quality. The byproducts of the fermentation process include aldehydes, esters, methanol, glycerin, organic acids and higher alcohols [19].

Carbonyl compounds such as methanal, ethanal, propanal, butanal, pentanal, hexanal, 2-methylpropanal, acetone, glyoxal and furfural worsen the organoleptic properties of alcoholic beverages. Their source is fat decompose reactions. In controlling the synthesis of aldehydes, the most important role is played by the temperature and pH of the fermentation environment [20, 21]. The agricultural distillates from molasses may have several times higher concentrations of acetaldehyde in comparison with cereal distillates. Part of the hydrocarbons can be converted by bacteria to aldehydes. In the final product, as a result of condensation of aldehydes with alcohols, they may appear in a small amount of acetals [22]. In higher concentrations of acetals, the smell of the beverage significantly worsens [19, 24].

Methyl alcohol is formed from pectin by the decomposition into methanol via pectic acid. As a result of the enzymatic hydrolysis of the polygalacturonic acid methyl ester, free carboxyl groups and methanol are formed [4, 23]. Pectin is a mixture of carbohydrates and is usually found in generative organs of plants (fruits, seeds). Therefore, the greatest amount of methanol will be composed of spirits derived from fruits, mainly apples and plums. Higher

alcohols adversely affect the taste of alcoholic beverages. They are produced in the amino acid dezamination process (isoleucine, leucine and valine) [4, 10, 11]. They are characterised by a higher boiling point than ethanol, which is why they can be removed relatively easily from the product. The following fermentation products appear in the largest concentrations: propanol, butanol, pentanol and 3-methylbutanol. Their formation depends on the type of yeast and raw material used and the presence of oxygen, nitrogen and sulphur in the solution. As in the case of esters, a low content of fusels in an alcoholic beverage may have a beneficial effect on its taste [4, 19, 23].

Esters are formed from the combination of activated fatty acids and alcohols formed during the fermentation process. Despite their presence in trace amounts in final products, they give them a characteristic smell [19]. The synthesis of esters can be controlled to some extent by the change conditions run of the fermentation process: hydrostatic pressure, reactor size, nitrogen level, wort oxygenation and temperature, and selection of raw materials. The influence of yeast race on ester production and their relation to each other is extremely important. Proper selection of yeast may increase or decrease nearly five times the amount of esters in the final product [24].

Sulphur compounds and organic acids may also be present in the fermentation product. In drinks, most common are methanethiol, dimethyl sulphide, dimethyl disulphide, dimethyl trisulfide, methyl thioesters and thiols, which are formed from the reaction of yeast metabolism or non-enzymatic reactions [25, 26]. Acetic, lactic and butyric acids may be produced (as a byproduct) by some yeast. The increase in temperature and pH of the fermentation environment is conducive to the development of microflora of bacteria responsible for the synthesis of organic acids. Their excessive amount and reduced pH, the overall efficiency of the fermentation process may be reduced [19,27].

The raw spirit obtained from the fermentation process is further purified. Depending on the degree of purity of the final product, ordinary, selective, luxury and neutral spirits are distinguished [4, 10, 11].

2.3. Separation of ethyl alcohol

2.3.1. Distillation

Ethyl alcohol can be extracted from the fermentation product using a classic distillation process that involves evaporation of the liquid and condensation of the resulting vapour. The mash is heated to about 78.3°C, where the ethanol is evaporated. Vapours are condensed by lowering their temperature. The raw spirit obtained after the distillation process is a mixture of many components, as the other compounds contained in the feed evaporates together with the ethanol. Separation of these substances from ethyl alcohol can be achieved using the rectification process. Ethyl alcohol forms azeotropic mixture with water, which cannot be separated by straight distillation [4]. In agricultural distillery, multiple straight distillations are carried out, producing 65–88% of crude spirit, which also includes volatile fermentation products [3].

Straight distillation is divided into equilibrium, periodic and continuous distillation. Equilibrium distillation is runs continuous and is mainly used to separate low volatility substances from high volatility substances. Periodic distillation is used for the production of special alcoholic beverages e.g. fruit vodkas. In the continuous distillation, due to the short residence time of the substance in the elevated temperature zone, the decomposition and polymerisation of the particles are avoided [28–30].

2.3.2. Rectification

Rectification is a separation process that uses the difference in volatility of individual components of liquid mixtures. The separation occurs in a column, when it flows countercurrent and without diaphragm flows down the liquid stream, and the vapour to top of the column. A higher temperature is at the bottom of the column, while the lower one is at the top. The steam passing through the column is enriched in a more volatile component and the liquid in the less volatile component has a higher boiling point [31].

The rectification can be carried out both periodically and continuously. The most common way is a continuous process, and periodic rectification is used to separate small amounts of mixtures. In the continuous process, the raw material is fed continuously to the column powered shelf, which is divided into the reinforcement and the stripping part. Continuous distillate and depleted liquid of constant composition are also obtained [29].

Periodic rectification consists of a one-time filling of the boiler with a liquid mixture, which is heated to the boil. Further distillate fractions are collected at the top of the column from the condenser. As a result, we get fractional mixtures that differ in condensing temperatures. The rectification process produces purified spirits that contain more than 96% ethyl alcohol but less than 97.2%. Hydro-selection or molecular sieves are used to obtain 100% of the spirit.

3. Quality requirements

In the process of alcoholic fermentation, in addition to ethyl alcohol, other compounds are formed, which give the spirits a characteristic flavour and aroma as well as affect their quality. Raw spirit should meet the requirements set out in Table 4 in accordance with the agricultural distillery quality standard [32].

The standard of the quality of rectified alcohol divides it into three species: luxury, perfect and ordinary (Table 5). In order to obtain pure vodkas, the spirit with demineralised water should be mixed in the characteristic proportions for the specific product. Species of vodka also include juices, macerates, distillates, dyes, sugar syrups and other flavours. The requirements set by the quality standard for pure spirits are given in tables 6 and 7 quality requirements for the species of vodkas [13, 33–35].

Table 4. Requirements for the quality of agricultural distillate [32]

Features	Agricultural distillate			
	molasses	cereal	potato	fruit
Transparency	transparent, without sludge and turbidity, mechanical impurities are allowed, drooping to the bottom after 2 hours			
Colour	colourless, yellowish or greenish shades are allowed			colourless, yellowish shade is allowed
Smell and taste	specific, characteristic of the raw material used, without extraneous odours and odours			
Power, % vol	≥ 88			≥ 65 and ≤ 86
Fusels[1], g*dm^{-3} 100%	*	≤ 5.0 in spirits for the production of okowit, in others *	*	≤ 5.0
Aldehydes[2], g*dm^{-3} 100%	≤ 0.3	≤ 0.1		≤ 0.2
Molasses alcohol, g*cm^{-3} 100%	*	≤ 0.08 in spirits for the production of okowit, in others *	*	≤ 0.8
Acids[3], g*dm^{-3} 100%	≤ 0.1	≤ 0.08		≤ 0.2
Hydrogen cyanide, mg*dm^{-3} 100%	*			≤ 0.3 in stone's fruit distillate, in others *
Dry residue after evaporation, g*dm^{-3} 100%	< 0.08	*		
Presence of furfural	unacceptable in the production of vinegar, in others *			
Pyridine, g*dm^{-3} 100%	< 0.02 in the distillate for the production of vinegar, in others*			

* is not normalised

[1)] in calculated on a mixture of isoamyl and isobutyl alcohol

[2)] in calculated on the acetaldehyde

[3)] in calculated on acetic acid

Table 5. Alcohol quality requirements [33]

Features	Rectified alcohol		
	ordinary	perfect	luxury
Ethanol, % vol	≤ 96	≤ 96.5	
Fusels [1], g*dm^{-3}100%	≤ 0.005	≤ 0.002	≤ 0.001
Aldehydes [2], g*dm^{-3}100%			
Esters [3], g*dm^{-3}100%	≤ 0.05	≤ 0.03	
Acids[4], g*dm^{-3}100%	≤ 0.0020	≤ 0.015	
Methyl alcohol, g*100 cm^{-3}100%	≤ 0.05		≤ 0.03
Lang's attempt, min	≥ 18	≥ 20	
Dry residue, g* dm^{-3}100%	≤ 0.015	≤ 0.010	
Furfural	unacceptable		
Volatile nitrogen base [5], g* dm^{-3}100%	< 0.001		

[1)] in calculated on isobutyl alcohol
[2)] in calculated on the acetaldehyde
[3)] in calculated on ethyl acetate
[4)] in calculated on acetic acid
[5)] in calculated on nitrogen

Table 6. Quality requirements of pure vodkas [24]

Vodkas, %vol		Aldehydes[1]	Fusels [2]	Acidity	Dry residue	Methanol, g*100 cm^3 100%
no more than, g*dm^{-3} 100%						
Ordinary	38–70	0.008	0.005	*	*	≤ 0.1
Perfect		0.004	0.003	*	0.035	≤ 0.05
Luxury		0.003	0.002	*		≤ 0.03
Mixed		like for pure vodkas, made from a lower alcohol quality				

* is not normalised
[1)] in calculated on the acetaldehyde
[2)] in calculated on amyl alcohol
[3)] in calculated on acetic acid

Table 7. Quality requirements for species vodkas [35]

Species products	Power, %vol	Extract, g*dm^{-3}	Fusels [1], g*dm^{-3} 100%	Methanol, g*100 cm^{-3} 100%	Hydrogen cyanide, mg*dm^{-3}
Dry regular	30–50	do 50	≤ 1.5	≤ 0.25	≤ 3
Dry natural	30–75		≤ 5	≤ 0.4 (slivovitz –0.8)	
Dry natural mixed			≤ 4	≤ 0.4	
Semi-dry	30–45	51–120	≤ 1.5	≤ 0.25	
Semi-sweet		121–220			
Sweet		221–330			
Liqueur	25–45	> 330			
Creams	18–25	> 400	*		
Cocktails		do 400	≤ 1.5		
Aperitifs	18-35				

* is not normalised

[1)] in calculated on amyl alcohol

4. Summary

- Current problems of agricultural distillery craft in Poland include the import of cheap rectified and dehydrated spirits from other EU countries, which causes a dramatic drop in market prices that do not ensure the return of production costs.
- The increase in excise duty in 2014 has contributed to the closure of the distillery craft; at present, there are about 100 (95% of small and medium-sized companies) in the country, most of which are producing distillate for bioethanol production.
- In most cases, raw materials containing starches, such as cereals and potatoes, are used in ethanol production. Despite the evaporation and mashing stages, the yield of ethyl alcohol is higher.
- Bacteria and yeast can be used in the fermentation process. They have the same efficiency of ethanol production from glucose but vary in speed. Yeasts allow for five times faster fermentation.
- Saccharomyces cerevisiae is the most widely used yeast species in the distillery industry. Mixed or genetically modified strains that are resistant to up to 25% vol. are becoming more and more popular.

References

[1] Modrzejewski F., *Farmacja stosowana*, PZWL, Warszawa 1961.
[2] Skomorowski T., *Księga wynalazków, rękodzieł i przemysłu, Tom II*, Wyd. Przyroda i Przemysł, Warszawa 1875.
[3] Łączyński B., *Skrócony kurs gorzelnictwa rolniczego*, Wydawnictwo Sigma-NOT, Warszawa 1993.
[4] Jarociński J., Jarosz K., *Gorzelnictwo i drożdżownictwo*, Wydawnictwo Szkolne i Pedagogiczne, Warszawa 1994.
[5] Kupczyk A., *Perspektywy rozwoju polskich gorzelni rolniczych*, Rynki Alkoholowe 7/2007, 60–65.
[6] Duda-Chodak A., *Otrzymywanie i ocena jakości bioetanolu*, Katedra Technologii Fermentacji i Mikrobiologii Technicznej, 2010.
[7] Marczak H., *Znaczenie bioetanolu w wypełnianiu obowiązku stosowania paliw odnawialnych w transporcie*, Inżynieria Ekologiczna 28/2012, 102–110.
[8] Dynkowska W., Boros D., *Czynniki warunkujące przydatność ziarna różnych zbóż do produkcji energii odnawialnej*, Biuletyn Instytutu Hodowli i Aklimatyzacji Roślin 251/2009, 67–81.
[9] Kotarska K., Dziemianowicz W., *Wpływ różnych warunków fermentacji alkoholowej melasy na jej intensyfikację i jakość otrzymywanego spirytusu*, Żywność. Nauka. Technologia. Jakość 2(99)/ 2015, 150–159.
[10] Konkol S., *Podstawy mikrobiologii żywności*, ALMANACH Technologia Żywności produkcja piekarsko-ciastkarska 2/2010, 81–82.
[11] Mehdikhani P., Bari M., Hovsepyan H., *Screening of Saccharomyces cerevisiae for high tolerance of ethanol concentration and temperature*, African Journal of Microbiology Research 5(18)/2011, 2654–2660.
[12] Morton J., *Glycolysis and Alcoholic Fermentation*, Acts & Facts 9(12)/1980, 56–58.
[13] Ministerstwo Skarbu Państwa – Polski rynek wódki, http://www.msp.gov.pl/pl/przekształcenia/serwis-gospodarczy/wiadomosci gospodarcze/28430,Polski-rynek-wodki.html (access: 10.06.2017).
[14] Strąk E., Balcerek M., *Wybrane technologie wykorzystywane w przemyśle gorzelniczym*, Acta Sci. Pol., Biotechnologia14(1)/2015, 33–44.
[15] Gumienna M., Lasik M., Czarnecki Z., *Wykorzystanie odpadów przemysłu spożywczego do produkcji alkoholu etylowego*, Brom. Chem. Toksyk. 3/2009, 969–974.
[16] Bizukojć M., *Mikrobiologiczne i biochemiczne ujęcie wytarzania wybranych biopaliw*, http://www.proakademia.eu/gfx/ baza_wiedzy/ 30/m_bizukojc_ artykul3.pdf (access: 160.4.2017).
[17] Stevenson B.J., Liu J-W., Kuchel P.W., Ollis D.L., *Fermentative glycolysis with purified Escherichia coli enzymes for in vitro ATP production and evaluating an engineered enzyme*, Journal of Biotechnology 157/2012, 113–123.

[18] Zhou M., Zhou J., Tan M., Du J., Yan B., Wong J.W.C., Zhang Y., *Enhanced carboxylic acids production by decreasing hydrogen partial pressure during acidogenic fermentation of glucose*. Bioresource Technology 245 A/2017, 44–51.
[19] Biernacka P., *Metody kompleksowej analizy składu produktów ubocznych procesu fermentacji alkoholowej w półproduktach i produktach spirytusowych*, Rozprawa doktorska, WCh, KChA, Politechnika Gdańska 2012.
[20] Kłosowski G., Czupryński B., *Przyczyny powstawania związków karbonylowych w spirytusie surowym ze szczególnym uwzględnieniem aldehydu octowego*, Przemysł Fermentacyjny i Owocowo-Warzywny 5/1993, 8–10.
[21] Kłosowski G., Czupryński B., Kotarska K., Wolska W., *Charakterystyka zanieczyszczeń chemicznych obniżających jakość spirytusu surowego cz. I*, Przemysł Fermentacyjny i Owocowo-Warzywny 6/2003, 20–21.
[22] Łączyński B., *Jakie są przyczyny ponadnormatywnej zawartości w spirytusie surowym aldehydów i jak zjawisku temu można przeciwdziałać w warunkach gorzelni rolniczej?* Przemysł Fermentacyjny i Owocowo-Warzywny 2/1995,13–15.
[23] Kłosowski G., Czupryński B., Kotarska K., Wolska W., *Charakterystyka zanieczyszczeń chemicznych obniżających jakość spirytusu surowego cz. II,* Przemysł Fermentacyjny i Owocowo-Warzywny 9/2003, 37–38.
[24] Verstrepen K.J., Derdelinckx G., Dudour J.P., Winderickx J., Thevelein J.M., Petorius I.S., Delavaux F.R., *Flavor activ esters: adding fruitiness to beer*, Journal of Bioscience and Bioegineering 96/2003, 110–118.
[25] Moreira N., Mendes F., Pereira O., Pinho P.G., Hogga T., Vasconcelos I., *Volatile sulphur compounds in wines related to yeast metabolism and nitrogen composition of grape musts*, Analitica Chimica Acta 458/2002, 157–167.
[26] Swiegers J.H., Pretorius I.S. *Modulation of volatile sulfur compounds by wine yeast*, Applied Microbiology and Biotechnology 74/2007, 954–960.
[27] Schlegel H.G., *Mikrobiologia ogólna*, PWN, Warszawa 2003.
[28] Świca K., Lepa Ł., Zabek M., i inni, *Saccharomyces cerevisiae jako drożdże o szerokim zastosowaniu w życiu codziennym i przemyśle*, Katedra Technologii Fermentacji i Mikrobiologii Technicznej, 2010.
[29] Wilczura-Wachnik H., Okresowa kolumna rektyfikacyjna, Uniwersytet Warszawski.
[30] Bandrowski J., Troniewski L., *Destylacja i rektyfikacja*, Państwowe Wydawnictwo Naukowe, Warszawa 1980.
[31] Ziółkowski Z., Destylacja i rektyfikacja w przemyśle chemicznym, Wydawnictwo Naukowo-Techniczne, Warszawa 1978.
[32] PN-A-79523:2002. Destylat rolniczy.
[33] PN-A-79522:2001 Spirytus rektyfikowany.
[34] PN-A-79530:1995 Wyroby spirytusowe gatunkowe. Wspólne wymagania i badania.
[35] PN-A-79531:1995 Wyroby spirytusowe czyste. Wspólne wymagania i badania.

TECHNICAL TRANSACTIONS 3/2018
CZASOPISMO TECHNICZNE 3/2018
CHEMISTRY
DOI: 10.4467/2353737XCT.18.037.8299
SUBMISSION OF THE FINAL VERSION: 20/2/2018

Justyna Jamróz
Dawid Jankowski (jankowski@indy.chemia.pk.edu.pl)
Chair of Chemical and Process Engineering, Faculty of Chemical Engineering and Technology, Cracow University of Technology

Project of laboratory rectification column for periodic work

Projekt laboratoryjnej kolumny rektyfikacyjnej do pracy okresowej

Abstract

The paper presents a project of a rectification column working periodically for the production of ethyl alcohol. The project contains an assumption of a steady concentration of the distillate. Thee range of change in the reflux from 1.7 to 13 required to obtain that concentration was calculated. With 50 litres of feed containing 9 mol % ethanol, about 10 litres of distillate was obtained, which has 70 mol % ethanol. Ceramic Raschig rings with a diameter of 15 mm, which, in a column, occupy a height of 0.5 m, were used. The height of the entire column was 0.67 m.

Keywords: rectification column, ethanol, heat balance column

Streszczenie

W artykule przedstawiono projekt kolumny rektyfikacyjnej pracującej okresowo do otrzymywania alkoholu etylowego. W projekcie założono stałe stężenie destylatu i wyznaczono zakres zmiany liczby powrotu od 1,7 do 13, konieczny do jego otrzymania. Z 50 litrów surówki zawierającej 9% mol. etanolu uzyskano ok. 10 litrów destylatu o stężeniu 70% mol. Jako wypełnienie zastosowano ceramiczne pierścienie Raschiga o średnicy 15 mm, które zajmują w kolumnie wysokość 0,5 m. Natomiast wysokość całej kolumny, jest równa 0,67 m.

Słowa kluczowe: kolumna rektyfikacyjna, etanol, bilans ciepła kolumny

1. Nomenclature

R – reflux
x_i – mole fraction in the liquid phase
y_i – mole fraction in the gas phase
n_t – the number of theoretical shelfs
α_{AB} – the average coefficient of relative volatility
ρ – density
M_i – molar mass
T – temperature
a – specific surface of fill
$\dot{V}_{\text{mol}}$ – molar flow of steam
ϕ – degree of usability of the surface
S_k – cross-sectional area of the column
r_r – molar heat of evaporation
i_i – enthalpy
D, S, W – mol quantity of distillate, feed and decoction
t_{wS} – boiling temperature of feed
t_o – surrounding temperature
C_{wi} – specific heat of feed
V_{mol} – amount of steam moles
$\dot{Q}_{sk}$ – heat stream from condensed vapor
k – heat transfer coefficient
d_z – external diameter
d_w – internal diameter

2. Introduction

After World War II, in Poland, there were about 1200 agricultural distilleries eligible for commissioning and 15 industrial distilleries adapted for the processing of molasses. The agricultural distilleries were grouped into 'State Agricultural Enterprises' and industrial distilleries into the 'POLMOS' State Enterprise of the Spirits Industry. These companies also included alcohol rectification plants, vodka factories and yeast plants. The production of spirits in post-war Poland was ran by three distilleries, which converted cellulose alkalis to alcohol. These operated within the paper industry, and several branches producing fruit spirits.

In 1991, the Ministry of Finance approved the introduction of a certain amount of spirits into gasoline. This allowed producing a fuel containing 5% anhydrous alcohol. It was practically insignificant for the alcohol market. Bioethanol is a dehydrated alcohol, which improves car fuels by increasing their octane number, reducing the concentration of carbon monoxide by 20–30% and the concentration of hydrocarbons in the exhaust gases by 10%. In 1997, more

than 50% of domestic ethanol was consumed in addition to fuel, which accounted for about 100 million litres of spirits. In subsequent years, the consumption was already smaller [1].

In 2007, the production of agricultural distilleries in the EU accounted for about 10% of ethanol production. Apart from Poland, agricultural distillate was also obtained on a larger scale in Austria and Germany. In other countries, the distilleries operated rather locally. Agricultural distilleries in Poland operate depending on the demand reported from the fuel and spirits sector. Table 1 shows the production of agricultural distillate (raw spirit), the number of active agricultural distilleries and the production of bioethanol in Poland in 1994–2006. In 2009, the consumption of bioethanol as an additive for gasoline was at 293.6 thousand m^3, of which 56% came from domestic production; the remainder came from imports. The total production of ethanol in Poland in this year amounted to 205.8 thousand m^3. In 2009, there were 179 distilleries craft, of which 14 produced bioethanol as a bio component [2, 3].

Table 1. Production of agricultural distillate (raw spirit), number of active agricultural distilleries and bioethanol production in Poland in 1994–2006 [3]

Year	Agricultural distillate, mln l	Active agricultural distilleries	Bioethanol, mln l
1994	210	b.d.	27.0
1995	245	950	63.0
1996	278	900	100.9
1997	240.6	840	110.6
1998	208	700	99.8
1999	167.2	620	88.5
2000	173.3	380	51.5
2001	181	350	69.4
2002	210	330	82.8
2003	210–219.6	300	76.2
2004	195	242	45.2–48.5
2005	ok. 230–241.5	220	110.8
2006	259	217	161

In 2016, 98 million litres of vodka were produced. This is less than in 2012, when more than 102 million litres were produced, or 2011 when the production was over 103 million litres. The reason for the reduced production was a 15% increase in excise duty on strong alcohol, which was introduced on January 1, 2014 (Fig. 1). In Poland, excise duty represents 75% of the price of a half a litre of vodka. In the year 2016, the spirits sector transferred 11 billion zlotys to the state budget [4].

A different situation concerns the export of Polish vodka, which has been developing dynamically since the accession of Poland to the EU. Over the last 10 years, its sales have increased three fold. In 2003, Polish producers sold 15 million litres of vodka abroad. In 2013, the export of spirit drinks increased from 44.5 million litres to 50.75 million litres. In 2017, exports amounted to 51 million litres. Polish spirits are best sold in France and in the USA (Fig. 2) [5].

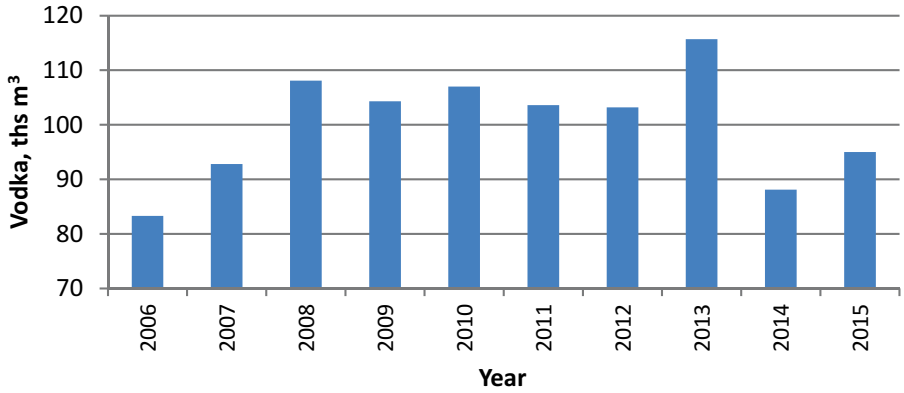


Fig. 1. Annual production of vodka in years 2006–2015 [4]

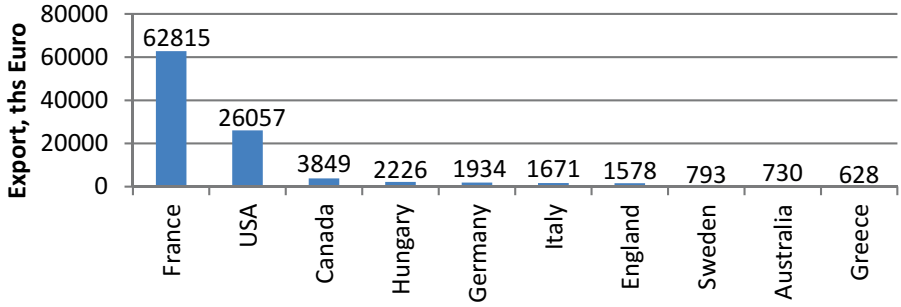


Fig. 2. The largest recipients of Polish vodka [5]

3. Project of periodic rectification column

In periodic rectification, the boiler is filled with a liquid mixture, which is heated to the boil point. The steam passing through the column is enriched in a more volatile component and the liquid in the less volatile component with a higher boiling point. At the top of the column, where the temperature is lower, the vapor is condensed and is collected as a distillate. As a result, we obtain mixtures of fractions that differ in condensing temperatures. In the boiler, a residue rich in less volatile component remains. The periodic rectification is used to separate small amounts of mixtures.

3.1. Project assumptions

The calculations assumed the process of periodic rectification carried out at a constant composition of distillate, for which reflux varies from R_p to R_k. The feed was 50 dm^3 in volume and contained a 25% vol. ethanol in water solution. The concentration of ethanol in the distillate was set at 0.7 mol·mol^{-1} (x_D), and in the residue – 0.01 mol·mol^{-1} (x_W).

3.2. Determining the reflux

To calculate the minimum reflux for the beginning and end of the process, the following formula (1) was used:

$$R_{\min} = \frac{x_D - y_i^*}{y_i^* - x_i} \tag{1}$$

where index 'i' is the composition of the feed (S) at the beginning of the rectification or composition of the residue (W) at its end. The star was the determined equilibrium concentrations corresponding to real concentrations.

The real reflux was calculated from equation (2). The coefficient of excess reflux (β) for the working calculations is taken in the range of 1.2–2.5 [6]. In order to more accurately select the reflux, an optimisation calculation would have to be made. For the purposes of this project, the value of β = 2 was assumed.

$$R = \beta \cdot R_{\min} \tag{2}$$

In order to receive a constant composition of distillate $x_D = 0.7$ mol·mol^{-1} (it was assumed), the reflux during rectification should be changed from 1.7 to 13.4 (Tab. 2). The total amount of distillate received (D) is 0.27 kmol (Fig. 3).

Table 2. The reflux for the beginning and the end of rectification

R_{minp}	R_p	R_{mink}	R_k
0.8	1.7	6.7	13.4

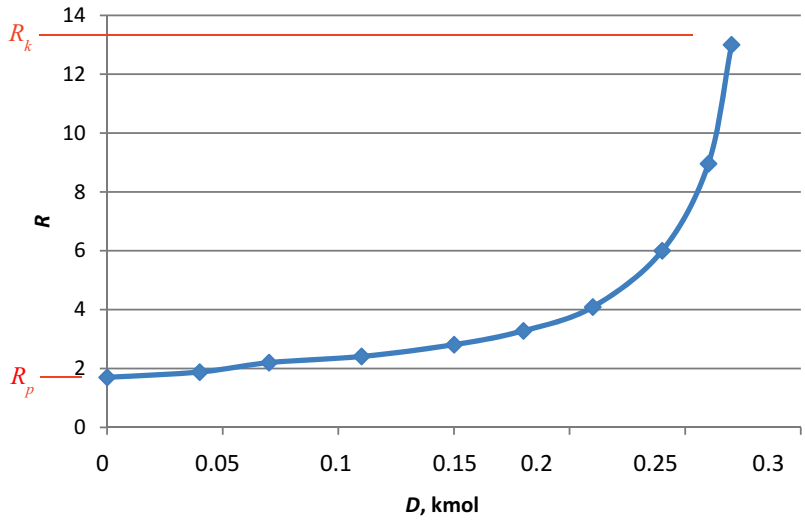


Fig. 3. Change of the amount of distillate depending on the change of the reflux

3.3. Determining the number of theoretical shelfs

3.3.1. McCabe-Thiele Method

McCabe-Thiele's graph method for determining theoretical shelfs is based on the introduction of equilibrium triangles between the operating line at the beginning of the rectification and the liquid-vapor equilibrium line. Figure 2.2 indicates the existence of four equilibrium steps. The first one is a boiler; therefore, the number of theoretical shelfs (n_t) is three (Fig. 4).

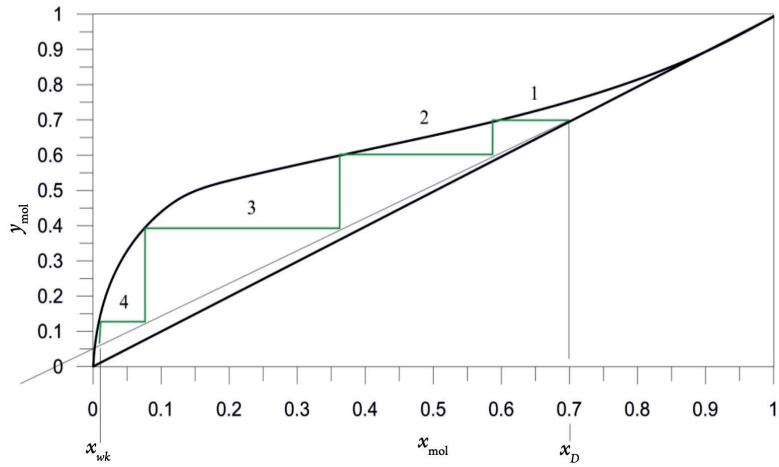


Fig. 4. Determination of theoretical shelfs by McCabe-Thiele methodology

3.3.2. Analytical method of Fenske

Fensky's analytical method uses average relative volatility at the boiling point of ethanol and water. The number of theoretical shelfs is calculated from the equation (3).

$$n_t = \frac{\log\left[\frac{x_D \cdot (1-x_{Wp})}{x_{Wp} \cdot (1-x_D)}\right]}{\log(\alpha_{AB})} - 1 \tag{3}$$

The calculated relative volatility factors at the boiling point of ethanol and water are 2.31 and 2.25, respectively. The number of theoretical shelfs determined by this method equals three.

3.4. Selection of column fill

As a fill, Raschig ceramic rings of Ace ChemPack were selected. Table 3 shows the characteristics of the rings used [7].

Table 3. Specification of Raschig ceramic rings

Specification of Raschig ceramic rings	
ring diameter, mm	15
ring height, mm	15
wall thickness of the ring, mm	2
specific surface (a_r), $m^2 \cdot m^{-3}$	330
free volume, $m^3 \cdot m^{-3}$	0.7
the number of items in 1 m^3	250 000
bulk density, $kg \cdot m^{-3}$	690

3.5. HETP and the height of the fill

In order to calculate the height equivalent of the theoretical plate ($HETP$), the equations of Granville (4) and Delzenne (5) were used. The fill height was determined for the average value of $HETP$ [7].

$$HETP = \left(8.8 \cdot 10^{(-4)} \cdot \frac{a_6}{a_r} \right) \cdot \frac{M_{S\rho S}}{T} \tag{4}$$

$$HETP = HETP_o \cdot \left[1 + 0.7 \cdot \log\left(\frac{D_k}{D_o} \right) \right] \tag{5}$$

In the formulas, a surface area of the 6 mm (a_6) diameter rings equal 790 $m^2 \cdot m^{-3}$, $HETP_o$ = 0.762 m, and the diameter of D_o = 0.915 m for the selected reference column, were assumed. The diameter of the column is 0.7 m (D_k). The average $HETP$ was 0.16 and the fill height was 0.5 m (h_w).

3.6. Mass transfer coefficient

The mass transfer coefficient (6) was calculated for the gas phase as the product of the mass transfer unit (h_{oG}) (7) and the number of mass transfer units (N_{oG}) (8) [9].

$$K_{oG} = h_{oG} \cdot V_{oG} \tag{6}$$

$$h_{oG} = \frac{\dot{V}_{\text{mol}}}{h_w \cdot a_r \cdot \varphi \cdot S_k} \tag{7}$$

$$N_{oG} = \int_{y_p}^{y_k} \frac{1}{y^* - y} \cdot dy \tag{8}$$

The mass transfer coefficient in the gas phase assumes the value:

$$K_{oG} = 2 \cdot 10^{(-3)} \left[\frac{\text{kmol}_A}{m^2 \cdot s \cdot \frac{\text{kmol}_A}{\text{kmol}}} \right]$$

3.7. Column height

The total height of the column (9) was calculated as the sum of the heights: above the mist eliminator – 0.07 m (h_1), the mist eliminator – 0.02 m (h_2), the condenser – 0.05 m (h_3), space under the grate – 0.01 m (h_4) the grate – 0.02 m (h_5) and the fill column – 0.5 m (h_w).

$$H_k = h_w + h_1 + h_2 + h_3 + h_4 + h_5 \tag{9}$$

$$H_k = 0.67 \text{ m}$$

3.8. Heat balance of rectification column

In the aim, which was determined to calculate the heat loss from the rectification process to the environment, the heat balance of the rectification column was prepared (10) [10, 11].

$$Q_w + Q_p = Q_d + Q_{sk} + Q_{\text{wyw}} + Q_{\text{str}} \tag{10}$$

The heat needed to produce the right amount of steam (11).

$$Q_w = r_D \cdot \int_0^D R dD + r_D \cdot D \tag{11}$$

Heat use to warm the feed in the boiler (12).

$$Q_p = C_{Ws} \cdot (t_{wS} - t_o) \cdot S \cdot M_S \tag{12}$$

Condensation heat of vapor (13).

$$Q_{sk} = r_D \cdot V_{\text{mol}} \tag{13}$$

Heat of residue, which remained in the boiler after the process end (14).

$$Q_{\text{wyw}} = i_w \cdot W \tag{14}$$

Heat received with distillate (15).

$$Q_d = i_D \cdot D \tag{15}$$

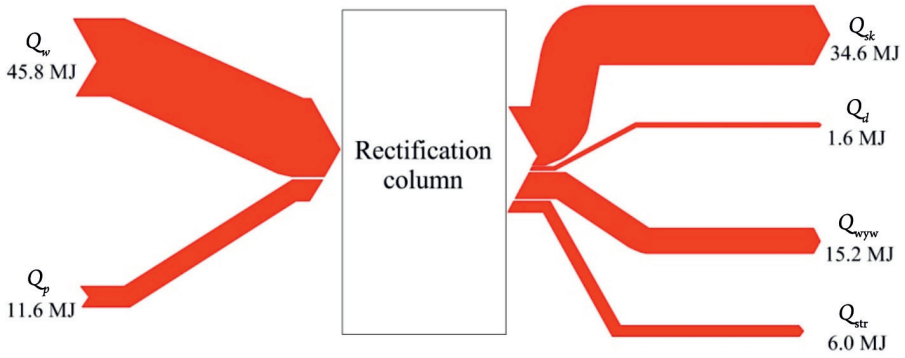


Fig. 5. Chart Sakey's for rectification column

The amount of heat lost to the environment during the rectification process based on the energy balance of the column is equal to 6 MJ (Q_{str}) and represents about 10% of all outgoing streams. Most of the energy was required to condense the resulting vapours (Fig. 5).

3.9. Selection of the pot and condenser for the column

3.9.1. Condenser

In order to select the condenser, it was necessary to calculate the heat exchange surface (16).

$$A = \frac{\dot{Q}_{sk}}{k \cdot \Delta T_m} \tag{16}$$

where ΔT_m is the logarithmic difference of temperature in the condenser.

Logarithmic diameter of the coil (17).

$$d_n = \frac{d_z - d_w}{\ln \frac{d_z}{d_w}} \tag{17}$$

Coil length (18).

$$L = \frac{A}{\pi\left(\frac{d_n}{d}\right)} \tag{18}$$

In addition, from the following formula (19), the amount of water flowing through the condenser was calculated at 57.2 kg·h^{-1}.

$$\dot{m}_{\text{wody}} = \frac{\dot{Q}_{sk}}{c_w \cdot \Delta T_m} \tag{19}$$

From the above formulas, a condenser cooling area equal 0.11 m^2 was calculated along with a coil diameter of 9 mm and a coil length of 3 m. For these values, a condenser was selected in the form of a tube with an internal coil (Tab. 4).

Table 3. Condenser characteristics [12]

Condenser	
tube diameter of the condenser, mm	70
length of pipe, mm	400
coil length, mm	3000
coil diameter, mm	10

3.9.2. Pot

The pot of the column was selected from available materials with the assumed volume of 60 dm^3. The tank is made of stainless steel, acid resistant AlSi304. Its specification is shown in Table 5.

Table 4. Characteristics of the pot [13]

Pot	
volume, dm^3	60
height, mm	340
diameter, mm	480
wall thickness, mm	1.4
net weight, kg	5

4. Conclusions

- Ethyl alcohol can be obtained in the laboratory using a rectification column for period work, which is suitable for a small amount of feed.
- At work, 25% vol. ethanol in feed was assumed. At such concentrations of ethanol, there is resistance to gamma-rayed Saccharomyces cerevisiae strains only.
- Periodic rectification can be carried out in two different ways by establishing a constant concentration of ethanol in the distillate by changing the reflux or vice versa.
- The calculations show that a change in reflux between 1.7 and 13 would allow the ethanol concentration in the distillate to be maintained at about 70 mol%. This method can be technically difficult to achieve and requires constant process parameters control, but it is possible to obtain a high concentration product already in the first rectification cycle. The concentration of ethanol in the residue remains small.
- The project used ceramic Raschig rings with a diameter of 15 mm, which occupy a column height of 0.5 m. The height of the entire column is 0.67 m.
- Heat losses in the column were obtained at an acceptable level of 10%.

References

[1] Łączyński B., *Skrócony kurs gorzelnictwa rolniczego*, Wydawnictwo Sigma-NOT, Warszawa 1993.
[2] Jarociński J., Jarosz K., *Gorzelnictwo i drożdżownictwo*, Wydawnictwo Szkolne i Pedagogiczne, Warszawa 1994.
[3] Kupczyk A., *Perspektywy rozwoju polskich gorzelni rolniczych*, Rynki Alkoholowe v 7/2007, 60–65.
[4] Biernacka P., *Metody kompleksowej analizy składu produktów ubocznych procesu fermentacji alkoholowej w półproduktach i produktach spirytusowych*, Rozprawa doktorska, Politechnika Gdańska, Gdańsk 2012.
[5] Ministerstwo Skarbu Państwa – Polski rynek wódki, http://www.msp.gov.pl/pl /przekształcenia/serwis-gospodarczy/wiadomoscigospodarcze/28430,Polski-rynek-wodki.html (access: 10.06.2017).
[6] Pawłow K.F., Romankow P.G., Noskow A.A., *Przykłady i zadania z zakresu aparatury i inżynierii chemicznej*, WNT, Warszawa 1971.
[7] AceChemPack – ceramiczne pierścienie Raschiga, http://www.tower-packing.com/ceramic /ceramic_rasching_rings_random_packings.html (access: 10.06.2017).
[8] Ziółkowski Z., *Destylacja i rektyfikacja w przemyśle chemicznym*, Wydawnictwo Naukowo-Techniczne, Warszawa 1978.
[9] Hobler T., *Dyfuzyjny ruch masy i absorbery*, Wydawnictwo Naukowo-Techniczne, Warszawa 1976.
[10] Ciborowski J., *Inżynieria procesowa*, Wydawnictwo Naukowo-Techniczne, Warszawa 1973.
[11] Raznjevic K., *Tablice cieplne z wykresami*, Wydawnictwo Naukowo-Techniczne, Warszawa 1983.
[12] Mojamaszyna – skraplacz, http://mojamaszyna.pl/chlodnica-destylator-3-metry-rurki-wezownicy-fi-10-i6600199556.html (access: 10.06.2017).
[13] Destylandia – zbiornik 60 l, http://destylandia.pl/index.php?id_product=2068
[14] &controller=product (access: 10.06.2017).

TECHNICAL TRANSACTIONS 3/2018
CZASOPISMO TECHNICZNE 3/2018
CHEMISTRY
DOI: 10.4467/2353737XCT.18.038.8300
SUBMISSION OF THE FINAL VERSION: 27/2/2018

Sylwia Piątkowska
Monika Gwadera (mgwadera@chemia.pk.edu.pl)
Chair of Chemical and Process Engineering, Faculty of Chemical Engineering and Technology, Cracow University of Technology

The effect of external mass transfer resistance on the adsorption rate in a liquid phase

Wpływ zewnętrznych oporów przenoszenia masy na szybkość adsorpcji w fazie ciekłej

Abstract

The paper presents the results of studies on the adsorption of a dye from a water solution on activated carbon. An activated carbon with the best adsorption properties for the dye used in the studies was selected experimentally. The equilibrium and kinetic measurements for the selected system were conducted. Constants of linear, Langmuir and Freundlich isotherms were determined. The obtained results have shown that increasing the flow rate of the solution in an adsorber results in increasing the adsorption rate.

Keywords: adsorption, activated carbon, dye removal

Streszczenie

W pracy zaprezentowano wyniki badań adsorpcji barwnika z roztworu wodnego na węglu aktywnym. Dokonano eksperymentalnego doboru węgla aktywnego o najlepszych właściwościach adsorpcyjnych dla stosowanego w badaniach barwnika. Wykonano badania równowagi i kinetyki adsorpcji dla badanego układu. Wyznaczono stałe w równaniu izotermy liniowej, Langmuira i Freundlicha. Otrzymane wyniki wskazują, że zwiększenie natężenia przepływu roztworu w adsorberze skutkuje zwiększeniem szybkości adsorpcji.

Słowa kluczowe: adsorpcja, węgiel aktywny, usuwanie barwnika

1. Nomenclature

A – absorbance
b – constant
C – concentration of the dye in the liquid phase
C_0 – initial concentration of the dye in the solution
C_e – equilibrium concentration of the dye in the solution
D_e – effective diffusion coefficient
k_F – constant
k_l – mass transfer coefficient on the fluid side
K – equilibrium constant
L – characteristic linear dimension
m_c – mass of activated carbon
n – rotational speed of the peristaltic pump
q – concentration of the adsorbed component in adsorbent grains
q_e – equilibrium content of the dye in adsorbent grains
q_∞ – constant
Q_v – flow rate of the liquid
t – time
V_s – volume of the dye solution
v – constant

2. Introduction

Adsorption is one of the basic surface phenomena; it is the occurrence of changes in the concentration of substances on the surface of adjacent phases. Due to the nature of the acting forces, there are two main types of it: chemical and physical adsorption. In the case of chemical adsorption, there are forces, which determine the formation of a chemical bond on the surface of a solid [1]. The chemical bond formed because of a chemical reaction involved in chemisorption makes this type of adsorption stronger than in the case of physical adsorption [2]. The heat of this process is of the same order as the heat of a chemical reaction [1]. Physical adsorption is caused by the forces of intermolecular interactions. Between the molecules of the substance, there are Van der Waals forces, which equilibrate inside of the phase. Moreover, particles located on the surface of contacting phases are subjected to unbalanced adhesion forces directed perpendicularly to the boundary surface [1].

Adsorption is a process used in various fields of chemical applications, technology and industry. It is applied for the purification of air and water, in recovering industrial solvents and in drying processes. It is also possible to measure the size of particles and pores in powdered substances using the adsorption process [3]. Adsorption on activated carbon [4–7] is primarily used to dispose of organic compounds, both natural and man-made, from water. Activated carbon as an adsorbent is also productive in terms of removal of certain types of

viruses, some inorganic impurities, and the binding of chlorine and chloramines. Adsorption allows one to remove impurities causing color, odor and taste from water.

When molecules of a gas or liquid hit the surface of a solid, some of them stick to it and are adsorbed, while other bounce off. At first, the velocity of adsorption is high, because the whole surface of the adsorbent is free from any molecules. The velocity becomes lower with time due to the increasing coverage of the surface of solid with the molecules of adsorbate. In the meantime, the velocity of desorption (a process, where previously adsorbed molecules are peeled from the adsorbent) becomes faster. In some cases, the velocity of adsorption will become so low, and the velocity of desorption so fast, that an equilibrium will be reached, and both velocities will be equal. Generally speaking, it is a dynamic equilibrium, because the number of adsorbed and desorbed molecules is the same [2]. The adsorption equilibrium state for the adsorbate-adsorbent system is dependent on the pressure of the gas and the temperature of the whole process. It can be presented as an adsorption isotherm at a constant temperature, an adsorption isobar at a constant pressure or an adsorption isostere. The widely used method, which represents the equilibrium state of the adsorption system, is the adsorption isotherm because of its simplicity and versatility.

The mass transfer process under consideration may be divided into five stages, which are theoretically important from the kinetics point of view [8]:

1. Transport of the molecules of the adsorbate in the solution to the boundary layer solution-adsorbent.
2. Diffusion in the boundary layer near the surface of the adsorbent.
3. Diffusion in the pores of the adsorbent to its active sites.
4. Surface diffusion.
5. Proper adsorption, during which the adsorbate molecules are placed at active adsorbent sites.

The rate of each stage is dependent on the velocity of the individual components. In this case, the slowest step limiting the rate of adsorption is the process of multistage diffusion, which is the transport of adsorbate molecules to adsorbent active sites.

In this paper, particular attention was paid to the process of adsorption of the dye on activated carbon [9]. The purpose of this work was to determine the equilibrium and kinetics of the adsorption of the dye Direct Fast Navy Blue BR 200% from an aqueous solution on the active carbon ORGANOSORB 10. The study of adsorption equilibrium was aimed at determining the equilibrium curve, i.e. adsorption isotherm. The dependence of the dye content in the solid phase on its concentration in the liquid phase was obtained by measurements. Constants of the linear, Freundlich and Langmuir models were found. In the case of adsorption kinetics, the measurements focused on the temporal changes of dye concentration in the aqueous solution for different flow rates of the solution in the adsorber filled with activated carbon grains. The purpose of these studies was to determine the effect of the flow rate on the changes of dye concentration in the liquid and solid phase over time.

3. Materials and Methods

3.1. Adsorbate – Direct Fast Navy Blue BR 200%

The adsorbate used for the study was the Direct Fast Navy Blue BR 200% (Granat bezpośredni BR 200%) produced by the Boruta company. It is applied primarily in the textile industry for dyeing mainly cellulose fibers, such as cotton, wool, viscose or sisal and silk [10]. Its properties are given in Table 2.1 [10] and structure is shown in Fig. 1 [11].

Table 1. General characteristic of Direct Fast Navy Blue BR 200%

General look	Dark gray powder
pH	9.5–10.5
Solubility at 20°C	30 g/dm^3
Solubility at 60°C	40 g/dm^3
Bulk density	650 kg/m^3

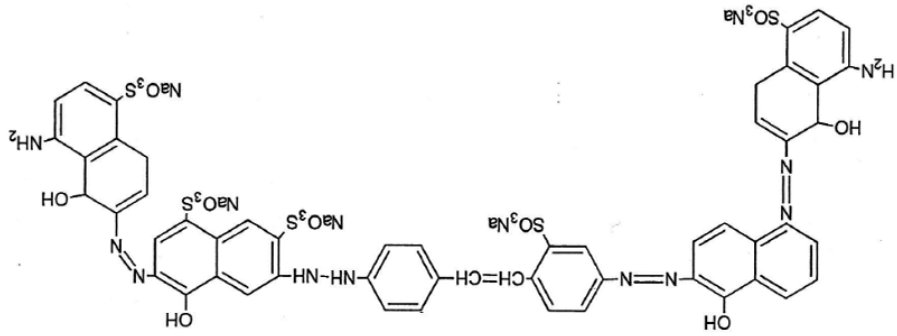


Fig. 1. Chemical structure of Direct Navy Blue BR 200%

3.2. Adsorbent – choosing the best

Four solutions of Direct Fast Navy Blue BR 200% with distilled water with a concentration of 5 mg/dm^3 were made in flat-bottom flasks. The volume of the solutions was equal to 100 cm^3. Different types of active carbon were added to each flask in the amount of 0.5 g:

- BA 10 – produced from hard coal using the vapor-gas method. The diameter of the granulate is about 3 mm. It is characterized by a large surface area, very good mechanical strength and high sorption capacity [12].
- BA 20 – formed activated carbon made from hard coal by the vapor-gas method. The diameter of the granules is 3-4 mm. It has a basic character. It is characterized by high adsorption capacity and very good mechanical strength. It is widely used in water and wastewater treatment [13].
- ORGANOSORB 10-CO – made from coconut shells. Its grain size is 2.36–0.6 mm. It is used to remove impurities of organic origin and in liquid filtration [14].

- ORGANOSORB 10 – made from coconut shells [15] and specially prepared for the purification of water (drinking, process) and sewage. High adsorption capacity eliminates organic pollutants to trace amounts [16].

After a few days, the concentration in each solution was measured. It was observed that the concentration of the solution with ORGANOSORB 10 was the lowest. This meant that this type of active carbon adsorbent had the best adsorption properties for the considered dye and that is why it was chosen for further experiments. Its physical and chemical properties are given in table 2.2 [14].

Table 2. Physical and chemical properties of ORGANOSORB – 10

Basic material	**Bituminous coal – granulated**
Specific surface area	Min. 950 m^2/g
Iodine number	Min. 950 mg/g
Moisture	Max. 5 %
Grain	0.6–2.36/0.425–1.70 mm

3.3. Adsorption Equilibrium Measurements

The adsorption equilibrium was examined by the static method. For this purpose, 8 flat-bottom flasks were prepared, to which 100 cm^3 of a solution of a known concentration of the dye Direct Fast Navy Blue BR 200% and 0.3 g of the pure adsorbent ORGANOSORB 10 were introduced. The samples were maintained at room temperature (20°C) and were regularly shaken mechanically (Fig. 2). The frequency of the shaker was set to 300 rpm. Before and after each shaking, the absorbance of solutions was measured in a spectrophotometer Shimadzu UV-2600 in a measuring cell with an optical path length of 50 mm. After about 10 days, the concentration of the solutions in flasks stopped changing, which can be assumed as being equivalent to an equilibrium between the adsorbate and the adsorbent.

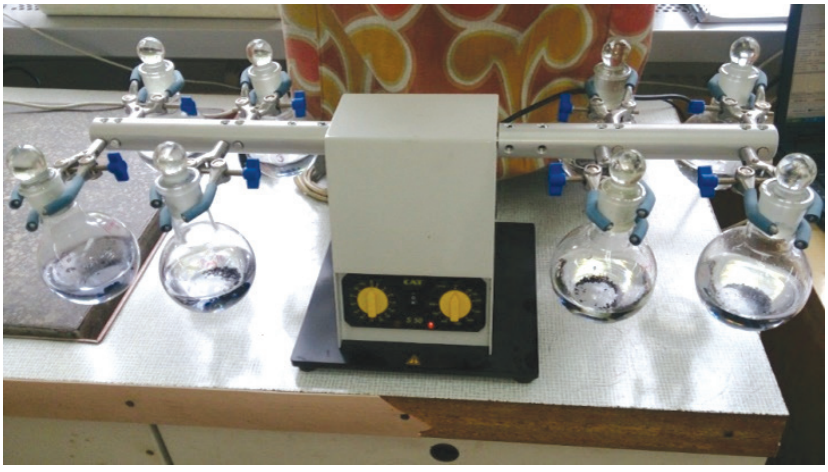


Fig. 2. Flat bottom flasks placed in a shaker during equilibrium measurements

3.4. Adsorption Kinetics Measurements

The test stand (Fig. 3) at which the dye adsorption kinetics studies were performed consisted of a 600 cm^3 beaker, a magnetic stirrer and a peristaltic pump together with the adsorber and connecting pipes forming a closed circuit. The peristaltic pump was set to the appropriate value. The examined solution was continuously mixed in a beaker with a magnetic stirrer to make the concentration uniform in the entire volume of the tank. Subsequently, the solution was pumped by the peristaltic pump to the adsorber. After passing through the adsorber filled with activated carbon, the solution returned to the beaker. Ten series were made, with changing the flow rate, as presented in Tab. 3. In each measurement, the amount of active carbon in the adsorber was about 2 g. Each measurement took an hour including a sample zero, that is the solution before any contact with active carbon. Every 5 minutes, 15 cm^3 of the solution was pipetted from the beaker, placed in a measuring cell of the spectrophotometer in order to measure the absorbance of the solution. All measurements were made for a wavelength equal to 584 nm. After measuring the absorbance, the solution from the measuring cell was poured back into the stirrer tank. The results were reported using the Shimadzu UV Probe program and saved to the hard disk.

Fig. 3. Laboratory equipment for adsorption kinetics measurements

Table 3. Pump speed and the flow rate of the liquid through the adsorber

Measurement No.	Pump rotations [rpm]	Q_v [cm^3/min]
1	35.625	19.688
2	67.742	35.806
3	110.901	58.638
4	136.341	72.128
5	169.780	89.617
6	203.220	107.107
7	236.660	124.597
8	270.100	142.087
9	303.540	159.576
10	336.980	177.066

4. Results

4.1. Equilibrium Studies

The graph in Fig. 4 shows how the concentration of the dye changed over time. As can be seen, the concentration lowers with time for all the considered water solutions of the dye. After some time, the lines stabilize and it can be assumed that the equilibrium was reached. For the lowest concentrations of the dye in water, after a few days, the concentration was so low that it was not possible to measure the absorbance accurately. Possibly, if the spectrophotometer was equipped with a measuring cell with an optical path length of 100 mm, such small concentrations could be measured. The graph also shows that mechanical shaking greatly influences the rate of adsorption. Only few hours of shaking can result in a similar decrease in concentration as can be observed after 24 hours for solutions without shaking.

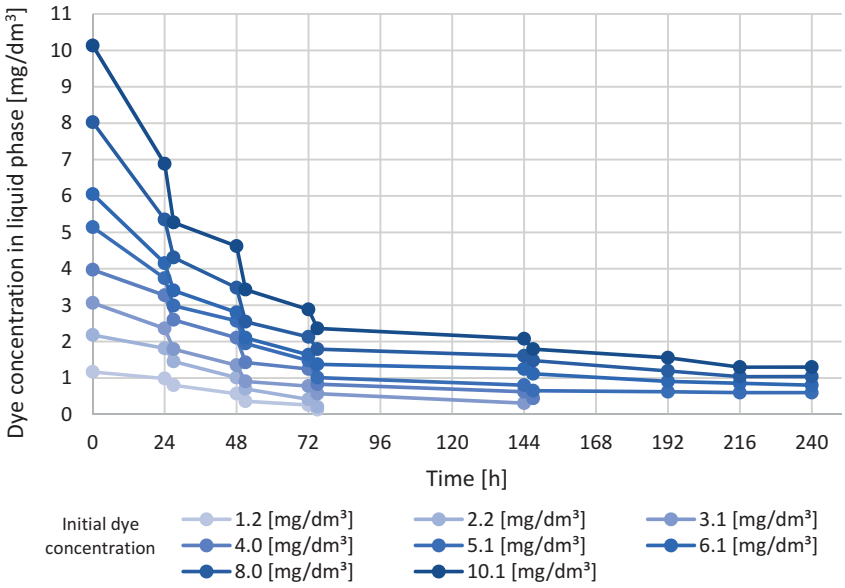


Fig. 4. Temporal concentration changes for all equilibrium measurements

The graph in Fig. 5 shows the relationship between dye concentration in adsorbent grains q_e and in the solution C_e. The points mark experimental results. The dye concentration values in the adsorbent grains q_e were calculated from the mass balance:

$$q_e = \frac{v_s(c_0 - c_e)}{m_c} \tag{1}$$

The lines in Fig. 5 refer to adsorption isotherms: linear, Freundlich and Langmuir. Parameters of these models were calculated by linear regression on the basis of the experimental results and are given in Table 4. As can be seen in Fig. 5, the experimental points correspond most to the Freundlich isotherm.

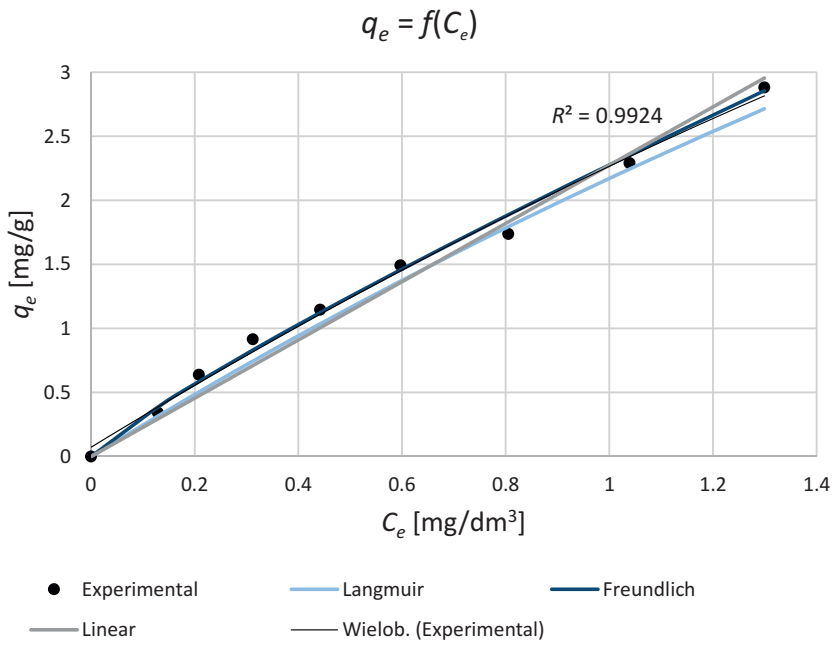


Fig. 5. Adsorption isotherm of Direct Fast Navy Blue BR 200%

Table 4. Constants of adsorption isotherms

Adsorption isotherm				
Linear $q_e = KC_e$	Langmuir $q_e = q_\infty \frac{bC_e}{1+bC_e}$		Freundlich $q_e = k_F C_e^{\nu}$	
K	q_∞	b	k_F	ν
2.275	16.556	0.151	2.278	0.865

4.2. Kinetics Studies

Table 5 as well as graphs 6 and 7 show the results of the adsorption kinetics studies. They refer to measurement 5 (according to Tab. 3), i.e. to fluid velocity equal to 89.617 cm^3/min. The dye concentration in the liquid decreases during the adsorption (Fig. 6), while concentration in the solid phase increases (Fig. 7). Both concentrations seek to achieve equilibrium values, respectively C_e and q_e.

The graphs in Figs. 8 and 9 show a comparison for all the performed kinetics measurements. It can be concluded that the flow rate influences the adsorption rate. When the flow rate of liquid phase, which bathes the adsorbent, is increased, the liquid phase concentration decreases for a given time (Fig. 8). When the liquid phase flow rate is increased, the solid phase concentration also increases for a given time (Fig. 8). As we increase the turbulence of the liquid around the grains, the mass transfer coefficient k_l increases and the external resistance of the mass transport decreases. Adsorption occurs more intensively and effectively.

Table 5. Results of experiment No. 5: values for the flow rate equal to 89.617 cm^3/min

TIME [min]	ABSORBANCE [–]	C [mg/dm^3]	q [mg/g]
0	0.183	4.753	-
5	0.163	4.233	0.061
10	0.153	3.974	0.092
15	0.138	3.584	0.139
20	0.130	3.376	0.163
25	0.118	3.064	0.201
30	0.111	2.883	0.222
35	0.104	2.701	0.244
40	0.099	2.571	0.259
45	0.092	2.389	0.281
50	0.087	2.259	0.296
55	0.081	2.104	0.315
60	0.077	2.000	0.327

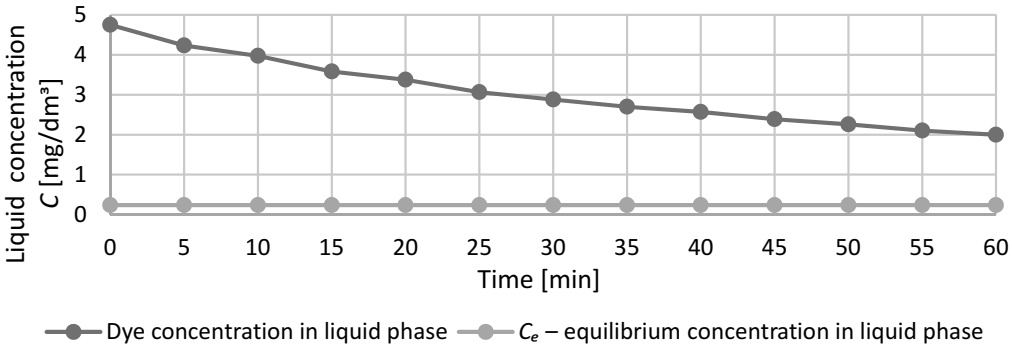


Fig. 6. Concentration of dye in the liquid phase vs. time for measurement 5

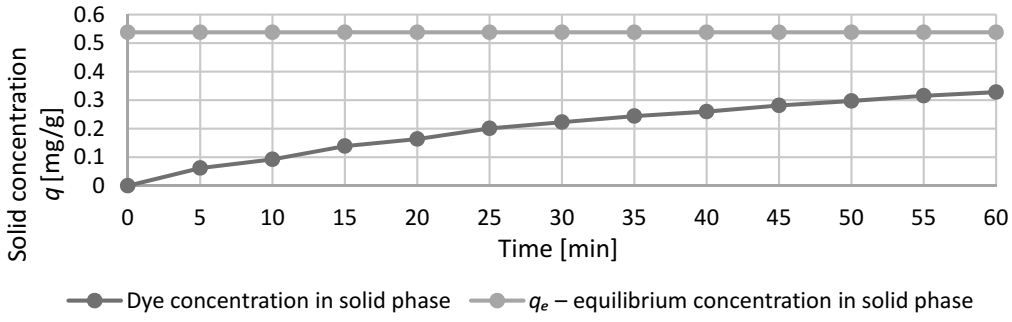


Fig. 7. Concentration of dye in solid phase vs. time for measurement 5

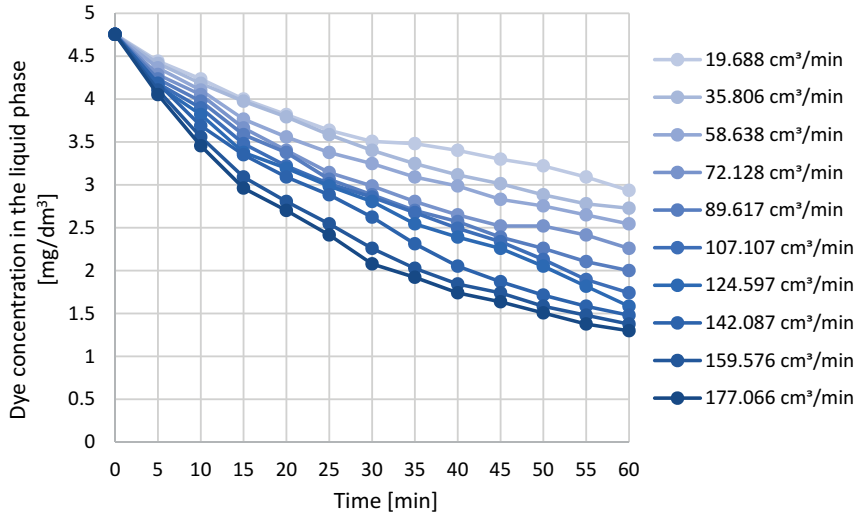


Fig. 8. Temporal changes of concentration in the solid phase for all the kinetics measurements

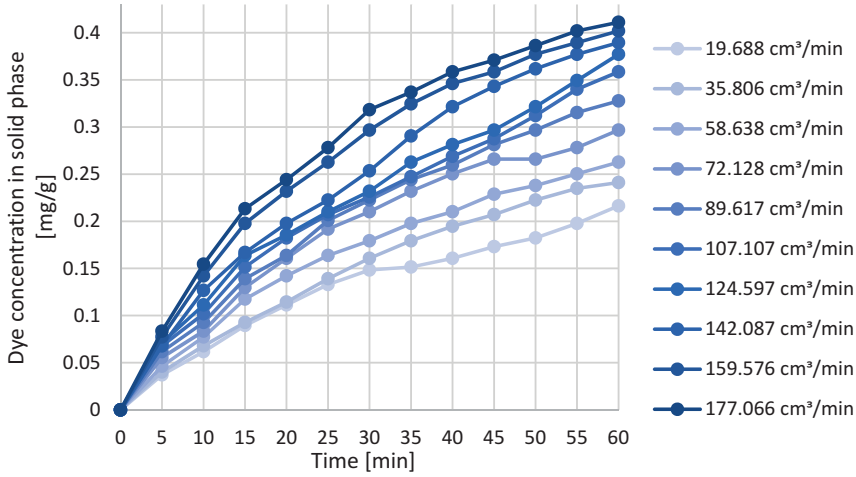


Fig. 9. Temporal changes of concentration in the liquid phase for all the kinetics measurements

5. Summary

- The results of the studies presented in this work show that the activated carbon ORGANOSORB 10 effectively removes the dye Direct Navy Blue BR 200% from its aqueous solution.
- As a result of the adsorption equilibrium analysis, it was shown that experimental points correspond the most to the Freundlich isotherm.
- Achieving equilibrium in the adsorption systems under investigation is greatly accelerated by mechanical shaking, which increases the speed of the liquid phase

movement and the mass transfer coefficient, and intensifies the transfer of matter between the fluid and the adsorbent. Short shaking causes a much greater decrease in the dye concentration in the solution than longer-lasting adsorption without the use of external factors.

- With increasing the flow rate of the solution in the adsorber, the concentration of the dye in the liquid phase decreases and its concentration in the solid phase increases for a given time. This can be explained by the fact that a higher flow rate causes the solution to be in contact with the adsorbent particles faster. Large fluid turbulence allows for an easier exchange of mass. This results in a higher mass transfer coefficient and lower external mass transfer resistance. However, the maximum flow speed is limited by residence time adequate for efficient adsorption and pressure drop of the fluid.

References

[1] Paderewski M.L., *Procesy adsorpcyjne w inżynierii chemicznej*, Wydawnictwa Naukowo-Techniczne, Warszawa 1999.

[2] Bansal R.Ch., Goyal M., *Activated Carbon Adsorption*, Taylor &Francis Group, Boca Raton 2005.

[3] Ponec V., Knor Z., Cerny S., *Adsorption of solids*, Butterworth Group, London 1974.

[4] Bahamon D., Carro L., Guri S., Vega L.F., *Computational study of ibuprofen removal from water by adsorption in realistic activated carbons*, Journal of Colloid and Interface Science 498/2017, 323–334.

[5] Cermakova L., Kopecka I., Pivokonsky M., Pivokonska L., Janda V., *Removal of cyanobacterial amino acids in water treatment by activated carbon adsorption*, Separation and Purification Technology173/2017, 330–338.

[6] Yongil Kim, Jae-Kwang Kim, Christoph Vaalma, Geun Hyeong Bae, Guk-Tae Kim, Stefano Passerini Youngsik Kim, *Optimized hard carbon derived from starch for rechargeable seawater batteries*, Carbon 129/2018, 564–571.

[7] Sanja Panić, Ákos Kukovecz, Goran Boskovic, *Design of catalytic* carbon *nanotube-based reactor for* water *denitration – The impact of* active *metal confinement*, Applied Catalysis B: Environmental 225/2018, 207–217.

[8] Kowal A., Świderska-Bróż M., *Oczyszczanie wody*, Wydawnictwo Naukowe PWN, Warszawa 2005.

[9] Gwadera M., Kupiec K., *Modeling of adsorption in a mixing tank*, Technical Transactions, 1-Ch/2015, 3–16.

[10] Safety data sheet of dye Direct Fast Navy Blue BR 200% [online], http://www.boruta-zachem.pl/pl/kolor/produkty/barwniki/barwniki-bezposrednie-i-helionowe?rid=1170 (access: 27.05.2017).

[11] Przybojewska B., *An evaluation of the genotoxic properties of some chosen dyes using the micronucleus test in vivo*, Mutation Research/Genetic Toxicology 367/1996, 93–97.

[12] *Formowane węgle aktywne BA 10 i BA 11*, http://program.republika.pl/produkty/ba10iba11.htm (access: 27.05.2017).
[13] *Formowany węgiel aktywny BA 20*, http://program.republika.pl/produkty/ba20.htm (access: 27.05.2017).
[14] *Węgiel aktywny*, http://www.filterpol.pl/wegiel.html (access: 27.05.2017).
[15] *Filtracja na węglu aktywnym*, http://gc2000.pl/urzadzenia-filtracyjne/filtracja-na-weglu-aktywnym (access: 27.05.2017).
[16] *Safety data sheet of activated carbon ORGANOSORB 10*, http://www.filtrybb.pl/img/cms/pdf/Organosorb10_PL.pdf (access: 27.05.2017).

TECHNICAL TRANSACTIONS 3/2018
CZASOPISMO TECHNICZNE 3/2018
CIVIL ENGINEERING
DOI: 10.4467/2353737XCT.18.039.8301
SUBMISSION OF THE FINAL VERSION: 6/2/2018

Andrzej Flaga (aflaga@pk.edu.pl)
Renata Kłaput
Agnieszka Kocoń
Institute of Structural Mechanics, Faculty of Civil Engineering, Wind Engineering Laboratory, Cracow University of Technology

Wind tunnel tests of wind pressure distributions for four different tent halls

Badania modelowe rozkładu ciśnienia wiatru czterech różnych hal namiotowych

Abstract

This papers concerns measurements of wind pressure distributions on the roofs and side walls of tent hall models. Four tent halls of different shapes and constructions were investigated in the boundary layer wind tunnel at Cracow University of Technology, Poland. On the basis of these measurements, different schemes of wind pressure coefficient distributions for these structures were determined, including mean and extreme values of the coefficients. The obtained results, their analyses and comparisons are of great importance from a structural design point of view for such types of structures.

Keywords: wind tunnel test, tent halls, wind pressure coefficients

Streszczenie

Praca zawiera wyniki badań rozkładu ciśnienia wiatru na dachach i ścianach bocznych hal namiotowych. W tunelu aerodynamicznym Politechniki Krakowskiej przebadano cztery hale namiotowe różnych kształtów. Na podstawie tych pomiarów określono schematy rozkładów współczynników ciśnienia wiatru oraz średnie i ekstremalne wartości współczynników. Otrzymane wyniki oraz ich analizy i porównania mają duże znaczenie z punktu widzenia projektowania tego typu konstrukcji.

Słowa kluczowe: tunel aerodynamiczny, hale namiotowe, współczynniki ciśnienia wiatru

1. Introduction

The wind action on four tent halls of various roof shapes and constructions was determined from wind tunnel experiments conducted in the boundary layer wind tunnel at Cracow University of Technology. Investigations concerned the variables of wind angle, terrain roughness and roof shape.

Tent halls are made of aluminium profiles covered with a two-sided textile membrane enhanced with PCV layers. These kinds of structures are widely used in low-rise industrial buildings, sport facilities, warehouses. However, their relatively light weight and flexibility make them vulnerable to wind action; therefore, knowledge of wind pressure distribution on the side walls and roofs is essential in the structural design of these objects.

Many researchers have investigated the problem of wind action on objects of different shapes [3, 4, 8]. On the basis of these investigations, it was concluded that the geometry of a structure has an influence on pressure distribution on the building walls. The local wind pressures on walls of different low-rise buildings have been investigated by, inter alia, Gavanski & Uematsu [2], Alrawashdeh & Stathopoulos, 2015 [2]. In the case of light structures, peak local wind pressures are the most important in terms of object safety – this was considered in Pratt & Kopp [6] and Saathoff & Melbourne [7].

The main goal of this paper is to present the results of wind tunnel tests on four different types of tent halls. For each hall model, wind pressures were measured on the external surfaces of the roofs and side walls of the tent halls. On the basis of wind pressure at the measurement points and the reference pressure, the mean and extreme wind pressure coefficients were calculated. The obtained values of aerodynamic coefficients can be used in static-strength calculations of major structural elements.

2. Description of tent halls models

The models used in the wind tunnel tests were made at the following scales: 1:37.5 – model A; 1:62.5 – model B; 1:71.4 – model C; 1:16 – model D. The scales were chosen in such a way as to obtain models within a similar range of dimensions? The elements of the A-C models were made from extruded polystyrene XPS plates and the D model was made from laminated Plexiglas which had been cut by a laser. Each model consists of side walls, a roof and a base. Figure 1 presents the schematics of the tent halls models.

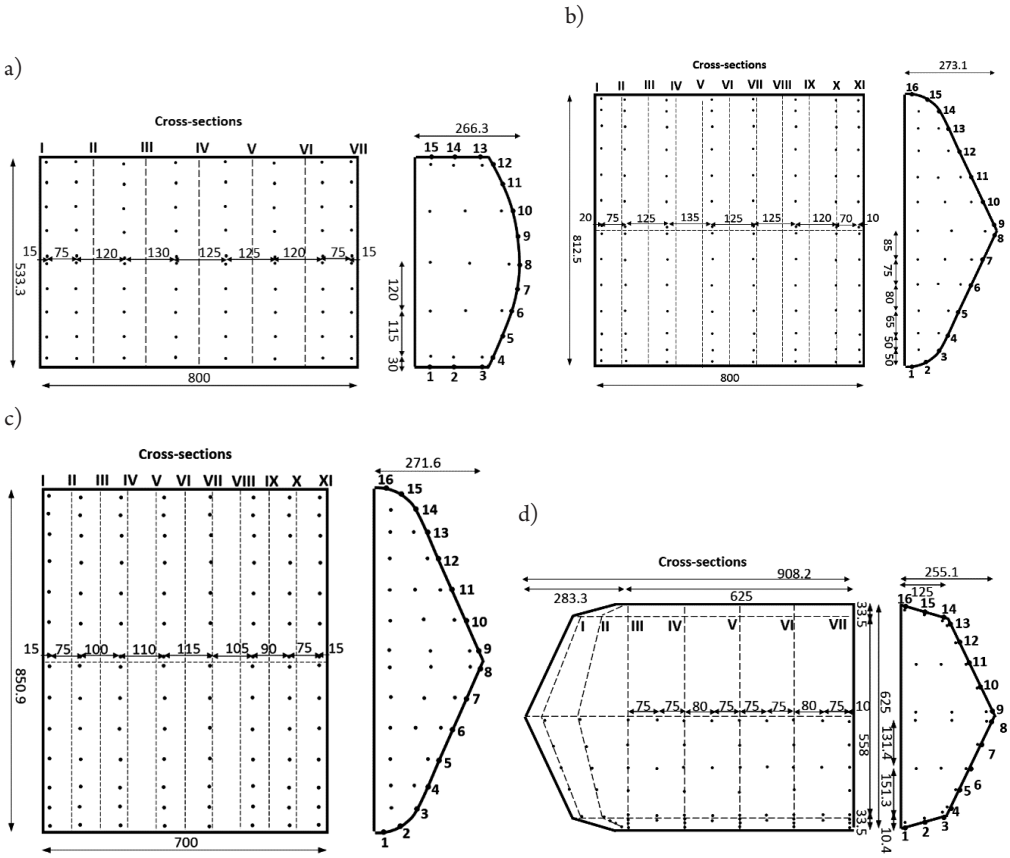


Fig. 1. Schematics of the tent halls models investigated in the wind tunnel tests: a) model A, b) model B, c) model C, d) model D

3. Description of the research

3.1. Simulation of boundary layer

The presented tests were conducted in the boundary layer wind tunnel a the Wind Engineering Laboratory at Cracow University of Technology. The basic dimensions of the working section of the wind tunnel are: 2.20 m (width), 1.40 m (height), 10.00 m (length). In the initial part of the investigations, the structure of the wind flow was determined. The wind profile was formed with the use of a barrier with a height of 20 cm. Thermo-anemometers were used to measure the mean and fluctuation of the wind velocity at 6 points located in the working section of the wind tunnel at heights from 5 cm to 30 cm above the floor level. Using power-low form of wind profile and data obtained from measurements, the following wind profile parameters were obtained:

$$V(z) = V_{ref}\left(\frac{z}{z_{ref}}\right)^{\alpha}$$

$$z_{ref} = 0.3\ \text{m},\ V_{ref} = 14.7\,\frac{\text{m}}{\text{s}},\ \alpha = 0.18.$$

where:

z_{ref} – reference height [m],
α – exponent depend on terrain roughness,
V_{ref} – reference wind velocity.

The obtained wind profile and turbulence intensity profile are shown in Fig. 2 a) and b). The red points mark values from wind tunnel tests and the black line marks function determined by least-square regression. The turbulence intensity I_v [%] on the reference level (z_{ref} = 0.3 m) was 19%. The power spectral density (Fig. 2c) was made from the velocity generated with time intervals of 0.002 s over 20 s (resulting in 10,000 samples) and compared with the Davenport spectrum.

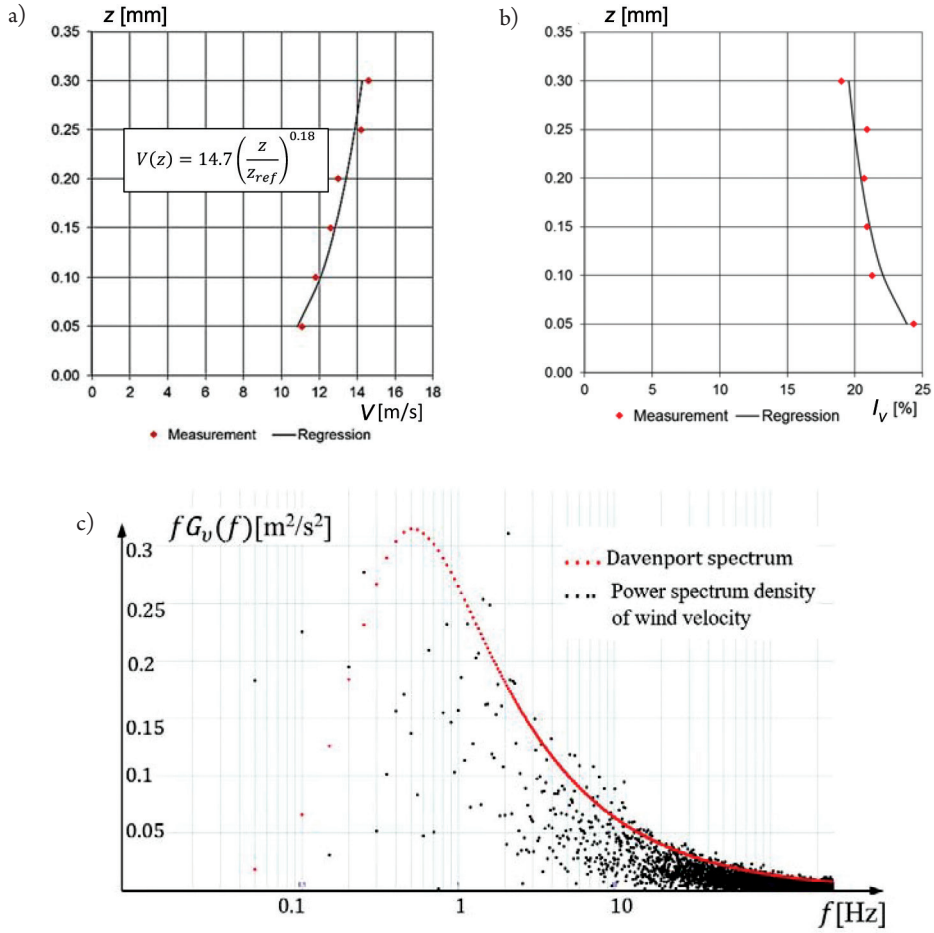


Fig. 2a) vertical profile of mean wind velocity in the wind tunnel; b) turbulence intensity profile; c) power spectrum density of wind velocity for measurement point at reference height and Davenport spectrum determined by least-square regression

3.2. Characteristics of the wind tunnel tests

Models of tent halls were placed in the working section of the wind tunnel on a round turnable table with a diameter of 2 m enabling the change of the angle of wind onflow onto the examined objects. The angle was changed in increments of 45 degrees for different measurement conditions. Figure 3 shows the orientation of the model with respect to the wind directions and Fig. 4 presents the tent halls in the measuring position.

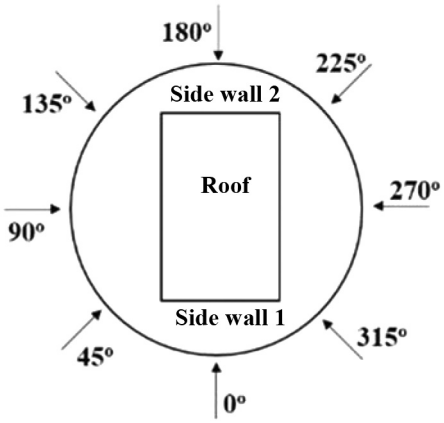


Fig. 3. Orientation of the model with respect to the wind directions

Due to the symmetry of the tested tent halls, the measurement points were distributed only on half of the model. At each of these points, wind velocity pressures as a function of time were measured. The wind velocity reference pressure (q_{ref}) was measured at the reference point as a difference of the total pressure obtained from a Prandl tube and the static pressure obtained from a static pressure probe.

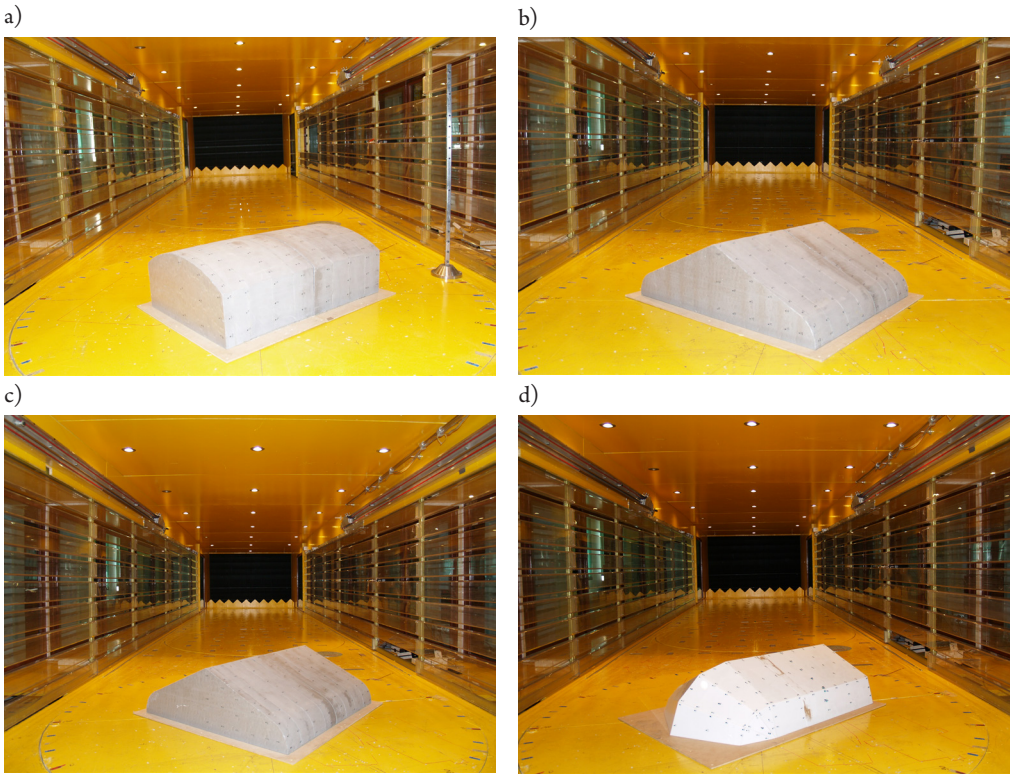


Fig. 4. View of the tent hall models in the working section of the boundary layer wind tunnel: a) model A, b) model B, c) model C, d) model D

The measurements were performed using the tent hall models equipped with pressure sensors distributed at various heights above the ground on the outer walls of the models. The sensors have been connected to the pressure scanners making possible simultaneously collection of instantaneous wind pressure time series. The scan rate of each series was 200 Hz and data was recorded over a period of 30 s. A 64-channel parallel type pressure scanner was used for the measurements.

Measurements at each point were then interpolated to obtain the distribution of wind pressures on the whole roof; thus, the pressure coefficients were determined for different tent hall cross sections.

4. Basic denotations and definitions

The wind pressures were measured on the external surfaces of the roofs and side walls of the models. The mean wind pressure coefficients were calculated according to the formula:

$$C_{pe} = \frac{p_e}{q_{ref}} \tag{1}$$

where:

p_e – mean wind pressure,

q_{ref} – reference pressure of the onflowing air at the model height.

Minimum and maximum local values of wind pressure coefficients were calculated on the basis of wind pressure and the standard deviation of mean wind pressure using the following formulas:

$$C_{pe}^{\max,l} = \frac{p_e + g_p^l \cdot p_e^\sigma}{q_{ref}};\ C_{pe}^{\min,l} = \frac{p_e - g_p^l \cdot p_e^\sigma}{q_{ref}} \tag{2}$$

where:

p_e^σ – standard deviation of instantaneous wind pressure $p_e(t)$;

g_p^l – local peak pressure factor.

Coefficients calculated according formulae (1) and (2) were determined for 8 angles of wind flow in accordance with Fig. 3.

In the structural design of the main bearing structures (frames) of tent halls, the following important questions should be considered:

1. Maximum and minimum local wind pressures never occur simultaneously at different points on the surface of a tent hall. Thus, in the design process of the main bearing structures (frames) of the tent halls, some average values of local wind pressures can be adopted.
2. Tent halls are temporary objects; therefore, using a simplified deterministic approach to calculate the wind action on these engineering objects is justified.
3. Usually, the load-bearing frames of a tent hall are similar in terms of geometry, construction material, solutions of structural nodes, and cross sections of frame columns and transoms.
4. The number of frames and number of air onflow cases for each tent hall requires around 200 calculations for static issues alone, which are related to mean wind action.

The number of calculation significantly increases if combinations of the following main actions on the structure are taken into account: self-weight, live loads, wind action, snow load, thermal actions.

5. Wind action is a random spatial-temporal process. In order to properly calculate such an action, it is necessary to know the spatial-temporal or space-frequency characteristics of these processes which can be obtained, for example, in model tests in wind tunnels, which extremely complicate not only the model tests but also the subsequent processing and analysis of tests results.

Taking all this into account, it is appropriate to use a simplified calculation approach of wind action on tent halls based on the quasi-static deterministic model proposed in this paper. This is characterised by the following assumptions:

1. The basis for determining the deterministic quasi-static actions is formula (1) and formula (2) in which the local peak pressure factor g_p^l related to the bearing structures (frames) of the investigated tent halls is of the order of 1. Thus, the following formulae were used in further considerations:

$$C_{pe}^{\max,l} = \frac{p_e + p_e^{\sigma}}{q_{ref}};\ C_{pe}^{\min,l} = \frac{p_e - p_e^{\sigma}}{q_{ref}} \tag{3}$$

2. The maximum and minimum values from all coefficients – $C_{pe}^{\max,l}$ and – $C_{pe}^{\min,l}$ are determined for each wind direction and for a given frame – these are denoted as: $C_{pe}^{\max}$ i $C_{pe}^{\min}$.
3. A similar procedure can be used for the average values of the coefficient C_{pe} determining the maximum and minimum values for each wind direction and for a given frame: $C_{pe,\max}$ and $C_{pe,\min}$.
4. The differences of the extreme coefficients are then determined, i.e.:

$$\left(C_{pe}^{\max} - C_{pe,\max}\right) \text{ and } \left(C_{pe}^{\min} - C_{pe,\min}\right) \tag{4}$$

These coefficients are schematically presented in Fig. 5.

5. The above quantities are the basis for determining the equivalent surface action for a given frame from according to the formulas:

$$p_{eq}^{\max} = \left(C_{pe} + C_{pe}^{\max} - C_{pe,\max}\right) q_{ref} \tag{5}$$

$$p_{eq}^{\min} = \left(C_{pe} + C_{pe}^{\min} - C_{pe,\min}\right) q_{ref} \tag{6}$$

6. Equivalent, maximal and minimal wind action for a given frame and for a given wind direction, per unit length of the frame column or transom, are determined by the formulas:

$$w_{eq}^{\max} = p_{eq}^{\max} L\gamma \tag{7}$$

$$w_{eq}^{\min} = p_{eq}^{\min} L\gamma \tag{8}$$

where: L – distance between adjacent span frames or half of the distance between the outer frame and the adjacent frame; γ – wind action coefficient (partial safety coefficient), which is proposed to be adopted at the level of 1.5, due to the adopted, simplified deterministic model of wind action on the tent halls. This coefficient according to Eurocode PN/EN 1991-1-4, is usually assumed at a level of 1.3 for typical buildings.

Taking into account the above remarks, coefficients: C_{pe}, $C_{pe}^{\max}$, $C_{pe\,\max}$, $C_{pe}^{\min}$, $C_{pe\,\min}$ were considered in the further data processing and analysis of tests results.

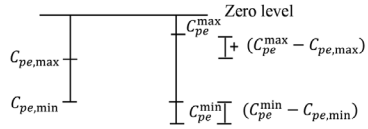


Fig. 5. Schematic presentation of respective extreme coefficients

5. Results of the wind tunnel tests

The results in Figs. 6–9 present distributions of mean wind pressure coefficients on the external surfaces of all the tent hall models. Because of the symmetry of the model, only results at 3 angles of wind attack (0°, 45°, 90°) for the A, B and C models, and at 5 angles of

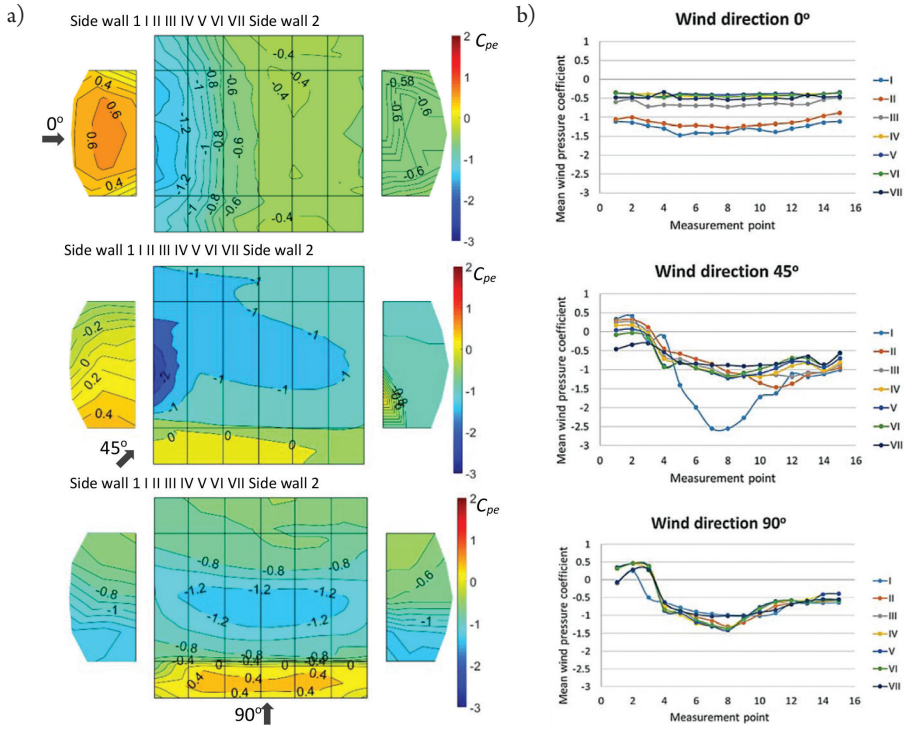


Fig. 6a) contour map of C_{pe} distributions on external surfaces; b) mean wind pressure coefficient for all cross sections of tent hall model 'A' at wind directions 0°, 45°, 90°

wind attack (0°, 45°, 90°, 135°, 180°) for the D model have been illustrated and subsequently analysed. Furthermore, the extreme values of the mean wind pressure coefficients (C_{pe}) and the peak wind pressure coefficients (C_{pe}^{max}, C_{pe}^{min}) for both of the side walls, the edge cross sections and the middle cross sections are summarised in Tables 1–4.

Table 1. The extreme values of the mean wind pressure coefficients C_{pe} and peak wind pressure coefficients C_{pe}^{max}, C_{pe}^{min}, on the side walls, edge cross section and middle cross section of model 'A'

Angle [°]	Cross section I (edge)				Cross section IV (middle)			
	C_{pe}		C_{pe}^{max}	C_{pe}^{min}	C_{pe}		C_{pe}^{max}	C_{pe}^{min}
	max	min			max	min		
0	–1.1	–1.5	–0.2	–2.3	–0.4	–0.5	–0.1	–0.7
45	0.4	–2.5	0.8	–2.6	0.2	–1.2	0.4	–1.3
90	0.3	–1.0	0.6	–1.6	0.5	–1.4	0.7	–1.5
Angle [°]	Side wall 1				Side wall 2			
	C_{pe}		C_{pe}^{max}	C_{pe}^{min}	C_{pe}		C_{pe}^{max}	C_{pe}^{min}
	max	min			max	min		
0	0.7	0	1.0	–0.2	–0.6	–0.6	–0.5	–0.8
45	0.5	–0.5	0.8	–0.6	–0.8	–1.0	–0.7	–1.1
90	–0.5	–1.4	–0.3	–1.8	–0.5	–1.4	–0.2	–1.6

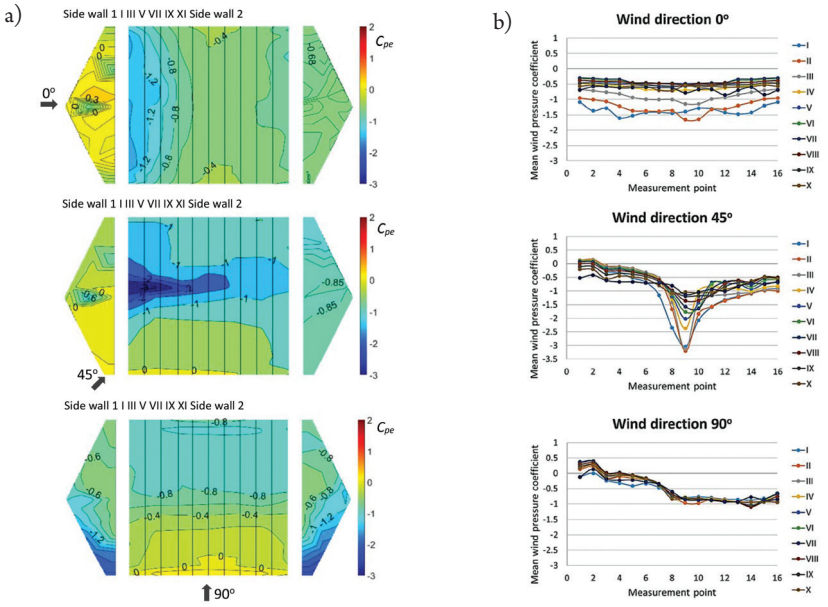


Fig. 7a) contour map of C_{pe} distributions on external surfaces; b) mean wind pressure coefficient for all cross sections of tent hall model 'B' at wind directions of 0°, 45°, 90°

Table 2. The extreme values of the mean wind pressure coefficients C_{pe} and peak wind pressure coefficients C_{pe}^{max}, C_{pe}^{min}, on the side walls, edge cross section and middle cross section of model 'B'

Angle [°]	Cross section I (edge)				Cross section VI (middle)			
	C_{pe}		C_{pe}^{max}	C_{pe}^{min}	C_{pe}		C_{pe}^{max}	C_{pe}^{min}
	max	min			max	min		
0	–1.1	–1.6	–0.4	–2.1	–0.3	–0.5	–0.2	–0.6
45	0.1	–3.0	0.6	–3.2	0.1	–1.8	0.2	–1.9
90	0	–0.9	0.3	–1.2	0.4	–1.1	0.6	–1.2

Angle [°]	Side wall 1				Side wall 2			
	C_{pe}		C_{pe}^{max}	C_{pe}^{min}	C_{pe}		C_{pe}^{max}	C_{pe}^{min}
	max	min			max	min		
0	0.4	–0.7	0.9	–0.7	–0.6	–0.7	–0.4	–0.7
45	0.3	–1.1	1.1	–1.1	–0.8	–1.1	–0.6	–1.1
90	–0.4	–2.1	0.2	–2.1	–0.5	–2.4	–0.1	–2.4

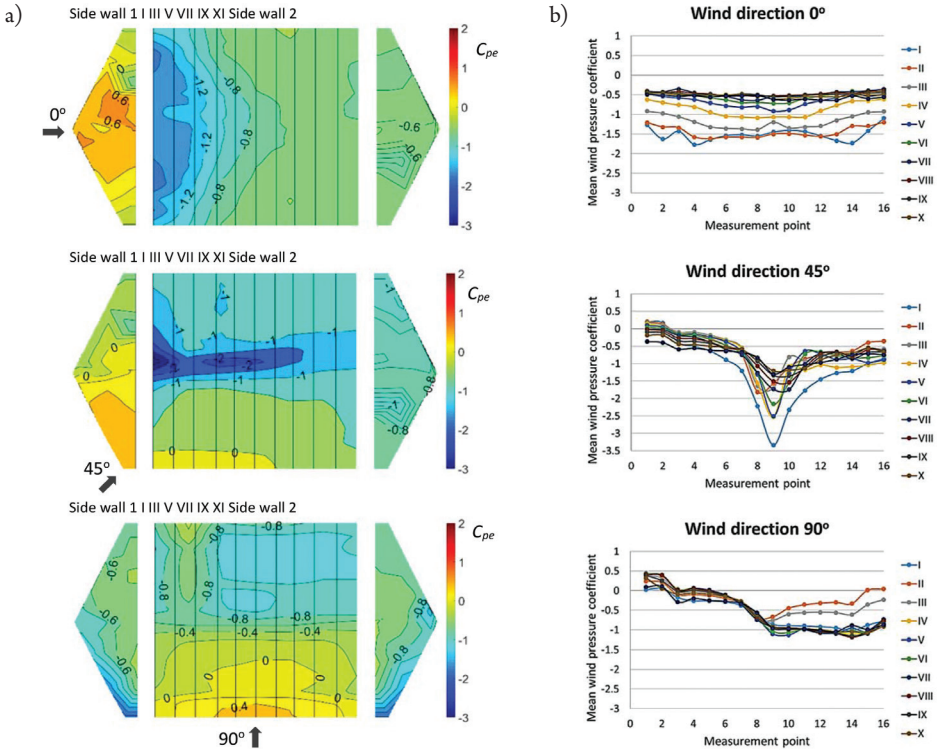


Fig. 8a) contour map of C_{pe} distributions on external surfaces; b) mean wind pressure coefficient for all cross sections of tent hall model 'C' at wind directions of 0°, 45°, 90°

Table 3. The extreme values of the mean wind pressure coefficients C_{pe} and peak wind pressure coefficients C_{pe}^{max}, C_{pe}^{min}, on the side walls, edge cross section and middle cross section of model 'C'

Angle [°]	Cross section I (edge)				Cross section VI (middle)			
	C_{pe}		C_{pe}^{max}	C_{pe}^{min}	C_{pe}		C_{pe}^{max}	C_{pe}^{min}
	max	min			max	min		
0	–1.1	–1.8	–0.7	–3.6	–0.4	–0.7	–0.2	–1.1
45	0.2	–3.3	0.6	–3.6	0	–2.2	0.2	–2.5
90	0.1	–1.0	1.3	–1.7	0.4	–1.1	0.6	–1.2

Angle [°]	Side wall 1				Side wall 2			
	C_{pe}		C_{pe}^{max}	C_{pe}^{min}	C_{pe}		C_{pe}^{max}	C_{pe}^{min}
	max	min			max	min		
0	0.7	–0.2	1.1	–0.6	–0.4	–0.8	–0.4	–0.9
45	0.5	–0.4	0.8	–1.0	–0.6	–1.0	–0.6	–1.3
90	–0.3	–1.9	0	–2.6	–0.3	–1.8	–0.1	–2.5

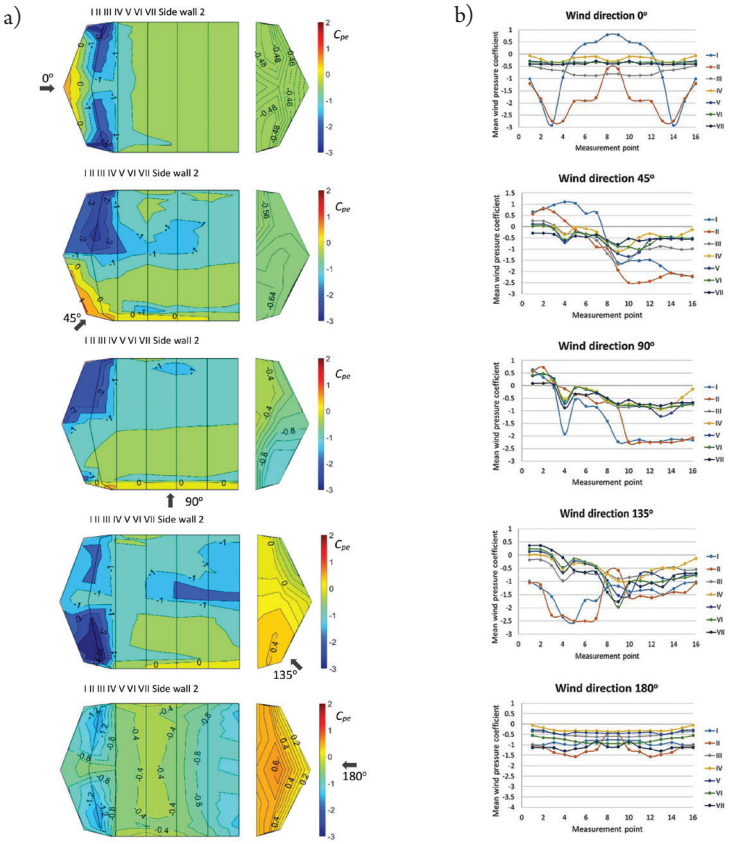


Fig. 9a) contour map of C_{pe} distributions on external surfaces; b) mean wind pressure coefficient for all cross sections of tent hall model 'D' at wind directions of 0°, 45°, 90°

Table 4. The extreme values of the mean wind pressure coefficients C_{ep} and peak wind pressure coefficients C_{pe}^{max}, C_{pe}^{min} on the side walls, edge cross section and middle cross section of model 'D'

Angle [°]	Cross section III (edge)				Cross section V (middle)			
	C_{pe}		C_{pe}^{max}	C_{pe}^{min}	C_{pe}		C_{pe}^{max}	C_{pe}^{min}
	max	min			max	min		
0	–0.5	–0.9	–0.4	–1.0	–0.3	–0.4	–0.2	–0.5
45	0.3	–1.7	0.4	–2.0	0.1	–1.3	0.3	–1.6
90	0.5	–0.9	0.7	–1.1	0.5	–1.2	0.7	–1.5
135	–0.2	–1.0	–0.1	–1.3	0.1	–1.5	0.3	–2.0
180	–0.4	–0.6	–0.3	–0.8	–0.3	–0.5	–0.1	–0.8

Angle [°]	Side wall 1 (cross section I, II)				Side wall 2			
	C_{pe}		C_{pe}^{max}	C_{pe}^{min}	C_{pe}		C_{pe}^{max}	C_{pe}^{min}
	max	min			max	min		
0	0.8	–2.9	1.4	–3.0	–0.4	–0.5	–0.4	–0.6
45	1.1	–2.5	1.6	–2.7	–0.5	–0.7	–0.4	–0.8
90	0.7	–2.2	1.3	–2.4	–0.3	–1.1	–0.1	–1.6
135	–0.6	–2.5	–0.3	–2.7	0.4	–0.4	0.6	–0.5
180	–0.5	–1.6	–0.3	–1.9	0.6	0.2	0.9	–0.1

In the case of the A model, the following conclusions can be drawn from the measurement analysis: the highest negative pressure appears in the half part of all the cross sections. The highest value of the C_{pe}^{min} coefficient is –2.6 – this occurred in the middle part of the roof with a wind direction of 45° in the first cross section. The highest value of the C_{pe}^{max} coefficient is 1.0 – this occurred on side wall I with a wind direction of 0°. The highest wind action occurred at the wind attack angle of 45° for edge the cross section and 90° for middle cross section.

In the case of the B model, the highest value of the C_{pe}^{min} coefficient is –3.2 – this occurred at the top of the roof and with a wind direction of 45° in the first cross-section of tent hall. The highest value of the C_{pe}^{max} coefficient is 1.1 – this occurred on side wall I with a wind direction of 45°. The highest wind action occurred with wind attack angles of 45° for both the edge cross section and the middle cross section.

The shape of tent hall *C* is similar to that of the previous hall (B), they differ only in the length with hall B being longer than hall *C*. The difference for both objects is negligible. The highest value of the C_{pe}^{min} coefficient is –3.6 – this occurred at the top of the roof with a wind direction of 45° and in the first cross section of the tent hall. The highest value of the C_{pe}^{max} coefficient is 1.3 – this occurred on the side wall with a wind direction of 90° and in the first cross section of the tent hall.

On the basis of C_{pe}^{min} the measurement results obtained for tent hall D, the following conclusions can be formulated: the highest value of coefficient is –3.0 – this occurred at a wind direction of 0° in the first cross section of the tent hall. The highest value of the C_{pe}^{max} coefficient is 1.6 – this occurred at a wind direction of 45° and in the first cross section of tent hall D.

Peak wind pressure coefficients C_{pe}^{max}, C_{pe}^{min} on the external surfaces for all tent halls are presented in Table 5.

Table 5. Peak wind pressure coefficients C_{pe}^{max}, C_{pe}^{min} on the external surfaces of the tent halls

Tent hall		A		B		C		D	
Wind direction	Surface	C_{pe}^{max}	C_{pe}^{min}	C_{pe}^{max}	C_{pe}^{min}	C_{pe}^{max}	C_{pe}^{min}	C_{pe}^{max}	C_{pe}^{min}
0°	side wall 1	1.0	–0.2	0.9	–0.7	1.1	–0.6	1.4	–3.0
	roof	0.5	–2.3	0.2	–2.2	–0.1	–2.1	0	–1.0
	side wall 2	–0.5	–0.8	–0.4	–0.7	–0.4	–0.9	–0.4	–0.6
45°	side wall 1	0.8	–0.6	1.1	–1.1	0.8	–1.0	1.6	–2.7
	roof	0.8	–2.6	0.6	–3.3	0.6	–3.1	0.4	–2.0
	side wall 2	–0.7	–1.1	–0.6	–1.1	–0.6	–1.3	–0.4	–0.8
90°	side wall 1	–0.3	–1.8	0.2	–2.1	0	–2.6	1.3	–2.4
	roof	0.8	–1.6	0.6	–1.6	1.3	–1.8	0.7	–1.5
	side wall 2	–0.2	–1.6	–0.1	–2.4	–0.1	–2.5	–0.1	–1.6

Figure 10 presents a comparison of the measured extreme values of the peak wind pressure coefficients C_{pe}^{max}, C_{pe}^{min} on the side walls and roof of all of the tent halls. For the wind direction perpendicular to the side wall (0°), the pressure distribution on the analysed roofs slightly depends on the shape of the tent hall. In the edge cross section accrues the lowest value of the mean pressure wind pressure coefficients up to –1.8 for the C tent hall.

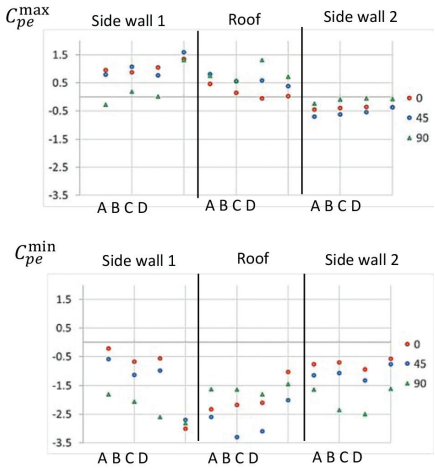


Fig. 10. Distribution of the extremal value of C_{pe}^{max}, C_{pe}^{min} coefficients on the side walls and roof of tent hall models A, B, C, D

Maximum negative wind pressure appears in the edge cross section for all of the analysed halls. For the A, B, and C tent halls, the maximum negative pressure occurs in the middle of the cross section. Tent hall D presents slightly different results because the highest negative pressure can be observed in the lower part of the wall.

The mean external wind pressure coefficient has extreme values in the case of hall A for angles of wind onflow of 90° for the middle cross section and 45° for the edge cross sections. During the analysis of the results, differences appearing for C_{pe} for B and C can also be observed. These halls have similar geometry – they differ only in the length with hall B being longer than hall C. The negative pressure for C is higher than for B and this trend is visible in results for all of the cross sections. It can be concluded that the length of the hall has an influence on the mean external wind pressure coefficient – the smaller the length, the higher the value of C_{pe}.

To summarise, in most of the analysed cases, the extremal mean wind pressure coefficient appears in the middle part of the edge cross section of tent halls. The shape of the hall cross section has a significant influence on the value of this coefficient. The highest values of C_{pe} appears for the 45° angle of wind onflow; therefore, it is essential to take into consideration this information in the design of such structures.

6. Conclusions

The measurements of wind pressure distributions on the surfaces of tent halls of different shapes allow the following conclusions to be drawn:

- The shape of the tent hall has a significant influence on the wind action on the analysed surfaces.
- The highest values of negative wind pressure appear on the roof edge, especially in the edge cross section for an angle of wind attack of 45°. Sharp roof edges generate higher values of wind pressure coefficients.
- The highest values of positive wind pressure occur on the longitudinal side walls for wind onflowing at an angle of 90°.

References

[1] Alrawashdeh H., Stathopoulos T., *Wind pressures on large roofs of low buildings and wind codes and standards, J. Wind Eng. Ind. Aerodyn.* 147, 2015, 212–225.

[2] Gavanski E., Uematsu Y., *Local wind pressures acting on walls of low-rise buildings and comparisons to the Japanese and US wind loading* provisions, J. Wind Eng. Ind. Aerodyn., 132, 2014, 77–91.

[3] Hoxey R.P., Moran P., *A full-scale study of the geometric parameters that influence wind loads on low-rise buildings*, J. Wind Eng. Ind. Aerodyn., 13 (1–3), 1983, 277–288.

[4] Kim K.C., Ji H.S., Seong S.H., *Flow structure around a 3-D rectangular prism in a turbulent boundary layer*, J. Wind Eng. Ind. Aerodyn., 91, 2003, 653–669.
[5] PN-EN 1991-1-4: Action on structures. Part 1–4: General action (in Polish).
[6] Pratt F.N., Kopp G.A., *Velocity field measurements above the roof of a low-rise building during peak suctions*, J. Wind Eng. Ind. Aerodyn., 133, 2014, 234–241.
[7] Saathoff P.J., Melbourne W.H., *Effects of free-stream turbulence on surface pressure fluctuations in a separation bubble*, J. Fluid Mech., 337, 1997, 1–24.
[8] Stathopoulos T., *Wind loads on low-rise buildings – A review of the state of the art*, Eng. Struct., 6, 1984, 119–135.
[9] Tielman H.W., *Wind tunnel simulation of wind loading on low-rise structures: a review*, J. Wind Eng. Ind. Aerodyn., 91, 2003, 1627–1649.

TECHNICAL TRANSACTIONS 3/2018
CZASOPISMO TECHNICZNE 3/2018
ELECTRICAL ENGINEERING
DOI: 10.4467/2353737XCT.18.040.8302
SUBMISSION OF THE FINAL VERSION: 5/3/2018

Ilona Lagun
Adrian Nakonechnyy
Department of Computer Systems and Automatic, Lviv Polytechnic National University, Ukraine

Volodymyr Samotyy (vsamotyy@pk.edu.pl)
Department of Automatic Control and Information Technology, Cracow University of Technology

Using fuzzy logic to optimise the selection of mother wavelets

Wykorzystanie logiki rozmytej dla optymalizacji wyboru falek bazowych

Abstract

The effectiveness of signal processing using discrete wavelet transformation depends on the correct choice of basic wavelet function. A model of multi-criteria optimisation of selection of wavelet function was developed based on fuzzy logic. An experimental study of the model was carried out for the most common types of signals selected from the MATLAB database package.

Keywords: selection of criteria, wavelet function, multi-criteria optimisation, fuzzy set, fuzzy criteria

Streszczenie

Skuteczność przetwarzania sygnału przy użyciu dyskretnej metody falkowej zależy bezpośrednio od prawidłowego wyboru falki bazowej. Opracowano model optymalizacji multicriterialnej doboru falki bazowej na podstawie logiki rozmytej. Wykonano badania eksperymentalne opisanego modelu dla najbardziej popularnych typów sygnałów wybranych z środowiska MATLAB.

Słowa kluczowe: wybór kryteriów, funkcja falkowa, optymalizacja wielokryterialna, zbiór rozmyty, kryteria rozmyte

1. Introduction

The efficiency of processing signals using a discrete wavelet transform is directly dependent upon the selection of the corresponding basic wavelet functions [12].

When choosing a mother wavelet, mainly, properties such as carrier size, the number of zero moments, and the smoothness of the basic functions need to be taken into account. However, such characteristics only have mathematical descriptions of the basic functions and do not enable the obtaining of explicit recommendations for their practical use for the analysis and processing of certain types of signals. Thus, it is expedient to form a certain set of wavelet functions, the optimal choice of which will be determined by certain criteria.

2. Formulation of the problem

In work [1], the properties of wavelets with final and infinite support are generalised. This approach facilitates the basic wavelet function selection process which is dependent upon usage. The amount recommendations reduce the error possibility during the process of basic wavelet selection. Sometimes, standard wavelet plural functions are used [2] – these include Haar (db1), Daubechies (db), Coiflets (coif) and Symlet (sym). Unfortunately, the rules of basic wavelet function selection have not been established in [2]. The basic wavelet function choice is executed using the previous experience. The basic wavelet function selection remains a problem. The best basic wavelet function selection has been executed for Electroencephalogram (EEG) data plural. Several standard wavelet plural including Daubechies, Symlet, Coiflets, Morlet Mexicanhat and Meyer have been reviewed. The most widespread wavelet functions for biomedical signal processing have also been reviewed. The optimal choice for the basic function is absent. In work [3], the choice of basic wavelet could be used for increasing accuracy of approximation and automation speed increasing To improve electrocardiogram (ECG) approximation accuracy the analyses of the basic wavelet functions Cubic Spline Wavelet, Haar Wavelet, Db4 Wavelet and Db6 Wavelet have been conducted. The comparative analysis have been performed as a table. The analysis of different wavelet functions for discrete wavelet transformation is shown in [9]. It is used for watermarks drawing. In [10], the automation of wavelet function attempt have been occurred during gear wheels and bearings diagnostics. The 324 wavelet function has been investigated. The (Daubechies) (db44) function have the most suitable form for this case. The algorithm of automatic basic function choice have been performed. The problem of the silence of EEG signals using wavelet transformation has been reviewed in [13]. The investigation has been performed using the standard EEG data set. To improve the characteristics, the wavelet functions choice db4, sym7, coif3, bior3.9. have been executed. The wavelet function choice was made on the basis of five criteria: MSE, RMSE, SNR, the improved SNR and PRD. In this article, it is emphasised that using this criteria, the unambiguous choice of required function could not be achieved. In this work for unambiguous choice of wavelet functions optimization the fuzzy logic have been used.

Currently, in the technique of signal processing, criteria have been developed based on energy, entropy and correlation dependences [7]: Energy to Shannon Entropy Relations (EER)); entropy (Mutual Information to Relative Entropy Relations (IER)) and correlation (correlation coefficient (Cr)). The energy criterion foresees the calculation of the following ratio:

$$\mathrm{EER} = \frac{E_C}{En} \tag{1}$$

where:

E_C – signal energy in the wavelet space;

En – Shannon entropy for the wavelet coefficients of the signal decomposition.

The entropy criterion is based on the calculation of the next dependence

$$Cr = \frac{\mathrm{COV}_{S\psi}}{\sigma_S \cdot \sigma_\psi} \tag{2}$$

where:

$\mathrm{COV}_{S\psi}$ – the mutual covariance of discrete signal sequences and the basic wavelet function;

σ_s, σ_ψ – the standard deviations of these sequences.

The correlation criterion involves calculating the relationship:

$$\mathrm{IER} = \frac{I(S,C)}{D(S||C)} \tag{3}$$

where:

$I(S; C)$, $D\left(S||C\right)$ – mutual information and relative entropy between the signal and its wavelet coefficients, respectively.

The analysis showed that it is not always possible to reach the uniqueness of the choice of the basic wavelet functions by the above-mentioned criteria [7]. Therefore, a more generalised criterion is required. At present, the solution to the multi-criteria optimisation problem, in which it is impossible to totally optimize all conflicting criteria by 100%, but only each of them to some extent, allows the use of fuzzy logic. The purpose of this work is to analyse basic methods of optimisation and construct a multi-criterial optimisation model for choosing basic wavelet functions under uncertainty.

3. Multi-criteria optimisation choice of basic wavelet functions

In general, the fuzzy solution $\tilde{D}$ is the result of the intersection of local criteria $\tilde{G}_1 \div \tilde{G}_3$. In the case of non-equilibrium criteria, we have:

$$\tilde{D}=\left\{\frac{\min\limits_{i=1,n}\left(\mu_{G_i}\left(x_1\right)\right)^{w_i}}{x_1},\frac{\min\limits_{i=1,n}\left(\mu_{G_i}\left(x_2\right)\right)^{w_i}}{x_2},\cdots,\frac{\min\limits_{i=1,n}\left(\mu_{G_i}\left(x_k\right)\right)^{w_i}}{x_k}\right\} \tag{4}$$

where:

w_i – the coefficient of relative importance of the criterion G_i.

Weights are normalised with the following condition:

$$w_1+w_2+w_3+\ldots+w_n=1. \tag{5}$$

To determine the coefficients w_i, comparative criteria matrices are formed. As a result, the best basic wavelet function is the function with the highest degree of membership:

$$\mu_D\left(x^*\right)=\max_{i=1,2,3,\ldots,n}\mu_D\left(x_i\right) \tag{6}$$

Since the dimensions of the criteria and their scale are different, it is impossible to compare their quality. In view of this, it is necessary to make all scales the same and dimensionless by the following normalisation [5]:

$$G_i=\frac{G_i-G_{i,\min}}{G_{i,\max}-G_{i,\min}}. \tag{7}$$

The next step is to present the criteria in the form of a fuzzy set on the universal set of basic wavelet functions X [4].

$$\tilde{G}_i=\left(\frac{\mu_{G_i}\left(x_1\right)}{x_1},\frac{\mu_{G_i}\left(x_2\right)}{x_2},\cdots,\frac{\mu_{G_i}\left(x_k\right)}{x_k}\right) \tag{8}$$

where:

$\mu_{G_i}\left(x_j\right)$ – a number in the range '0-1', the degree of membership of the basic low-wavelength function fuzzy set $\tilde{G}_i$. The higher the value $\mu_{G_i}\left(x_j\right)$, the higher the estimation of the base wavelet function x_j by the criterion G_i.

Each of the criteria can be represented as a simple fuzzy value of G_i and a linguistic variable $G_i=\left\{G_{iL},G_{iM},G_{iH}\right\}$, where L, M and H denote the concept of 'low', 'middle' and 'high' for the corresponding i-th criterion.

On the basis of the analysis of the effectiveness of the criteria performed on the results of the processing of different types of signals, the following expert pair comparisons are formed:

1. Weak advantage of G_2 over G_1;
2. Significant advantage of G_2 over G_3;
3. Significant advantage of G_1 over G_3.

The practical realisation of the multi-criterial optimisation problem of choosing a base wavelet function is performed on the basis of the FIS-editor of systems of fuzzy output from

the Fuzzy Logic Toolbox, which is part of the package of applied mathematical modelling software Matlab R2011b.

It is known that in the process of constructing fuzzy-out systems, the methods of Mamdani and Sugeno have become the most commonly used. An analysis of both methods showed the feasibility of using a fuzzy model based on the Mamdani method.

The level of effectiveness of the criterion for choosing basic wavelet functions can be described as follows: low (L), middle (M), high (H). At the phasing stage, membership functions for the sets of input and output linguistic variables are given. The set of criteria $G = \{G_1, G_2, G_3\}$ is represented by three linguistic variables:

$$EER=\{EER\_L, EER\_M, EER\_H\};$$

$$IER=\{IER\_L, IER\_M, IER\_H\}; Cr=\{Cr\_L, Cr\_M, Cr\_H\};$$

Subsequently, the sigmoid function is selected as a membership function for each input linguistic variable and the trapezoidal function is selected as a membership function for the output linguistic variable. In the next stage, a rule base is formed in the form of a structure with three inputs and one output based on expert pair comparisons (Table 1).

Table 1. Rules for forming a multi-criterial optimisation model for choosing a basic wavelet function

No.	Entrance			Exit
	G_1(EER)	G_2(IER)	G_3(Cr)	Y(WV)
1	2	3	4	5
1	low	low	low	low
2	low	low	middle	low
3	low	low	high	low
4	low	middle	low	low
5	low	middle	middle	middle
6	low	middle	high	middle
7	low	high	low	middle
8	low	high	middle	middle
9	low	high	high	middle
10	middle	low	low	low
11	middle	low	middle	low
12	middle	low	high	middle
13	middle	middle	low	middle
14	middle	middle	middle	middle

tab. 1 (cont.)

1	2	3	4	5
15	middle	middle	high	middle
16	middle	high	low	middle
17	middle	high	middle	middle
18	middle	high	high	high
19	high	low	low	low
20	high	low	middle	middle
21	high	low	high	middle
22	high	middle	low	middle
23	high	middle	middle	middle
24	high	middle	high	middle
25	high	high	low	high
26	high	high	middle	high
27	high	high	high	high

Finally, in the process of dephasing is used the centre of weight method for a discrete set of values of the membership function. At the dephasing stage, the implementation of a fuzzy output system makes it possible to obtain an estimate of the efficiency of the basic wavelet function.

4. Research results

As a result of the construction of the model, the dependence of the level of efficiency of the basic wavelet functions on the two criteria in the form of surfaces was obtained (Fig. 1) wavelet functions, which are defined as optimal for processing each test signal. For the evaluation of the developed models, families of orthogonal functions were selected with a compact Daubechies carrier (db1 ... db20), Coiflets (coif1 ... coif5), Symlets (sym1 ... sym20) and test signals from the Matlab package: blocks, bumps, doppler, heavy sine, trsin, wcantor [5].

For each of the test signals and the selected basic functions, the Matlab package evaluates the values for the energy, entropy and correlation criteria. As a result, arrays of values for each of the criteria are formed. For each array of values, the developed multi-criteria optimisation model was applied and appropriate dependencies were obtained.

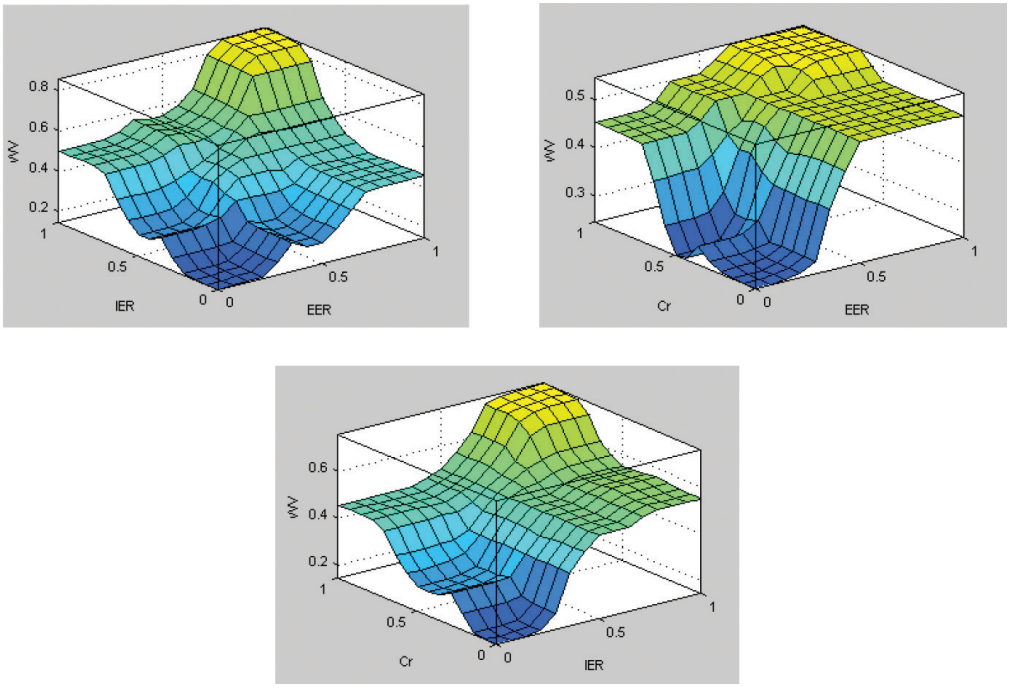


Fig. 1. The surface of a fuzzy model system relative to: a) input variables EER, IER; b) the input variables EER, Cr; c) the input variables Cr, IER

Table 2. Effective base wavelet functions, selected on the basis of multi-criteria optimisation for processing test signals

Test signals	Base wavelet functions
'blocks'	db1, sym1
'bumps'	sym15, sym19
'doppler'	sym9, sym13
'heavy sine'	db2, db3
'trsin'	sym7
'wcantor'	db15, db17, db18, db19, db20

5. Conclusions

In this paper, the solution of the ambiguity problem of the choice of basic wavelet functions is proposed using various criteria by means of multi-criteria optimisation with the use of fuzzy logic.

The paper deals with the analysis of algorithms for solving the multi-criterial optimisation problem on the basis of the theory of fuzzy sets. The method of constructing a multi-criterial optimisation model using linguistic variables based on the Mamdani algorithm. As a function of

the membership of input variables, a sigmoid function is selected; as a function of the membership of the output variable, the trapezoidal function is selected. Rules for the model of the multi-criteria optimisation of the choice of the basic wavelet function were constructed on the basis of expert information according to the results of previous studies. The practical implementation of the model is based on the FIS-editor of Fuzzy Logic Toolbox fuzzy output systems, which is part of the Matlab R2011b software package of mathematical modelling. The developed model was used to determine the efficiency of the basic wavelet functions for six test signals of the Matlab package – this allowed the selection of optimal functions for the further elaboration of each test signal.

References

[1] Ahuja N., Lertrattanapanich S., Bose N.K., *Properties determining choice of mother wavelet*, IEE Proceedings – Vision, Image and Signal Processing, Vol. 152, 2005, 659–664 .

[2] Al-Qazzaz N.K., Bin Mohd Ali S.H., Ahmad S.A., Islam M.S., Escudero J., *Selection of Mother Wavelet Functions for Multi-Channel EEG Signal Analysis during a Working Memory Task*, Sensors, 2015, 15, 29015–29035.

[3] De Moortel I., Munday S.A., Hood A.W., *Wavelet Analysis: the effect of varying basic wavelet parameters*, Solar Physics, Vol. 222, 2004, 203–228.

[4] Chong E.K.P., Zak S.H., *An Introduction to Optimization*, 4th Edition, John Wiley & Sons, 2013.

[5] Hao Ying, *Fuzzy Control and Modeling: Analytical Foundations and Applications*, Wiley-IEEE Press, 2000.

[6] Kapil Tajane, Rahul Pitale, Jayant Umale, *Comparative Analysis of Mother Wavelet Functions with the ECG Signals*, Journal of Engineering Research and Applications, Vol. 4, 2014, 38–41.

[7] Lagun I., Nakonechny A. *Selection of wavelet basis for the effectiveness processing of signals,* Vestnik Brestskogo gosudarstvennogo tekhnicheskogo universiteta, No. 5, 2016, 69–73.

[8] Mallat S., *Wavelet tour of signal processing*, Third edition: The Sparse Way, 2008.

[9] Navdeep Goel, Gurwinder Singh, *Study of Wavelet Functions of Discrete Wavelet Transformation in Image Watermarking,* An International Journal of Engineering Sciences, Vol. 17, 2016, 154–160.

[10] Rafiee J., Rafiee M.A., Tse P.W., *Application of mother wavelet functions for automatic gear and bearing fault diagnosis*, Expert Systems with Applications, Vol. 37, 2010, 4568–4579.

[11] Reiner Horst, Tuy Hoang, *Global Optimization: Deterministic Approaches*, 3rd edition, Springer-Verlag GmbH, 1996.

[12] Thomas Weise, *Global optimization algorithms: theory and application*, 3nd Edition, Thomas Weise, 2011.

[13] Zaid Abdi Alkareem Alyasseri, Ahamad Tajudin Khader, Mohammed Azmi Al-Betar, *Electroencephalogram Signals Denoising using Various Mother Wavelet Functions: A Comparative Analysis*, Proceedings of the International Conference on Imaging, Signal Processing and Communication, At Penang, Malaysia, October 27, 2017, 100–105.

TECHNICAL TRANSACTIONS 3/2018
CZASOPISMO TECHNICZNE 3/2018
ELECTRICAL ENGINEERING
DOI: 10.4467/2353737XCT.18.041.8303
SUBMISSION OF THE FINAL VERSION: 12/2/2018

Krzysztof Wójcik (krzysztof.wojcik@mech.pk.edu.pl)
Krzysztof Wójcik, Production Engineering Institute, Faculty of Mechanical Engineering, Cracow University of Technology

Dariusz Pałka
Department of Applied Computer Science, Faculty of Electrical Engineering, Automatics, IT and Biomedical Engineering, AGH University of Science and Technology

Olaf Bar
Institute of Computer Science, Faculty of Mathematics, Physics and Technical Science, Pedagogical University of Cracow

Micro-ontology building – the main variants of the OTO method

Metoda "oto" tworzenia mikro-ontologii – podstawowe warianty algorytmu

Abstract

This article describes the main properties of an iterative method of simple knowledge structure creation. The method is based on an inductive learning scheme. The knowledge structure is built automatically and takes the form of a simplified ontology. Knowledge transformation plays a key role in the process of creating the knowledge structure. In order to regular describe many kinds of these transformations the article provides the relevant theoretical background. The task of finding the proper ontology (knowledge structure) is extremely complex. This paper highlights the necessity to investigate efficient search methods; additionally, the work draws attention to the advantages that arise from building the knowledge structure at the minimal possible size. The paper points to possible areas of the method application, especially in connection with problems of the automatic understanding of images and websites.

Keywords: pattern recognition, image understanding, machine learning

Streszczenie

Artykuł przedstawia podstawowe własności iteracyjnego procesu (nazywanego w pracy OTO) tworzenia struktury wiedzy. Budowana automatycznie struktura przyjmuje formę ontologii. Artykuł prezentuje podstawy teoretyczne opisywanego procesu. Kluczową rolę odgrywa w nim zestaw specyficznych algorytmów transformacji wiedzy. Opisywany proces jest ekstremalnie złożony obliczeniowo. Artykuł podkreśla konieczność opracowania bardziej efektywnych algorytmów numerycznych, uwypuklając jednocześnie korzyści z budowy ontologii w minimalnej, możliwej formie (mikro-ontologia). Proces budowy wiedzy przybliżono z pomocą odpowiednio dobranego przykładu. W pracy wskazano na możliwe obszary zastosowań metody, w szczególności dotyczące automatycznego rozumienia obrazów oraz rozumienia stron WWW.

Słowa kluczowe: rozpoznawanie obrazów, rozumienie obrazów, uczenie maszynowe

1. Introduction

The World Wide Web may be considered to be an environment in which the phenomena of an observation, understanding, or action may happen. As in the real world, the phenomena may be described using a suitable language (e.g. XML, Spanish, or graph language for image representation).

The phenomena may be also modelled by the use of an ontology. The creation of the ontology involves an operation of defining objects and concepts, and establishing relationships between them [9]. Despite the existence of several tools which support this process [2], initial assumptions mainly depend on the intuition and practical knowledge of a researcher [1]. This statement remains valid regardless of the kind of ontology domain and the level of abstraction of its elements. For instance, this may relate to ontology for the analysis of texts from the World Wide Web as well as the ontology for the system which search weapons in the pictures obtained from social networking services.

To overcome the above difficulties, we propose a certain general scheme of the ontology creation. We assume that the scheme enables the knowledge to be built with minimal prior assumption (inborn knowledge); additionally, the generated system should be able to automatically update the knowledge. The proposed method constructs the knowledge structure during an iterative scheme called an OTO (observation-transformation-operation). This approach is one of many possible variants of the widely known methods of structured knowledge building (see Literature remarks). In the paper, a brief example of the method usage is outlined (see Section 3.4). This is involved with problems of building a very 'low-level' concept structure corresponding to an abstract concept of an equinumerosity of sets

The same problems also concern the automatic creation of basic concepts for website-understanding systems (text as well as images).

The proposed approach can be regarded as a kind of an algebraization of the problem. Unfortunately, the considered method suffers from a lack of efficient methods for searching the proper knowledge structure within an extremely large search space

Let us articulate the main goal of the paper. This is to encourage researchers to investigate the efficient numerical methods for searching the optimal knowledge structure.

The authors are aware that a condition of this research is providing a expanded theoretical introduction which defines the key terms – this constitutes a major part of the paper and is an extension of works [12–14]. The presentation of a consistent theoretical background is the second main goal of the article.

As has been previously noted, the knowledge creation process is based on minimal assumptions and can produce a group of fundamental (very low level) concepts (e.g. 'equinumerosity of sets' see [15]). It is difficult to compare the proposed methodology to other known methods of knowledge building which utilise the basic human-defined concepts (these operate on the upper levels only). For this reason, the method comparison is omitted. Due to a lack of efficient optimisation methods (see above), we have also omitted a comparison of the computational efficiency of our approach and existing methodologies; however, some similarities between them should be highlighted.

1.1. Literature remarks

This section presents key points that relate the proposed approach to well-known techniques of artificial intelligence (AI) – these are grouped according to some selected sub-disciplines of AI.

1.1.1. Image Understanding Techniques (IU)

From the many variants of IU methods, we will focus on techniques which rely on the extraction of semantic content from the phenomenon. This is achieved through the automatic reasoning process, which generates and verifies hypotheses regarding the phenomena [10]. This idea appeared in the 1980s, at present, however, mainly syntactic variants of this process have been developed [10, 11]. In this case, the knowledge is generated as some structured grammar [6, 7]. It can be observed that there is a strong similarity between this kind of knowledge description and the representation used in the proposed methodology (Section 3).

1.1.2. Learning from observations, inductive learning methods

According to proposed scheme, knowledge acquisition should occur automatically. One of the possible solutions is to base this process on the observation of the relationships between objects in a given reality. The first works using this approach appeared in the 1980s [4]. They introduced the idea of 'conceptual clustering' which, in a different shape, may also be regarded as a basis for our methodology. The 1990s brought the development of inductive learning methods. There was invented a methodology of Inductive Logic Programming (ILP) [5], which applies, as a formal description, a predicate calculus. The presented approach (Section 2) utilises this formalism in an analogous manner.

1.1.3. Automatic ontology building, semantic web methods

The literature [1] presents the main ideas of ontology engineering. The 'ontology', which is the central term, may be simply defined as a 3-tuple: a set of concepts, a set of relationships (between the concepts) and a set of instances [1]. The definition closely corresponds to the term 'hierarchical structure of concepts' introduced in Section 2. In recent years (especially since the beginning of this century) we have noticed a rapid growth in the use of Semantic Web methods [1]. This is caused mainly by their usage for semantic analysis and automatic understanding of the websites. There are several projects that deal with this matter, the best known of which is the OWL (Web Ontology Language) [9]. OWL is a knowledge representation language which enables the description of ontology elements like features of the classes and features of relationships, etc. It also allows the definition of some operations on the classes. The analogous operations concerning the knowledge structure are widely disputed in Section3.

2. Inductive scheme of knowledge structure building

Our first goal is to propose a method of defining the knowledge about the visible phenomenon. The created knowledge structure strongly depends on previously established assumptions. To decrease this impact, we can reduce them to a minimal form (this corresponds to the operation called 'phenomenological reduction' [3]). This reduction leads to the statement that the phenomenon consists of elementary, atomic parts – these are called primitive objects or instances. The objects correspond to primitive 'concepts' which are also called 'types' or 'classes'. The concepts are some kind of generalisation of objects (i.e. there exists a predefined, fundamental relation between objects and concepts called a membership). Furthermore, we can assume that the objects may be connected by some primitive relationships. We will use the following notations: $\mathbf{X}_0$, $\mathbf{X}$ – sets of objects; $\mathbf{C}_0$, $\mathbf{C}$ – sets of concepts; $\mathbf{D}_0$, $\mathbf{D}$ – sets of relationships between objects, in which the primitive objects, concepts, and relationships are denoted by $\mathbf{X}_0$, $\mathbf{C}_0$, and $\mathbf{D}_0$, respectively. They contain the whole initial knowledge. We can assume that all processes of knowledge increasing relate to the extension of these sets.

Since the concepts are the generalisation of objects, an occurrence of the specific objects in the given reality should be a condition of the building of new concepts (this universal idea is utilised in various kinds of inductive learning methods and learning based on observation [4, 5]). The objects which will be used in this process should be in some way significant; we assume that this will happen if they are connected by the relationships. The frequent, multiple presence of such objects should strengthen their ability to create concepts according to the simple rule: *this phenomenon occurs many times, so it must be important*. We will now be more precise about this.

Let us assume that we can select some primitive objects (from $\mathbf{X}_0$) and recognise them as particular instances corresponding to certain concepts (from $\mathbf{C}_0$). Additionally, we assume that it is possible to check all prior defined relationships between all the objects. Let us consider one such relationship, which will be indicated by:

r_i, where: $r_i \in \mathbf{D}_0$, $i \in \mathbf{I}$, $\mathbf{I} = \{1, 2, ..., u\}$ is a set of indices of relations.

We can assume that r_i has n arguments; thus, it may be satisfied by some n-tuple (sequence of n elements) indicated by:

t, $t \in X_0^n$ where: X_0^n – n-th Cartesian power of set $\mathbf{X}_0$.

Of course, this particular relationship can be satisfied by many other tuples, which may be denoted by:

t_{ik}, where: $i \in \mathbf{I}$, $k \in \mathbf{K}$, $\mathbf{K} = \{1, 2, ..., m\}$ is a set of indices of tuples which satisfy r_i relation (the first index of tuple *tik,* points to r_i relation).

On the basis of selected tuples, we propose the construction of a new concept definition. Let us try to construct a group (set) of such tuples. To each tuple, we will attach some

information that identifies the relationship which it satisfies. Let us create the pair (t_{ik}, i), which contains the selected tuple and the index i that points to the relationship.

We will define the proposed group (denoted by **G**) as an ordered set of pairs:

$$\mathbf{G} = \{(t_{ik}, i) : i \in \mathbf{I}, k \in \mathbf{K}\} \tag{1}$$

G corresponds to one particular occurrence of some collection of objects which are connected by relationships. A set of all possible groups **G** which may be created on the basis of the given scene will be denoted by $\mathbf{G}^*$, i.e. $\mathbf{G} \in \mathbf{G}^*$. Let us transform the group **G** by simply replacing each object in each tuple with the label of the object type. As a result, we obtain a set:

$$\mathbf{S} = \{(vtype_i, i) : i \in \mathbf{I}\} \tag{2}$$

where: $vtype_i$ – vector of labels of argument types of the ith relation, $vtypei \in \mathbf{T}^{n_i}$ (n_i th Cartesian power of **T**), **T** – set of labels, n_i – number of arguments of ith relation.

S describes an abstract (with regard to types rather than specific objects) arrangement of relations in the group. The groups that have identical or similar (according to some distance function) arrangements will be regarded as similar. A number of such groups may appear in the sequence of the input phenomena; therefore, let us consider a set of similar groups:

$$\hat{\mathbf{G}} = \{\mathbf{G}_j \in \mathbf{G}^*, j = 1, 2, ...\} \tag{3}$$

Based on set $\hat{\mathbf{G}}$, we may determine one of the most characteristic groups denoted by $\mathbf{G}_p$. This group may be transformed to some $\mathbf{S}_p$ set which represents its arrangement of relations (see above). The $\mathbf{S}_p$ contains information about the relationships and their types of arguments; therefore, it may be called a 'pattern'. We can denote a subtask of the $\mathbf{S}_p$ pattern creation with **FG**.

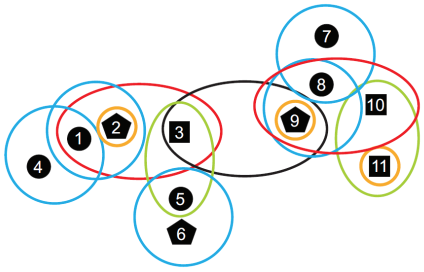


Fig. 1. Simple example of the group structure

Let us present a simple example of $\mathbf{S}_p$ pattern [13]. Figure 1 shows several indexed objects: $x_1, x_2, \ldots$ of three types (1 – circle, 2 – pentagon, 3 – square). A few relationships are also defined (r_1 – indicated by a red oval, r_2 – blue, r_3 – orange, r_4 – green, and r_5 – black oval). We can specify several tuples:

$$t_{31} = (x_2), t_{32} = (x_9), t_{33} = (x_{11}) \text{ (the } t_{31} \text{ denotes 1}^{st} \text{ tuple of relation } r_3)$$

$$t_{21} = (x_1, x_4), t_{22} = (x_1, x_2), t_{23} = (x_8, x_9), t_{24} = (x_8, x_7)$$

$$t_{41} = (x_3, x_5), t_{42} = (x_{10}, x_{11}),$$

$$t_{11} = (x_1, x_2, x_3), t_{12} = (x_8, x_9, x_{10}).$$

We can then point to two groups that are strongly connected by the relations:

$$\mathbf{G_1} = \{(t_{31},3), (t_{21},2), (t_{22},2), (t_{11},1)\}, \mathbf{G_2} = \{(t_{32},3), (t_{24},2), (t_{23},2), (t_{12},1)\},$$

then: $\hat{\mathbf{G}} = \{\mathbf{G_1}, \mathbf{G_2}\}$ and as $\mathbf{G}_p$ we may chose $\mathbf{G_1}$, thus:

$$\mathbf{S}_p = \{((2), 3), ((1, 1), 2), ((1, 2), 2), ((1, 2, 3), 1)\}.$$

Returning to the main problem, let us assume that pattern $\mathbf{S}_p$ has been created. We will try to use it in the definition of a new concept. Nonetheless, such a definition must be completed with some additional elements. We may create a new concept (abstract knowledge); however, only the creation of more concrete items – objects belonging to a certain concept – make it possible to interpret a given phenomenon. The new object may be defined as a combination of several sub-objects (according to pattern $\mathbf{S}_p$). The question arises of whether such a new object may be applied as an argument of old relationships, or must the old relationships be redefined before that. Generally, this is dependent upon how long the new objects should inherit the behaviour of their parents.

Let us consider some relationships that assign certain values to objects. These values may be regarded as object attributes (properties) [9]. We encounter a similar problem: should the new objects have the new kinds of attributes? If so, we ought to provide methods to calculate them. For example, we may use combinations of some standard transformations, e.g. copying, calculation of a sum or an average of parent objects attributes. Finishing these remarks, let us denote processes of defining the new relationships and object attributes using **FR** and **FA**, respectively. Finally, we can define the concept as a composition of elements:

$$(\mathbf{S}_p, \mathbf{FR}, \mathbf{FA}) \tag{4}$$

The new objects, which have been created on the basis of the new concepts, may be utilised for building further concepts. This process may be carried out cyclically. This results in creating a 'hierarchical structure of concepts' and a 'hierarchical structure of objects'. The concept structure includes new general knowledge and can be treated as a model of reality. Thanks to its hierarchical form, such knowledge may be easily interpreted, verified and adapted to other systems. Using the concept structure as a set of patterns, and employing the set of primitive objects which describe the unknown phenomenon, we can produce new object structure – this structure may (or may not) contain objects of particular types. We can postulate that the presence of these objects can be considered as a kind of system response – it may be directly employed in tasks of objects classification, understanding, control, etc.

We should notice, however, two crucial matters concerning the whole approach. The first is associated with the problem of knowledge evaluation. Let us present here only general ideas:

Evaluation by the directly defined performance function.

The hierarchical knowledge in the form of the concept structure may be relatively easily interpreted. In many cases, this allows the identification of concepts that correspond to crucial states of the system. The presence of such concepts in the knowledge structure might be expressed by a particular mathematical formula. This statement may be the basis for the definition of the performance function.

Evaluation with the help of an arbiter, teacher (supervised learning).

The evaluation of the concept structure may be performed by examining the system responses to certain sets of objects. The objects and the proper system replays (determined by the teacher) create pairs which may be treated as elements of the learning sequence [8].

The second crucial problem is the way of finding a suitable knowledge structure; this is dealt with in the next section.

3. OTO multistage process of knowledge building

The process of knowledge construction starts from establishing and defining primitive objects, concepts and relationships ($\mathbf{X}_0$,$\mathbf{C}_0$, and $\mathbf{D}_0$). This initial phase may be performed in several ways. However, the main part of the concept creation involves multiple executions of the **FG**, **FR**, and **FA** tasks – there is a huge number of possible variants of these. Consequently, finding a suitable structure should be treated as the optimisation of a criterion function in an extreme large, multi-dimensional search space. Trying to solve the problem, we will make attempts to reduce the search space size. First, let us again consider the stages of the method; these can be briefly enumerated as follows:

1. calculation of all possible relationships between all objects from **X** set.
2. determining $\hat{\mathbf{G}}$ set.
3. deriving $\mathbf{S}_p$.
4. establishing new relations and new definitions of object feature.
5. creating new objects.

Let us group these issues into three major stages:

A. The observation (corresponding to 1st–3rd items)

In this phase, we identify the relationships between the elements (objects) of the reality. This process provides the answers to the following questions:

What is visible at present? (i.e. in this step).

What properties do the visible objects have?

What kinds of relationships exist between the objects?

We only take into account the observed relationships; therefore, we 'keep a connection' with the observed reality. This leads to decreasing the size of the search space (hypothesis space).

B. The transformation of knowledge (corresponding to the 4th item)

The aim of this process is to change knowledge in order to better describe the observed phenomena in the future (i.e. in the next steps of the process). This means that the transformed knowledge may take into consideration such relationships which are not directly observed. Thus, the main kinds of these transformations should lead to knowledge generalisation – this increases the size of the search space.

C. The operation (corresponding to the 5th item)

The essence of this process is upgrading the current representation of the reality. The representation is conveyed by the structure of the objects. The operation process may lead to the creation of the new objects; therefore, it may significantly increase the size of the search space. The reverse process is also possible – elimination of the object may cause a decrease of the search space size.

The presented scheme may be considered as a kind of framework. Let it be denoted by OTO (Observation-Transformation-Operation).

Let us try to estimate the order of the number of all solutions generated by OTO. We assume for simplicity, that we have m subtasks on each pass of the OTO loop. In each i th subtask, we may choose one of p_i possible ways (the edges of the decision tree). Therefore, the number of possible solutions might be approximated by:

$$h_{oto} = \left(\prod_{i=1}^{m} p_i \right)^2 \tag{5}$$

where: n – the number of passes of the main OTO loop.

The expression (5) is an imprecise estimation, m and pi usually vary with passes of the loop. Let us assume that the size of the knowledge structure may be expressed by the number of concepts which have been included in it. For simplicity, we do not take into account the number of relationships between the concept and the number of arguments in particular relationships.

Additionally, let us presume that the number of concepts produced in each OTO loop is limited to a certain, small value (normally less than 5). Therefore, the size of the structure is roughly proportional to the number of passes of the OTO loop (denoted by n). Formula (5) clearly shows a radical, exponential growth in the number of possible solutions with an increasing knowledge size (proportional to n). We may also interpret that relationship as a growth of the computational complexity with the problem size (expressed by the number of concepts); thus, computational complexity is $2^{O(n)}$. This is a key problem of the whole presented approach. In practical, even in very simple implementations this number exceeds millions.

The OTO process is executed in distinct phases. This allows the beneficial use of some specific problem-solving techniques [8]. We will depict only a general view of selected techniques (the authors' paper [12] provides more details).

The bundle search methods

At each stage of concept creation, we can calculate the value of some local criterion functions (objective function) which assess the ability of the constructed structure to create useful knowledge. It especially refers to the task of the proper choice of set $\hat{\mathbf{G}}$. The value of the criterion function may depend on the number of groups that belong to set $\hat{\mathbf{G}}$ (a typical approach is 'the more the merrier'). The criterion function might be used to rank all possible solutions. In further phases of the searching process, we may only take into account a limited number of potential best solutions (see [8]).

Monte Carlo methods.

Some kinds of Monte Carlo methods that combine random searching and bundle methods can be used. First, we rank the possible edges of the decision tree according to the local criterion function. Then, one edge is selected randomly from the top of the ranking list.

In order to decrease the computational complexity of the search process, we should focus especially on the opportunity to generate the structure at a minimal size. To achieve this, let us consider the possible variants of subtasks of the OTO in more detail. This study will follow according to the main stages – observation, transformation, and operation.

3.1. Variants of the observation process

In our earlier proposal, the observation refers to calculating all relationships between objects from reality. However, the question is what we want to consider as the observed reality. In a broader sense, it may be everything that brings in information, not only the particular objects (instance) of a concept but also the concepts as general beings. We may doubt that the concepts, which are abstract knowledge, may be a source of information about a real phenomenon. Let us remember, however, they have been produced according to some real objects.

The observation of the concepts may involve finding the similarity between them. This makes it possible to build a new concept after the identification of an analogy in the relation patterns of concepts created earlier. To sum up, with reference to the subject of the observation phase, we may list three, very general variants of the observation method.

Obser.1. (this abbreviation refers to a criterion of this simple classification; this notation will be continued in the next points).

1. Set of objects.
2. Set of concepts.
3. Set of relationships.

The next criterion refers to some constraints of the observation processes.

Obser.2.

1. All the objects are observed at a particular stage.
2. Only objects that are significant in some way are observed. In this way, the observation process contains a kind of filtering.
3. One-off object usage.
   Let us presume that an object denoted by $x2$ has been created based on object x_1. Let us then assume that we want to create another object x_3 from two objects $x1$ and x_2. This case is an example of many possible kinds of multiple inheritance (object x_3 contains redundant information). The logical solution is to ignore all the relationships that connect $x1$ and x_2 at the observation stage. This variant of the observation leads to a significant decrease in the size of the search space.
4. Central object method with binary relationships.
   Let us simplify the creation of the **G** sets. The first assumption is that a specific object (called the 'central object') exists, which is an argument of all relations in **G**. Secondly, let us only take binary relationships into account. As a result, the complexity of the observation process grows with the number of objects like: $O(n^2)$ (for each potential central object from **X** we must analyse relationships to all other objects from **X**). The presented approach is efficient, but the open problem is the possibility of substitution of n-ary relationships by combinations of the binary ones.

3.2. Variants of the transformation process

The overview of the variant of the $\mathbf{S}_p$ set transformation will start from modifications of relationships in the pattern. Let us point to main ideas.

Tran.1.

1. Replacing one, or more, indices that identify the relationships in the $\mathbf{S}_p$, with the other indices, see expression (2). We assume, that the new relationship has identical argument types as the previous relationship.
2. Creation of a new relationship through modification of the old relationship.
   The modification may consist in changing the relation parameters (for example a relation *to_be_big* obviously depends on the given threshold). It may also refer to combining the relationships with the other ones. In this way, the set of objects that satisfy the relation can be widened or narrowed (the disjunction and conjunction of the relations).
   We will now consider the other kinds of transformations that modify the number of satisfying objects. Let us take into account binary relationship r_0. We can imagine the new unary relation r_1 that is satisfied by the particular object $\boldsymbol{x}$, if the relationship r_0 is held between the $\boldsymbol{x}$ and only one other object from set **X**. Similarly, we may define the new relationship r_2 which will be satisfied by x if r_0 is satisfied by x and all the objects from **X**. In the same exact way, we may define r_3 if r_0 is satisfied by x and none of the objects from **X**. This kind of modification will be called an 'excluding transformation'. An obvious, but important group of relationship

modifications refers to changing their type of argument. Such transformations are very easy to achieve: suffice is to change one data (index of argument type) in set $\mathbf{S}_p$. This modification may have far-reaching consequences. Let us presume that set $\mathbf{S}_p$ defines a concept indexed by *j*. In $\mathbf{S}_p$, we are changing the type of index of a particular argument of a chosen relation into *k*. Therefore, the concept indexed by *k* is now used to define the concept indexed by *j*. Thus, we may consider this process as the operation of joining one item of knowledge to another.

Let us consider one important, special case of changing the argument type. Let us presume that, as above, the concept indexed by *j* is defined according to some pattern $\mathbf{S}_p$. Technically speaking, it is possible to change some index of the argument type of a selected relation in $\mathbf{S}_p$ also into *j*. In this way, we obtain a 'recursive' type which is defined by itself. The recursive definition cannot 'go' to infinity; there is a necessity to use an alternative definition of the considered concept (see the following points). The modification of the argument type, including creating the recursive type may be considered as a kind of structural transformation. To sum up, let us list some important variants of the relationship transformation.

a) the change of the relation parameters
b) the modifications that the number of objects that satisfy the input relation
c) the simple change of the type of argument to another type
d) the creation of the recursive type (special case of the previous point)
e) the enlargement of the set of possible argument types of certain relationship (in an extreme case the use of objects of any type is allowed)
f) the narrowness of the set of the possible argument type.

To finish the topic of relation transformation, let us remember the **FA** process, leading to the creation of new object features (also regarded as relations) – these may be calculated in different ways. Two main strategies are:

Tran.2.

1. Inheritance of the method of calculating object features from parent concepts.
2. Establishing the new method of the computation of features.

At this time, let us discuss a few basic transformations that change the structure of $\mathbf{S}_p$ sets (structure of concepts). Let us assume that we have created a copy of one particular pattern $\mathbf{S}_p$. Afterwards, this copy is transformed using some of the described methods. In this way, two different $\mathbf{S}_p$ patterns create alternative concept definitions. This leads to knowledge increasing, and may be regarded as an extension of the concept definition (4).

Let us reflect on a reverse operation – the deletion of some concepts. A simple variant of the operation involves removing a given concept and all the concepts which have been created from it. Simultaneously, we have to delete all objects (instances) of cancelled concepts. This reduction has a significant impact on the observation process. The decrease of the number of concepts and objects leads to a radical reduction in the number of generated tuples, as a consequence, this limits the computational time of the whole process. We can propose several ways of selecting the concept to be removed, e.g. deleting such concepts which were

not used to produce the new concepts over a few runs of the OTO loop. To summarise this subsection, let us itemise the major variants of the described transformations.

Tran.3.

1. Permission (or not) for the following structural operations:
   a) the definition of alternative concepts,
   b) concept joining,
   c) concept deleting.

3.3. Variants of the operation process

The operation phase may concern two main processes – the creation and the removal of objects. The first may occur in the following main variants:

Oper.1.

1. Creation of all possible objects.
2. Creation of a reduced set of only the most dissimilar objects.

Other variants relate to the problem of recursive object creation (see above – creation of the recursive type).

Oper.2.

1. Permission (or not) for the creation of recursive objects. In order to avoid the creation of the infinity object structure, we should define specific stop criteria. For example, the object creating process may stop when:
   a) the currently created new object includes itself (among others) as an ancestor object
   b) the currently created new object is similar to the other objects which have been created in the previous passes of the loop (according to the distance function in the space of object attributes).

The second elementary operation refers to object removing.

Oper. 3.

1. Removal of the 'old' objects (which have been created in the earlier runs of the OTO loop).
2. Removal of the unused objects (which are not used in the creation of other objects over a given number of iterations of the OTO loop).

The presented description of the method variants does not pretend to be a strict classification – we have taken into consideration only the most important criteria.

4. A brief example of the application of the OTO scheme

As previously noted, the possibility to create an object of a certain type may be regarded as a system output. The newly constructed, potentially complex object is instantly, automatically classified (just as an object of the given type).

Paper [14] provides an example of the usage of the proposed approach for image recognition tasks. The generating system is able to identify specific spatial relationships between visible objects. A more complex classification is described in [15]. First, the system creates the abstract concept of an equinumerosity of sets. The system should build this concept with minimal prior assumptions (without using its well-known definition that utilises a bijection transformation). Let us briefly portray this problem.

Primitive objects will be considered; these are vectors having only one feature – *colour* (with two values: 0 and 1 or *blue* and *red*). Additionally, primitive binary relationships are defined: *equality* and *inequation* of the *colour* and unary relations: *having_blue_color* and *having_red_color*. The following variants of the concept-creating process are utilised (according to Section 3): **Obser.1.** 1 (1st variant); **Obser.2.** 2, 3, 4; **Tran.1.** 2b, 2c, 2d; **Tran.2.** 1; **Tran.3.** 1a; **Oper.1.** 2; **Oper.2.** 1a, 1b.

The search program has generated many appropriate solutions, we will depict only one:

typ 20 q1 pat:
typ 21 q1 pat: 13: 20 20
typ 22 q1 pat: 15: 21 21
typ 22 q1 pat: 15: 22 21
typ 24 q4 pat: neg 13: 22 20 15: 22 20 15: 22 21

This definition can be interpreted as follows:

typ 20 q1 pat:
typ 20 – a header of the definition of concept (type) no. 20, regarded as primitive objects, q1 – a performance value of this concept, (not significant), pat: – relation definitions (empty here);

typ 21 q1 pat: 13: 20 20
typ 21 – the definition of concept 21 (in further points, headers will be omitted), 13: 20 20 – relation 13 (*inequation*) held between two objects of type 20; the object of concept 21 is a pair of primitive objects (type 20) which have different colours, such a pair will be called a 'different pair'; the colour of this compound object is always 0 (*blue*) as a result of calculating the average colour of components in an integer domain; the first arguments of all binary relations are 'central objects' (Subsection 3.1),

typ 22 q1 pat: 15: 21 21
15: 21 21 – relation 15 (*equality*) held between two objects of type 21; the object of type 22 contains two objects of type 21 (different pairs); all objects of type 21 have *blue* colour, so relation 15 indicates here the existence of another object of type 21;

typ 22 q1 pat: 15: 22 21
15: 22 21 – relation 15 (*equality*) held between objects of type 22 and 21, this is a transformation of the previous concept, type 22 becomes the 'recursive type'; the object of extended concept 22 may be a set of different pairs (type 21);

typ 24 q4 pat: neg 13: 22 20 15: 22 20 15: 22 21
this is the 'excluding transformation' (denoted by 'neg') of relations:
13: 22 20 15: 22 20 15: 22 21;
object of type 24 may be interpreted as a set of different pairs (type 22) for which:
13:22 20 – another object (type 20) having a different colour does not exist,
15:22 20 – another object (type 20) having the equal colour does not exist,
15:22 21 – another pair does not exist (all pairs have *blue* colour).

The last conditions mean that an object of type 24 includes all of the different pairs, additionally, there exists no other object (type 20) that is not a component of them. As a consequence, the possibility of creating the object of type 24 indicates that two sets have the same cardinality. Figure 2 shows a tree structure of a chosen object of type 24.

The usage of the equinumerosity concept and performing of the operation of knowledge joining (see Subsection 3.2) enables the establishment of a percentage composition of some examined objects that are visible on the image [14]. The analogous usage of the method may relate to website interpretation. We know that the created object corresponds to the **G** group of relations which describe the observed situation (Section 2). We may say that the group contains the semantic content of the situation (phenomenon). We regard the extraction process of the semantic content of the given situation as a condition of its understanding (Section 1.1). As a consequence, the described method may be utilised in the automatic understanding of complex sentences or multipart pictures.

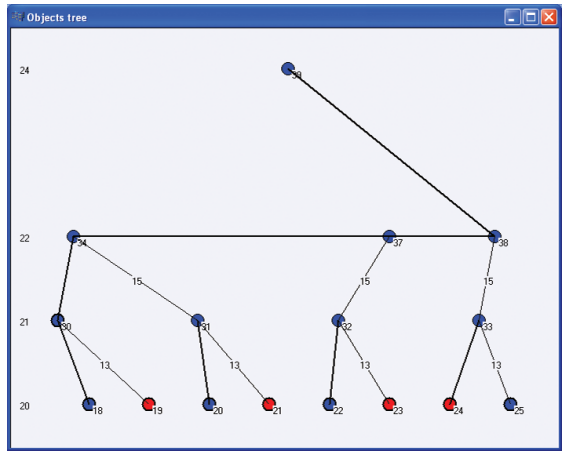


Fig. 2. An example of the tree structure of an object of type no. 24. All the objects on the same level are of an identical type (their numbers are written on the left). Bold line, leading to the particular object, indicate which one of its parent objects is the 'central object'. The other lines (edges) symbolise the binary relationships between the central object (mentioned above) and the other parent objects (the number of the relation is placed in the middle of the line)

5. Conclusions

The paper presents the multi-stage scheme of creating simple ontology. The creation process may start on the basis of minimal assumptions. The phases of the scheme, which correspond to knowledge generalisation and specification, allow convergence of the knowledge structure into the desired form. Consequently, the knowledge system has the ability to self-build and self-update. In this context, the presented method satisfies the requirements specified in Section 1. We can highlight some properties of the presented scheme, for example, the ability to construct knowledge without a model of given phenomena that extends the capability of classical knowledge systems (expert systems) [8].

The main drawback of the presented approach is the lack of efficient methods of searching the optimal knowledge structure. The invention of such methods remains the major challenge. In many parts of the paper, we accent the necessity of building the knowledge structure at the minimal possible size (micro-ontology). The article indicates several ways to achieve the knowledge simplification. The most important of these are:

- an ontology definition which uses relations between objects only;
- usage of binary relationships;
- applying the 'central object' method;
- using the operations of deletion of objects and concepts.

The usage of the OTO model may be very promising in areas where the knowledge has an evidently structural form, especially for the automatic understanding of the websites and the image recognition tasks.

References

[1] Davies J., Studer R., Warren P. (eds.), *Semantic Web Technologies Trends and Research in Ontology-based Systems*, John Wiley & Sons Ltd, 2006.

[2] Gennari J. et al., *The evolution of Protégé, An environment for knowledge-based systems development*, Int. Journal of Human-Computer Interaction, 58, 2003.

[3] Husserl E., *Logical Investigations*, Routledge, New York 2003 (first published in German as Logische Untersuchungen, M. Niemeyer, Hale 1900/1901).

[4] Michalski R.S., Carbonell J.G., Mitchell T.M. eds., *Machine Learning: An Artificial Intelligence Approach*, Springer Science & Business Media 2013.

[5] Muggleton S., *Scientific knowledge discovery using inductive logic programming*, Communica-tions of the ACM, Vol. 42, 1999.

[6] Piekarczyk M., Ogiela M.,R., *Matrix-based hierarchical graph matching in off-line handwritten signatures recognition*, Proc. of 2nd IAPR Asian Conference on Pattern Recognition, IEEE, 2013.

[7] Piekarczyk M., Ogiela M.R., *The Touchless Person Authentication Using Gesture types Emulation of Handwritten Signature Templates*, Proc. of 10th ICBWC, Coommunication and Applications BWCCA, Krakow 2015.

[8] Russell S., Norvig P., *Artificial Intelligence: A Modern Approach*, 3rd edn, Prentice Hall, En-glewood Cliffs 2010.
[9] Smith M.K., Welty C. (eds.), *OWL Web Ontology Language Guide*, http://www.w3.org/TR/owl-guide (access: 10.11.2016).
[10] Tadeusiewicz R., Ogiela M.,R., *Medical Image Understanding Technology*, Studies in Fuzziness and Soft Computing, Vol. 156, Springer-Verlag, Heidelberg 2004.
[11] Szczepaniak P.S., Tadeusiewicz R., *The Role of Artificial Intelligence, Knowledge and Wisdom in Automatic Image Understanding*, Journal of Applied Computer Science, Vol. 18, 2010.
[12] Wójcik K., *OTO Model of Building of Structural Knowledge – Areas of Usage and Problems*, Advances in Intelligent Systems and Computing, Image Processing and Communications Chal-lenges 4, Springer-Verlag Berlin, Heidelberg 2012.
[13] Wójcik K., *Indutive learning methods in the simple image understanding system,* ICCVG' 10 Proceedings of 10th International Conference on Computer Vision and Graphics, Part I, LNCS, Springer-Verlag Berlin, Heidelberg 2010.
[14] Wójcik K., *Hierarchical Knowledge Structure Applied to Image Analyzing System – Possibilities of Practical Usage*, ARES'2011 Proceedings of the IFIP WG 8.4/8.9 International Cross Domain Conference Viena, LNCS Springer-Verlag Berlin, Heidelberg 2011.
[15] Wójcik K., *Knowledge Transformations Applied in Image Classification Task*, Advances in Intelligent Systems and Computing, IPCC 5, Springer-Verlag Berlin, Heidelberg 2013.

TECHNICAL TRANSACTIONS 3/2018
CZASOPISMO TECHNICZNE 3/2018
ENVIRONMENTAL ENGINEERING
DOI: 10.4467/2353737XCT.18.042.8304
SUBMISSION OF THE FINAL VERSION: 26/2/2018

Bożena Babiarz (bbabiarz@prz.edu.pl)
Department of Heat Engineering and Air Conditioning, Faculty of Civil and Environmental Engineering and Architecture, Rzeszow University of Technology

Alicja Puszkarewicz
Department of Water Purification and Protection, Faculty of Civil and Environmental Engineering and Architecture, Rzeszow University of Technology

Selected causes of exploitation problems caused by secondary contamination of drinking tap water

Wybrane przyczyny problemów eksploatacyjnych wywołanych wtórnym zanieczyszczeniem wody wodociągowej

Abstract

The paper presents selected causes of problems related to the widely understood phenomenon of secondary contamination of drinking tap water caused by corrosion. The legal aspects, related to the formal responsibility for water quality in water supply systems: cold water and hot water, according to the way it is prepared, have been analysed. The main causes of the deterioration of water quality in water supply systems were characterized by the criteria for assessment of the corrosivity and aggressiveness of the water. Proposals for actions which should be undertaken to prevent secondary water contamination in water supply systems have been drawn up.

Keywords: drinking tap water, secondary contamination of water, aggressiveness

Streszczenie

W pracy zaprezentowano wybrane przyczyny problemów związanych z szeroko pojętym zjawiskiem wystąpienia wtórnego zanieczyszczenia wody wodociągowej, spowodowanego korozją. Przeanalizowano aspekty prawne związane z formalną odpowiedzialnością za jakość wody w instalacjach wodociągowych: wody zimnej i ciepłej wody użytkowej, w zależności od sposobu jej przygotowania. Scharakteryzowano główne przyczyny wystąpienia zjawiska pogorszenia jakości wody w instalacjach wodociągowych podając kryteria oceny korozyjności i agresywności wody. Opracowano propozycje działań, jakie należy podjąć w celu zapobiegania zjawiska wystąpienia wtórnego zanieczyszczenia wody w instalacjach wodociągowych.

Słowa kluczowe: woda wodociągowa, wtórne zanieczyszczenie wody, agresywność

1. Introduction

Modern technologies of water treatment allow water with better and better parameters to be introduced to water supply networks. However, this does not guarantee that consumers will receive water with equally good parameters. The direct effects of secondary contamination of tap water are suffered mainly by its consumers [4] but operational problems are also felt by water supply companies and, in the case of using district heating to prepare hot tap water, by heat-generating plants as well. In normal operational conditions, secondary contamination of water resulting from the supposed contact of tap water with the heating medium (heating water) is excluded due to the water heating in heating substations using a diaphragm. Heated water – hot tap water has no direct contact with heating water. Yet, problems appear connected with the occurrence of "red water" in cold and hot tap water systems. The main cause of this situation may be the technical condition and the configuration of the infrastructure responsible for supplying water to consumers as well as the lack of water treatment process stability [5]. The operating conditions of both the water supply network [3] and domestic distribution systems may also have an influence on water contamination. The highest risk is where water consumption is low. Additionally, the diameter of pipes and, above all, the ratio of the internal size of pipe surfaces to the amount of flowing water affect the content of metal ions and corrosion products. Secondary microbiological contamination of tap water connected with corrosion processes is possible. The presence of corrosion products on internal surfaces of pipes through which drinking water flows contributes to the creation of biofilm [5]. The intensity of the creation of this biofilm and its susceptibility to disinfection processes depend largely on the material of the system.

Tap water that stands for a long time in network and system tanks and pipes contributes to its contamination. Secondary microbiological contamination of water occurs when deposits come off the sanitary system walls. Biofilms created on surfaces of hot tap water domestic system pipes favour the development of *Legionella* bacteria. Disinfection of the network prevents secondary microbiological contamination [1, 15] but intensifies the corrosion process.

Water supply networks and domestic distribution systems are made mainly of materials which are prone to corrosion and deposit accretion. Such phenomena result in water quality deterioration [5] to the highest degree. Science has studied the corrosion of cast iron and steel in water supply networks for more than 80 years and in spite of the fact that this long period has allowed the problem to be described thoroughly, almost all scientists admit that corrosion of materials based on iron (steel, cast iron) is a very complex phenomenon.

Interactions between components of transported water and corrosion products present on pipework walls significantly change the quality of water mainly in a way that is undesirable to its consumers. Reactions taking place in corrosion deposits contribute to water quality deterioration in pipework, above all, as a result of the release of iron compounds, migration of elements from corroding materials and creation of compounds which have an undesirable smell. A straight majority of consumers' complaints to water suppliers is directed in connection with the occurrence of so-called "red" water in hot tap water domestic systems. This phenomenon is caused by the release of Fe(II) ions from cast iron and steel pipes

which are then oxidized with oxygen dissolved in the water or remaining disinfectants, creating coloured floccules $Fe(OH)_3$. The occurrence of "red water" is strictly correlated with the cloudiness of water so the increase of this parameter will indicate the development of corrosion in the network. Most frequently "red water" appears at the end of networks. Legal regulations concerning water quality in domestic distribution systems and water supply networks as well as the operation of water supply and heat-generating networks are contained in relevant legislation.

2. Legal aspects of water supply to consumers

Pursuant to Article 3 (2) of the Crisis Management Act of 26 April 2017 [13] (JL no 89(590) as amended), the water supply system belongs to the critical infrastructure of the state and is ensured by the efficient operation of all its bodies.

Water supply and sewage treatment companies operate under the Act of 7 June 2001 on Collective Supply of Water and Collective Sewage Treatment as amended [14] (JL no 139 of 2015). Pursuant to Article 5(1) of that Act: "A water supply and sewage treatment company is obliged to ensure the ability of water supply and sewage devices possessed by such a company to supply the required amount of water under an adequate pressure and to supply water and treat sewage in a continuous and reliable manner, and to ensure that the supplied water and treated sewage are of proper quality".

The priority for water supply and sewage treatment companies should always be the quality of supplied water. Pursuant to §3(1) of the Regulation of the Minister of Health of 13 November 2015 on the Quality of Water Intended for Human Consumption [11] (JL no 1989 of 2015): "Water is safe for human health if it is free from pathogenic micro-organisms and parasites to the degree which constitutes a potential risk for human health, any substances in concentrations which constitute a potential risk for human health, does not display aggressive corrosion characteristics and meets:

1) basic microbiological requirements described in Annex no 1 to the Regulation;
2) basic chemical requirements described in Annex no 2 to the Regulation".

Water supply to consumers is affected by numerous conditions connected with the organizational and legal structure of water supply whose demonstration is presented in Fig. 1.

The Regulation of the Ministry of Infrastructure of 12 April 2002 on Technical Conditions Buildings and their Location Should Meet [9] also notices the necessity to secure the quality of tap water, i.e. to secure water against secondary contamination. According to § 113 of that Regulation "Products used for a domestic water distribution system should be selected taking into account water corrosivity so that its quality and the durability of the system would not be impaired and such effects would not be brought about by interaction of materials of which such products are made. A domestic water distribution system should be equipped with devices protecting against secondary contamination of water according to requirements concerning reverse flow described in the Polish Standard concerning protection against reverse flow".

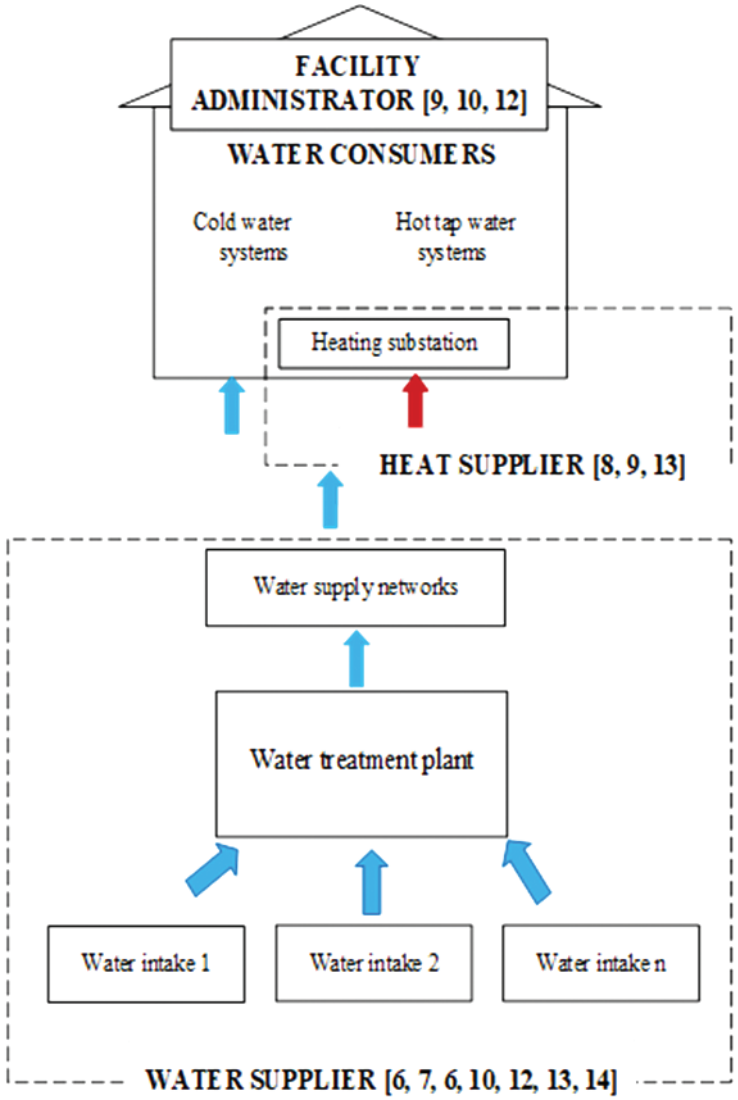


Fig. 1. Demonstration organizational and legal structure of water supply (own elaboration)

According to standard PN-EN 1717:2003 [7] indicated in the Regulation, domestic distribution systems should not cause contamination of public or private drinking water supply systems with deposits, harmful water or undesirable substances, irrespective of their design or construction. A choice of proper protection is possible after a thorough assessment of the risks in the draw-off point requiring protection against contamination caused as a result of reverse flow [4].

Conditions of supplying water intended for human consumption and rules of creating conditions for ensuring proper quality of water are governed by the Act of 7 June 2001 on Collective Supply of Water and Collective Sewage Treatment [14] (Article 5), according to

which: "1. A water supply and sewage treatment company is obliged to ensure the ability of water supply and sewage devices possessed by such company to supply the required amount of water under an adequate pressure and to supply water and treat sewage in a continuous and reliable manner, and to ensure that the supplied water and treated sewage are of proper quality.

1a. A water supply and sewerage treatment company is obliged to conduct regular internal controls of water quality."

Requirements concerning hot tap water are contained in the Regulation of the Ministry of Infrastructure of 12 April 2002 on Technical Conditions Buildings and their Location Should Meet [8]; they are, however, only requirements concerning water temperature, as cited below (§ 120): "A domestic hot water distribution system should enable to obtain water with the temperature not lower than 55°C and not higher than 60°C in draw-off points".

That Regulation regulates requirements concerning the amount of thermal energy needed to prepare such water, i.e. such amount should be maintained at a reasonably low level and devices for preparing hot water installed in buildings should meet requirements concerning energy efficiency.

According to the Regulation of the Minister of Economy of 15 January 2017 on Detailed Conditions for Heat Distribution System Operation [8], describing, among others, the way heat is sold (§ 2), "A heat distribution company – energy company producing heat in sources of heat exploited by such a company and transmitting, distributing and selling heat produced in those sources or purchased from another energy company".

So, heat distribution companies are suppliers of heat in the form of heating medium serving for heating tap water and are responsible for thermal energy parameters. Heating is effected using a diaphragm on devices in which only the process of heat exchange takes place between the heating medium and heated water which in normal conditions of operation excludes the contact of the heating medium with heated tap water. The cold water supplier is responsible for the quality of heated tap water; however, the influence of the condition of the domestic distribution system of the building for which its owner is responsible is not without significance.

The principles of using residential buildings with all their systems and technical devices are governed by the Building Law [12] of 4 July 1994 (JL no 1409 of 2013 as amended) and the Regulation of the Minister of Internal Affairs and Administration of 16 August 1999 on Technical Conditions of Using Residential Buildings [9].

According to Article 5(2) the Building Law [12], „A structure should be used according to its intended use and environmental protection requirements and maintained in proper technical and aesthetic conditions, not allowing that its utility capabilities and technical efficiency be deteriorated, especially with regard to requirements referred to in Point 1 (1–7)" including „adequate hygienic and health conditions and environmental protection".

According to § 3 of the Regulation of the Minister of Internal Affairs and Administration of 16 August 1999 on Technical Conditions of Using Residential Buildings [10] (JL.99.74.836, JL.09.205.1584): " ...domestic hot tap water distribution system – hot water pipework in a building along with its fittings and equipment which begins in the place in which the pipe is connected to the valve cutting off such a system from the heating substation or connection and

ends in hot water draw-off points; such a system also includes a local hot tap water distribution system". Moreover, §13(1) gives the conditions for and the manner of using technical devices and systems as well as products used for repairing and maintaining them which should not deteriorate the utility capabilities of the medium supplied through such devices and systems. According to Point 2, "technical and functional parameters of the medium supplied using the technical devices and systems to flats and rooms intended for common usage should be in line with values of parameters specified in separate regulations and relevant designs of such devices and systems". Chapter 8 § 28 of that Regulation regulates using a hot tap water system according to which: "The hot tap water system should, in the period of using water, enable the possibility to supply it with the temperature defined in separate regulations to draw-off points according to the conditions of using it, as assumed in the design".

3. Corrosive factors

The most important parameters which determine the intensiveness and rate of corrosive processes include [2]:

- **temperature** – this has a decisive influence on the effect of corrosive destruction of metals. On the one hand, it changes the concentration and even the composition of water, and on the other hand, it causes the increase of the degree of dissociation and a considerable rise in the movement of ions. This results in the decrease of the value of pH reaction. It has been determined that with an increase in the temperature of 1°C pH falls by 0.01. Additionally, the temperature rise increases the rate of diffusion of oxygen to the surface of metal; however, its solubility decreases. At a temperature > 80°C, a fall of the corrosion rate is observed. This phenomenon is connected with the occurrence of two processes: the increase of the corrosion rate with the rise in the temperature and at the same time the decrease of the insolubility of oxygen with the fall of the temperature of the system, and this process, after it exceeds a temperature of 80°C, obstructs the influence of the temperature factor. In closed (pressurized) systems from which oxygen cannot be released, the corrosion rate increases with the rise in the temperature up to the total absence of oxygen in water.
- **content of oxygen dissolved in water** – such parameter has a variable influence on releasing Fr(II); and the higher the content in tap water, the larger is the corrosion rate,
- **water pH** – numerous tests show that a higher value of water pH limits the rate of iron release; pH increase affects the rate of Fe(II) oxidization and limits the solubility of $Fe(OH)_3$ and iron carbonate,
- **alkalinity of water** – higher alkalinity limits corrosion but in some cases the occurrence of "red water" was also observed when the alkalinity of water was increased,
- **content of free carbon dioxide in water** – the smaller the amount of CO_2 in water containing oxygen, the lower the rate of releasing iron compounds.
- **salinity of water** – a high concentration of Cl^- and SO_4^{2-} ions increases the corrosion rate and some papers indicate that the increase of the corrosion rate is seen with the rise in the temperature of water; additionally, a higher concentration of nitrates will affect

the increase of releasing iron in chlorine-disinfected water; in oxygenated water, the release of iron is inhibited by insoluble ferric hydroxide oxides, calcium carbonate and siderite which form a protective layer hindering diffusion of Fe(II) ions,

- **stagnation of water in pipework** – this usually causes a fall in the oxygen content and an actual risk of the occurrence of "red water".

Analysing the above, it may be found that corrosion is limited (but not eliminated) by higher alkalinity of water and higher pH. Chlorides and sulphates create favourable conditions for corrosion, the lower the alkalinity of the water. Waterworks stations taking surface water should consider the fall in the alkalinity of the water which takes place in summer. Stagnation of water in pipework leads to releasing relatively large amounts of Fe(II), which, after the following flow, may result in "red water". Reports also show that when pipes have already been corroded, good water oxidization prevents such phenomena.

A proper protection against corrosion should include, above all, the identification of the cause and determination of the corrosion rate. On that basis actions undertaken to prevent and slow down the corrosion should be defined. Maintaining an appropriate pH reaction and the equilibrium of carbon and calcium as well as the adjustment of alkalinity of water prevent excessive corrosion of the network and protect against the creation of so called rust-coloured water. No water component remains unreactive to metal and each may speed up or slow down corrosion. The rate of corrosion is determined by interaction of protective and corrosive components contained in water. Water corrosivity in most cases is identified with its aggressiveness, which is not correct. Aggressive waters, indeed, are always corrosive but nonaggressive waters may also be corrosive.

4. Methods of assessing the aggressiveness of corrosive water

There are many indexes (Langeliera, Ryznara indexes or Strohecker formulas) used to assess the corrosivity of water but due to the complexity of the corrosion process, none of them ensures the full assessment of the rate of corrosion taking place in water environment [2]. Due to the common usage of galvanized steel pipes in hot and cold water systems, the values of the above-mentioned indexes were regarded as a good basis for forecasting the durability of system materials. As the range of materials used in domestic distribution systems has been significantly extended, it has become necessary to create the proper material selection system based on the knowledge of the degree of corrosive aggressiveness of tap water which will be supplied to a given system. Standard PN-EN 12502 [6]; Protection of metal materials against corrosion, is helpful in this. Guidelines to the assessment of the risk of corrosion in water distribution and storage systems. According to such document: "corrosion is a result of interaction between various factors, its volume may be only expressed as a risk of the occurrence of corrosive damage". The document is, thus, a guide and it does not contain any principles for using various materials in water supply networks. It may only serve to decrease the risk of the occurrence of corrosive damage by:

- helping to design, install and operate systems with regard to anti-corrosion protection,
- assessing the need of additional protection against corrosion for new and already operating systems,

- helping with analysing the causes of failure to prevent its recurrence.

The risk of corrosion caused by the impact of water is determined by indicators S_1, S_2. Indicator S_1 determines the risk of pitting corrosion [5];

$$S_1 = \frac{c(Cl^-) + c(NO_3^-) + 2c(SO_4^{2-})}{c(HCO_3^-)} \quad (1)$$

when:

$c(Cl^-)$ – concentrations of ions Cl^-,
$c(NO_2^-)$ – concentrations of ions NO_2^-,
$c(SO_4^-)$ – concentrations of ions SO_4^{2-},
$c(HCO_3^-)$ – concentrations of ions HCO_3^-.

Concentrations of specific ions are expressed in [mmol/dm^3]. The risk of pitting corrosion increases with an increase in the value of S_1:

If $S_1 < 0.5$ – the occurrence of pitting corrosion is unlikely and very likely for $S_1 > 3$. Bicarbonate and phosphate anions combined with calcium cations may behave like cathodic inhibitors. Concentrations necessary for that purpose are:

$$c(HCO_3^-) \geq 2.0 \frac{mmol}{1}, cCa^{2+} \geq 0.5\ mmol/1$$

Stagnation of water and sudden changes in temperature contribute to the occurrence of pitting corrosion. If water in cold water pipes after overnight stagnation in the room temperature is replaced with water with lower temperature, corrosion products tend to come off.

The value of parameter S_2 shows the risk of selective corrosion:

$$S_2 = \frac{c(Cl^-) + 2c(SO_4^{2-})}{c(NO_3^-)} \quad (2)$$

when:

$c(Cl^-)$ – concentrations of ions Cl^-,
$c(NO_3^-)$ – concentrations of ions NO_2^-,
$c(SO_4^{2-})$ – concentrations of ions SO_4^{2-}.

It is thought that the susceptibility to selective corrosion is low when the value of S_2 is below 1 or above 3 or when $cNO_3^- < 0.3$ mmo/l.

The conditions of system operation have no significant influence on the selective corrosion process but the effects of selective corrosion significantly affect water quality. The amount of dispersed solid particles of corrosion products which may be removed from pipe walls depends on the time of slow water flow and the volume of sudden turbulent flow. With the progress of the corrosion process, the colour of corrosion products changes from grey to brown.

The above indicators do not consider many parameters determining the corrosive nature of water; in practice, they are used for a "rough" assessment of the chemical stability of the water. Unstable water causes the accretion of calcium carbonate deposits on pipes and contributes to the occurrence of deposit corrosion, whereas corrosive water destroys pipes and contaminates them with corrosion products.

5. Summary

The main causes of operational problems arising from secondary contamination of tap water due to its corrosiveness are connected by factors including the instability of the quality of water supplied to consumers, the technical condition, configuration and plumbing conditions of the network (pressure changes), heating substations and domestic distribution systems. The lack of chemical stability of tap water may be caused by many factors, mainly the deteriorated quality of raw water (which results for example from low water level, long periods of high temperatures or lack of precipitation), the application of large doses of coagulant and lack of pH correction as well as possible disruptions in the automatic administration of reagents, etc. Stagnation of water and changes in temperature, the flow rate and pressure intensify the phenomenon of corrosion. At the same time, the occurrence of the above-mentioned circumstances combined with instability of the chemical constitution of water may cause corrosion and corrosion products coming off pipe walls which result in secondary contamination. That is visible in hot tap water system mainly in places in which the flow rate falls (heating) and water is accumulated (in tanks) as well as in final points of the cold water system, thus causing operational problems connected with the occurrence of "red water".

This problem could be solved if water treatment plants performed obligatory analyses of water quality with regard to its corrosive aggressiveness. A solution to this problem may be achieved by introducing additional legal requirements which would explicitly determine criteria and limits of tap water corrosivity to avoid consequences connected with water reaction to elements of domestic distribution systems and water supply networks. This paper suggests the possibility of using indicators S_1, S_2 and Langelier and Ryznar indexes in the assessment of the corrosive aggressiveness of water. That determines the quick identification of causes of water quality deterioration, enabling to avoid possible operational problems.

Technical actions preventing the occurrence of secondary contamination of water in domestic distribution systems include: the necessity to fix fittings protecting against contamination supplied from the water supply network on the cold water inlet to heating substations (mechanical filters) as well as fittings ensuring the possibility to rinse by-passes in heating substations, conducting regular assessments of domestic distribution systems and keeping them in a proper technical condition, maintaining a constant circulation in hot tap water systems in buildings and stabilization of pressure in the water supply network. It is essential to optimize the selection of materials of pipes, fittings and devices in the network, heating substations and domestic distribution systems depending on water composition, and to successively apply high-quality materials and fittings which are resistant to corrosion. Due to the undeniable influence of failures in domestic distribution systems and water supply networks, attention should be paid to the quality of work performed and the following of the regime during resolving the failure in the water supply network. Additionally, there is significance in the strict cooperation between water supply and heat distribution companies' services, domestic distribution system users and building owners with regard to the prevention of water quality deterioration in domestic distribution systems and avoidance of problems connected with such deterioration.

References

[1] Babiarz B., Szymański W., *Możliwości dezynfekcji termicznej instalacji ciepłej wody z zastosowaniem kolektorów słonecznych*, Ośrodek Informacji „Technika instalacyjna w budownictwie", Instal, Vol. 3, 2007, 36–38.

[2] Kowal A., Świderska-Bróż M., *Oczyszczanie wody*, PWN, Warszawa2007.

[3] Niedziołek M., Dąbrowski W., Żaba T., McGarity M., Głód K., *Charakterystyka awarii rurociągów tranzytowych na przykładzie przewodów biegnących z ujęcia w Dobczycach do Krakowa*, Czasopismo Techniczne, 1-Ś/2011.

[4] Opaliński S., Rybak-Wilusz E., Babiarz B., *Zabezpieczenie wody w instalacjach wodociągowych przed wtórnym zanieczyszczeniem wynikłym z niekontrolowanych wstecznych przepływów*, Ośrodek Informacji „Technika instalacyjna w budownictwie", Instal, Vol. 5, 2002, 7–12.

[5] Pierścieniak M., Trzcińska N., Słomczyński T., Wąsowski J., *Problem wtórnego zanieczyszczenia wody wodociągowej*, Ochrona Środowiska i Zasobów Naturalnych, Vol. 39, 2009.

[6] PN-EN 12502; Ochrona materiałów metalowych przed korozją. Wytyczne do oceny ryzyka wystąpienia korozji w systemach rozprowadzania i magazynowania wody.

[7] PN-EN 1717: 2003 Ochrona przed wtórnym zanieczyszczeniem wody w instalacjach wodociągowych i ogólne wymagania dotyczące urządzeń zapobiegających zanieczyszczaniu przez przepływ zwrotny.

[8] Rozporządzenie Ministra Gospodarki z dnia 15 stycznia 2007 r. w sprawie szczegółowych warunków funkcjonowania systemów ciepłowniczych (Dz.U. 2007 nr 16 poz. 92).

[9] Rozporządzenie Ministra Infrastruktury z dnia 12 kwietnia 2002 r. w sprawie warunków technicznych, jakim powinny odpowiadać budynki i ich usytuowanie (Dz. U. z 2002 r. Nr 75, poz. 690) z późn. zm.

[10] Rozporządzenie Ministra Spraw Wewnętrznych i Administracji z dnia 16 sierpnia 1999 r. w sprawie warunków technicznych użytkowania budynków mieszkalnych. (Dz.U.99.74.836, Dz.U.09.205.1584).

[11] Rozporządzenie Ministra Zdrowia z 13 listopada 2015 r. w sprawie wymagań dotyczących jakości wody przeznaczonej do spożycia przez ludzi (Dz.U. z dnia 27 listopada 2015 r. poz.1989).

[12] Ustawa Prawo Budowlane z dnia 4 lipca 1994 r. (Dz.U. z 2013r poz.1409 z późniejszymi zmianami.

[13] Ustawa z dnia 26 kwietnia 2007 r. o zarządzaniu kryzysowym (Dz. U. Nr 89, poz. 590 z późn. zm.).

[14] Ustawa z dnia 7 czerwca 2001 r. o zbiorowym zaopatrzeniu w wodę i zbiorowym odprowadzaniu ścieków wraz z późniejszymi zmianami (Dz.U. z 2015 poz. 139).

[15] Vranayová Z., Košičanová D., Očipová D., *Legionella prevention plan in water systems and its energetic consequences*, Czasopismo Techniczne, 1-B/2009.

TECHNICAL TRANSACTIONS 3/2018
CZASOPISMO TECHNICZNE 3/2018
ENVIRONMENTAL ENGINEERING
DOI: 10.4467/2353737XCT.18.043.8305
SUBMISSION OF THE FINAL VERSION: 5/3/2018

Jacek Dawidowicz (j.dawidowicz@pb.edu.pl)
Białystok University of Technology, Faculty of Civil and Environmental Engineering, Department of Environmental Technology and Systems

Andrzej Czapczuk (andrzej.czapczuk@fbitasbud.pl)
F.B.I. TASBUD S.A., Science and Research Division Warsaw

The reliability of rural water distribution systems in relation to the layout of the pipework within the network

Niezawodność wiejskich systemów dystrybucji wody w zależności od układu przewodów sieci

Abstract

In this paper, the *K* readiness index was adopted in order to evaluate the reliability of the rural water distribution system; complete systems were assessed using the partial survey method. Hydraulic calculations were performed, followed by reliability evaluation with our own NSW programme. In addition, the lengths of the ring and branched sections of the analysed water distribution systems were calculated and 256 computational variants were obtained. Taking into account the values of the *K* readiness index and the cable lengths, the influence of the pipework system on the level of reliability of a water distribution system with a pumping-hydrophore structure was determined. In the case of a branched system, a moderately strong negative correlation was identified between reliability and the length of pipework.

Keywords: rural water distribution system, reliability, stationary index of readiness *K*

Streszczenie

W niniejszej pracy do oceny niezawodności wiejskiego systemu dystrybucji wody przyjęto wskaźnik gotowości *K*, a ocenę całych systemów prowadzono metodą przeglądu częściowego. Wykonano obliczenia hydrauliczne, a następnie niezawodnościowe z pomocą autorskiego programu NSW. Wyliczono długości przewodów części pierścieniowej i rozgałęzionej analizowanych systemów dystrybucji wody. Uzyskano 256 wariantów obliczeniowych. Biorąc pod uwagę uzyskane wartości wskaźnika gotowości *K* oraz długości przewodów, starano się ustalić wpływ układu przewodów na poziom niezawodności systemu dystrybucji wody o układzie pompowo-hydroforowym. W przypadku przewodów rozgałęzionych uzyskano średnio silną ujemną korelację między wskaźnikiem niezawodności *K* a długością rurociągów.

Słowa kluczowe: wiejski system dystrybucji wody, niezawodność, stacjonarny wskaźnik gotowości *K*

1. Introduction

The water distribution system is one of the most important and expensive elements of the water supply system and its design is a decisive factor in the reliability of delivering water to consumers. The primary task of a water distribution system is to provide an uninterrupted supply of water at the right level of pressure and for the minimum cost of system construction and operation. Designing a water distribution system for a given terrain topography and within given spatial planning restrictions requires various technical solutions to be considered. The decision to build a water supply system is characterised by the following features: it is irreversible; it is non-routine; it requires huge financial investment; significantly, it changes the living conditions of the given community. These features make the design stage extremely important. The longevity of the water distribution system for many years of reliable water supply in the future is a decisive factor.

In practice, design variants are predominantly analysed from technical and economic standpoints, but reliability aspects are also increasingly taken into consideration nowadays.

2. Characteristics of rural water distribution systems

Rural water distribution systems are characterised by many features that distinguish them from urban systems, hence the need for them to be analysed separately from technical and reliability perspectives. These features are as follows [1, 2]:

- high fluctuations in hourly water consumption in the countryside,
- low water pressure required by consumers in rural settlements compared to the pressure required in cities,
- an extensive water supply network in rural areas which differs from the grid system for more compact urban developments,
- there is no water intake from many water pipes,
- the predominant share of the branched structure of the water supply network,
- no mains water pipes characteristic of the urban water distribution systems in the structure of the rural water network – transit pipework is more common, as is the case with group water supply systems.

The main components of the water distribution systems are the pipework and fittings, such as valves, reducers and air vents. The route of the pipework is mostly adapted to the rural traffic system. In rural conditions, water pipes rarely exceed diameters of 200 mm. In terms of maintaining pressure, the most common solution in Poland is the pump-hydrophore system.

3. Review of the research on the reliability of water distribution systems

Research on the reliability of water distribution systems has been ongoing for many years [3–6]. The theoretical basis for the application of the analysis of reliability in water supply issues is discussed in, inter alia [7–10]. A comparative analysis of reliability indicators for

rural and municipal water supply systems is presented in articles [11–13]. The reliability of water supply and sewage infrastructure in Poland is discussed in study [14]. The subject of the reliability of rural water supply systems was discussed in works [15–17].

A proposal for the method of determining the index of readiness for simple water supply subsystems was discussed in study [18]. Article [19] presents the methodology for determination of readiness index K using a multi-stage decomposition process. Study [20] describes a procedure for analysing the reliability of the water distribution systems during interruptions to supply resulting from the disconnection of water pipe sections for reasons of maintenance or repair, taking into account the probability of failure. In study [21], the theory of entropy was proposed for the evaluation of the reliability of the system. Entropy was used as a substitute measure of reliability and described the degree of fragmentation of the water flow between the source and the nodes in the water distribution system.

In paper [22], a methodology was presented to estimate the hydraulic reliability of the water distribution system; this can be defined as the probability that the system can provide the required flow rate at the required pressure. Because of the random nature of future water demand, along with the required pressure and the condition of the pipework, the reliability of future distribution systems is uncertain. Using the 'Monte Carlo' method, a simulation model with three basic components was developed, these components are the generation of random numbers, hydraulic network simulations and reliability calculations. Another simulation model for calculations of the reliability of the network and pumping stations was proposed in paper [23]. In study [24], a reliability assessment model was discussed; this combined the selection of pipe diameters and tank heights. An assessment of the reliability of water distribution systems along with an analysis of network gateway locations is described in articles [25, 26].

Hydraulic calculation methods are supplemented with various optimisation methods and reliability assessments [27–34].

A description of the management of the water distribution system with regard to its reliability is provided in works [35, 36]. Article [37] describes the situation encountered in many under-developed countries where the water supply system is unable to supply water to all consumers simultaneously. In such a situation, the reliability of the system is improved by optimising the water supply through the proper management of the operation and the introduction of maintenance schedules. Aspects of reliability and economics are described in articles [38–40]. In article [41], a model describing the probability of water shortage is proposed. Studies simulating simultaneous occurrences of water shortages were conducted. It was demonstrated that a single failure within the water distribution system ($k = 1$) has a considerable effect on the degree of severity vis-à-vis the scarcity of water. The model can be helpful in determining service standards for the amount of water supplied to consumers.

4. A study on the reliability of rural water distribution systems from the point of view of the pipework

In study [42], the necessity to create two basic models for assessing operational reliability within rural water systems led to the following suggestions:

- a model of node sub-systems, the capture, treatment, pumping, storage and regulation of water pressure,
- a linear subsystem that is a water distribution model.

The above division is related to the need for various methods of assessing the reliability of the node and pipework elements of a rural water supply system. Water distribution sub-systems are characterised by their complex reliability structure, which can be represented as follows:

- graphically, as in a block diagram,
- as a function of the structure, such as in a structure table or an analytical record,
- the full range of minimal ability paths, or inability paths.

Reliability in rural water distribution systems may be determined by one of the following methods [43, 44]:

- via the 'simple decomposition' method,
- via the 'complete decomposition' method,
- via the 'partial decomposition' method,
- via the 'complex decomposition' method,
- via minimum suitability paths, or non-suitability paths,
- via the 'matrix' method,
- via the 'mass service' method.

The 'simple decomposition' method consists of the consecutive transformation of structures which change the system into a number of sub-systems, each made up of simple structures.

The 'complete decomposition' method (also known as the 'system state review', 'complete review', or 'operational state review') consists of the decomposition of all suitable and non-suitable areas within the system.

The 'partial decomposition' method or the 'partial review' method is the same as the 'complete decomposition' method except that it is limited to the occurrence of one or more simultaneous failures. In the case of a rural water distribution system limited to one failure, it is possible to assess its reliability in detail.

The 'complex decomposition' method is based on the fact that the division of a system consisting of n-elements is made with respect to a selected group consisting of k $(1 < k < n)$ elements. 2^k subsystems containing n-k elements are obtained. Decomposition is performed until sub-systems with simple structures are formed.

The 'minimum suitability or non-suitability path' method consists of transforming the real system into one equivalent parallel-to-serial or serial-to-parallel system illustrated by a logical scheme.

'Matrix' methods are used in the case of dependent damages.

'Mass handling' methods are used for homogeneous systems.

In the case of water distribution sub-systems, there is a close relationship between the technical solution used and its reliability rating. The process of assessing the reliability of a water distribution system under fixed flow conditions consists of the following steps [16]:

- The hydraulic calculation of the water distribution system can be carried out in a number of ways, using various numerical calculation programmes that take into account sectional expenditures, the type of network structure and the possibility of using temporary reservations, or in other words, the co-operation of those networks with tanks or reservoirs.
- The simulation of failures consisting of the disconnection of successive sections of water pipes and hydraulic calculations for the existing conditions assuming failure for peak hour water demand.
- Observations are conducted of the values of calculated flows and loss of pressure within the nodes. In the event of the pressure dropping below the required value in the i-th step of the simulation resulting in a lack of water supply to the recipient, a correction of the diameter of the *i*-th section must be made. Faced with the unsuitability of such a solution, it would be necessary to change the structure of the water distribution system.
- Final hydraulic calculations should be made after correcting the diameters or the structure.
- Evaluation should be undertaken of the reliability of the water distribution network using the 'partial decomposition' method.

A reliability analysis for all variants of the distribution system was undertaken on thirty-three rural water systems located in the former voivodeships of Białystok, Toruń and Bydgoszcz.

For simulation purposes, the location of the water supply stations and the structure of the water distribution system were changed. Hydraulic calculations were made for all variants. Calculation results were automatically uploaded by our own NSW software – NSW being 'niezawodność systemów wodociągowych' (water supply system reliability) – used to calculate the reliability level of individual variants using the 'partial review' method. The NSW programme also defined the length of the water supply pipes along the network branches and in the rings [16].

The stationary index of readiness *K*, which determines the probability of the system retaining its suitability to work at any point in time *T*, was used as a criterion for assessing the level of reliability. This probability is expressed by the following formula:

$$K = \frac{T_p}{T_p + T_n} \tag{1}$$

where:

T_p – is the average working time of the water supply system [d],
T_n – is the average renewal time of the water supply system [d].

It follows that the readiness index K depends both on the likelihood of the system becoming impaired and on its renewability. The readiness index characterises the reliability of renewable facilities.

The partial review method consists of making a statement regarding only the most probable elementary states of the system. In the case of rural water distribution systems, only one failure of the water supply pipeline was accepted ($k = 1$). Simulations of failures of particular sections of the network pipelines were performed in order to determine the state of efficiency or disability for each elementary state.

In order to distinguish the above two states, N_{gr} was adopted, which was calculated on the basis of the formula [15]:

$$N_{gr} = \frac{L_Z}{L_C - L_t} \geq 10\% \tag{2}$$

where:

L_z – maximum distance between shut-off valves, which was adopted at 300 [m],
L_c – total length of the water supply network [m],
L_t – length of transit wires [m].

The basis for calculating the reliability of the rural water distribution system was the values of the readiness index K_1 of water supply pipes given for a 1 km length for different pipeline diameters and materials that were adopted from work [15] and included in Table 1.

Table 1. The values of the readiness index $K(L_1)$ for water pipes in rural water distribution systems for a 1 km length of water main [15]

Diameter and type of water pipe	$K(L_1)$
cast iron pipelines	0.9975
steel pipelines	0.9967
PVC pipelines ∅ 90 mm	0.9980
PVC pipelines ∅ 110 mm	0.9981
PVC pipelines ∅ 160 mm	0.9983

The readiness index $K(L_i)$ for particular sections of the pipelines, depending on their length was calculated from the formula [15]:

$$K(L_i) = K(L_1)^{(L_i/1000)} \tag{3}$$

where:

$K(L_i)$ – readiness index for the i-th section of the water pipe depending on the length,
L_i – the length of the i-th section of the water pipe.

During the simulation, the NSW program determines the water deficiency N_i during the failure of each i-th section of the water pipe. The values of N_i were compared with N_{gr}.

If $N_i > N_{gr}$, it was assumed that the system was in a state of disability. Otherwise, when $N_i < N_{gr}$, the system was in a state of efficiency. For all failures of pipelines in which the system was considered to be operational, a partial index of readiness K_i was calculated [15]:

$$K_i = [1 - K(L_i)] \bigcap_{(j=1)}^{n} K(L_j) \quad i = 1, 2, 3, n \quad j \neq i \tag{4}$$

where:

K_i – a partial readiness indicator for the failure of the i-th section of the water pipe,

$K(L_j), K(L_i)$ – readiness indicators for individual pipeline sections including their length L, calculated according to formula (3),

i, j – sections of water pipes.

If the water distribution system was in a state of disability, $K_i = 0$ was assumed. The index of readiness K for the rural water distribution system was calculated by summing the values of the partial index of readiness K_i for the failure of the next sections of water supply pipelines and the entire network in an undamaged state.

A total of 261 variants of rural water distribution systems were analysed with respect to their reliability. The analysis was simplified on the assumption that only staple or sectional expenditure and water uptake is proportional to the length of expenditure of the pipework at L, which, in the case of rural water distribution systems, does not seem to be too simplistic. Expenditures for transit pipework were assumed to be zero [16].

As a result of studies on the influence on the level of reliability of the water distribution system structure, the following dependencies were established for those systems with a pumping-hydrophore structure [16]:

- in the case of water distribution systems with a branched structure, the value of the stationary index of readiness K decreases, depending on the total length of the pipework according to the following formula:

$$K = 0.999858 - 1.05736 \cdot 10^{(-6)} \cdot L \tag{5}$$

$$r = -0.909107 \tag{6}$$

where in formulas (5–10):

K – stationary index of readiness,

L – total length of the pipework, range 2000–40 000 [m],

r – Pearson correlation coefficient.

- for systems with a mixed structure, this relationship assumes the following form:

$$K = 1.00204 - 5.42144 \cdot 10^{(-6)} \cdot L \tag{7}$$

$$r = -0.764953 \tag{8}$$

- branch pipework in rural water distribution systems:

$$K = 1.00076 - 7.15538 \cdot 10^{(-7)} \cdot L \tag{9}$$

$$r = -0.882719 \tag{10}$$

- in the case of pipework built into rings, no relationship between the length of the pipework and the value of the stationary index of readiness was found.

The obtained results show that reliability decreases with an increase in the length of the pipework. With the help of equations, it is possible to predetermine whether a given water distribution system will require additional reliability solutions before carrying out basic calculations.

5. Summary and conclusions

This paper discusses reliability issues in rural water distribution systems. Based on hydraulic calculations and the reliability of selected rural water distribution systems, the interrelationship was demonstrated between reliability levels, the pipework for the water supply system and the length of that pipework. In the case of manifolds, fairly strong relationships were found indicating a decrease in reliability, which is associated with the increase in the length of such pipework. There is no dependency on the pipework in the rings – this may be due to the small amount of such pipework found in rural group waterworks where branched systems predominate. On this basis, it can be stated that when designing water distribution systems, it is desirable to aim to shorten manifolds in order to increase reliability; however, it should be remembered that any correction to waterworks systems should be preceded by hydraulic analysis.

References

[1] Szpindor A., *Zaopatrzenie w wodę i kanalizacja wsi*, Arkady, Warszawa 1992 (in Polish).
[2] Dzienis L., Królikowski A., *Wodociągi i kanalizacje wiejskie,* Wydawnictwo Politechniki Białostockiej, Białystok 1991 (in Polish).
[3] Gheisi A., Forsyth M., Naser G., *Water Distribution Systems Reliability: A Review of Research Literature*, Journal of Water Resources Planning and Management, Vol. 142/2016, Issue 11, doi: 10.1061/(asce)wr.1943-5452.0000690.
[4] Mays L.W., *Review of reliability analysis of water distribution systems*, A a Balkema, Rotterdam 2016.
[5] Mays L.W., *Methodologies for reliability analysis of water distribution systems*, [in:] *Computer Modeling of Free-Surface and Pressurized Flows*, ed. M.H. Chaudhry, L.W. Mays, Kluwer Academic Publishers, Dordrecht 1994, Vol. 274, 485–517.
[6] *Niezawodność miejskich systemów zaopatrzenia w wodę*, ed. A. Wieczysty, Politechnika Krakowska, Kraków 1993 (in Polish).
[7] *Niezawodność i bezpieczeństwo systemów zbiorowego zaopatrzenia w wodę*, ed. J. Rak, Oficyna Wydawnicza Politechniki Rzeszowskiej, Rzeszów 2012 (in Polish).
[8] Rak J., *Wybrane zagadnienia niezawodności i bezpieczeństwa w zaopatrzeniu w wodę*, Oficyna Wydawnicza Politechniki Rzeszowskiej, Rzeszów 2008 (in Polish).
[9] Kwietniewski M., Roman M., Kłoss-Trębaczkiewicz H., *Niezawodność wodociągów i kanalizacji*, Arkady, Warszawa 1993 (in Polish).

[10] *Metody oceny i podnoszenia niezawodności działania komunalnych systemów zaopatrzenia w wodę*, ed. A. Wieczysty, Monografie KiŚ PAN, Kraków 2001 (in Polish).
[11] Królikowska J., Królikowski A., *Analiza porównawcza wskaźników niezawodności wiejskich i komunalnych systemów zaopatrzenia w wodę*, [in:] *Zaopatrzenie w wodę, jakość i ochrona wód: zagadnienia współczesne*, ed. M.M. Sozański, PZITS O/Wielkopolski, Poznań 2010, Vol. 2, 383–392 (in Polish).
[12] Atkinson S., Farmani R., Memon F.A., Butler D., *Reliability Indicators for Water Distribution System Design: Comparison*, Journal of Water Resources Planning and Management, Vol. 140/2014, No. 2, 160–168, doi: 10.1061/(asce)wr.1943-5452.0000304.
[13] Kwietniewski M., Roman M., *Establishing performance criteria of water supply systems reliability*, Journal of Water Supply Research and Technology-Aqua, Vol. 46/1997, No. 3, 181–184.
[14] Kwietniewski M., Rak J., *Niezawodność infrastruktury wodociągowej i kanalizacyjnej w Polsce*, Studia z zakresu inżynierii środowiska No. 67, Polska Akademia Nauk, Komitet Inżynierii Lądowej i Wodnej, Instytut Podstawowych Problemów Techniki, Warszawa 2010 (in Polish).
[15] Dzienis L., *Niezawodność wiejskich systemów zaopatrzenia w wodę*, Rozprawy Naukowe No. 4, Wydawnictwa Politechniki Białostocka, Białystok 1991 (in Polish).
[16] Dawidowicz J., Dzienis L., Królikowski J., *Optymalizacja wiejskich systemów zaopatrzenia w wodę*, Projekt badawczy KBN No. 606239101, Białystok 1992 (in Polish).
[17] Dawidowicz J., *Metoda projektowania wiejskich podsystemów dystrybucji wody z zastosowaniem analizy techniczno-ekonomiczno-niezawodnościowej (ATEN)*, Politechnika Białostocka, 1997, dysertacja (in Polish).
[18] Dąbrowski W., Niedziołek M., McGarity M., *Propozycja sposobu określania wskaźnika gotowości prostych podsystemów zaopatrzenia w wodę*, Czasopismo Techniczne, 1-Ś/2011, 43–55 (in Polish).
[19] Iwanejko R., Bajer J., *Dekompozycja wielokrotna jako metoda dokładnego wyznaczania niezawodności systemów złożonych*, Czasopismo Inżynierii Lądowej, Środowiska i Architektury, JCEEA, Oficyna Wydawnicza Politechniki Rzeszowskiej, Rzeszów 2017, Vol. XXXIV, Issue 64, 287–296, doi:10.7862/rb.2017.246 (in Polish).
[20] Ciaponi C., Franchioli L.,Papiri S., *Simplified Procedure for Water Distribution Networks Reliability Assessment*, Journal of Water Resources Planning and Management-Asce, Vol. 138/2012, No. 4, 368–376. doi: 10.1061/(asce)wr.1943-5452.0000184.
[21] Boccafoschi A., Rejtano B., *Reliability analysis of water supply systems by the entropy and equivocation theory*, Crc Press-Taylor & Francis Group, Boca Raton 2010.
[22] Bao Y., Mays L.W., *Model for Water Distribution System Reliability,* Journal of Hydraulic Engineering, Vol. 116/1990, Issue 9, https://doi.org/10.1061/(ASCE)0733-9429(1990)116:9(1119).
[23] Wagner J.M., Shamir U., Marks D.H., *Water distribution reliability – simulation methods*, Journal of Water Resources Planning and Management-Asce, Vol. 114/1988, No. 3, 276–294.

[24] Abunada M., Trifunović N., Kennedy M., Babel M., *Optimization and reliability assessment of water distribution networks incorporating demand balancing tanks,*. Procedia Engineering, Vol. 70/2014, 4–13.
[25] Bouchart F., Goulter I., *Selecting Valve Location to Optimize Water Distribution Network Reliability*, Proceedings of the 6th IAHR International Symposium on Stochastic Hydraulics, 1992, 18–20.
[26] Bouchart F., Goulter I., *Reliability improvements in design of water distribution networks recognizing valve location*, Water Resources Research, Vol. 27/1991, No.12, 3029–3040, doi: 10.1029/91wr00590.
[27] Gargano R., Pianese D., *Reliability as tool for hydraulic network planning*, Journal of Hydraulic Engineering-Asce, Vol.126/2000, No. 5, 354–364, doi: 10.1061/(asce)0733-9429(2000)126:5(354).
[28] Gheisi A., Naser G., *Water distribution system reliability under simultaneous multicomponent failure scenario*, Journal American Water Works Association, Vol.106/2014, No. 7, 83–84, doi: 10.5942/jawwa.2014.106.0075.
[29] Kwietniewski M., Roman M., *Reliability assessment of water supply systems*, Journal of Water Supply Research and Technology-Aqua, Vol. 46/1997, No. 5, 283–287.
[30] Shamsi U.M., *Computerized evaluation of water-supply reliability*, Ieee Transactions on Reliability, Vol.39/1990, No. 1, 35–41, doi: 10.1109/24.52638.
[31] Tanyimboh T.T., Burd R., Burrows R., Tabesh M., *Modelling and reliability analysis of water distribution systems*. Proceedings of the Water Science and Technology, Vol.39/1999, No. 4, 249–255, doi: 10.1016/s0273-1223(99)00078-5.
[32] Trifunovic N., *Decision Support Tool for Design and Reliability Assessment of Water Distribution Networks*, Crc Press-Taylor & Francis Group. Boca Raton 2012.
[33] Trifunovic N., *Diagnostics of Regular Performance of Water Distribution Networks and its Relation to the Network Reliability*, Crc Press-Taylor & Francis Group, Boca Raton 2012.
[34] Yannopoulos S., Spiliotis M., *Water Distribution System Reliability Based on Minimum Cut – Set Approach and the Hydraulic Availability*, Water Resources Management, Vol. 27/2013, No. 6, 1821–1836, doi: 10.1007/s11269-012-0163-5
[35] Tchorzewska-Cieslak B., *Water supply system reliability management*, Environment Protection Engineering, Vol. 35/2009, No. 2, 29–35.
[36] Wanakule N., Mok C.M., *Reliability based water supply management model*, Proceedings of the Second IASTED International Conference on Water Resources Management, 2007.
[37] Monte M.B.S., de Almeida A.T., *A Reliability-Based Approach to Maximize Availability in a Water Supply System*, Ieee Latin America Transactions, Vol.13/2015, No. 12, 3807–3812.
[38] Asahi C., Hagihara K., *Economic Valuation for Improving Supply Reliability: Risk Countermeasures for Water Quantity and Quality in Water Supply Systems*, [in:] *Coping with Regional Vulnerability: Preventing and Mitigating Damages from Environmental Disasters*, eds. K. Hagihara, C. Asahi, Springer-Verlag, Tokyo, Vol. 4/2016, 151–171.
[39] Zimoch I., Szymik-Gralewska J., *Application of Integrated Reliability-Economic Analysis in Management of Oversized Water Supply Infrastructure*, Annual Set The Environmental Protection, Vol. 37/2015, No. 4, 25–30.

[40] Damelin E., Shamir U., Arad N., *Engineering and economic evaluation of reliability of water supply*, Water Resources Research, Vol. 8/1972, No. 4, 861–877, doi: 10.1029/WR008i004p00861.
[41] Kwietniewski M., *Field reliability tests of water distribution system from the point of view of consumer's needs*, Civil Engineering and Environmental Systems, Vol. 23/2006, No. 4, 287–294, doi: 10.1080/10286600600888540.
[42] Sołowjew A.D., *Analityczne metody w teorii niezawodności*, WNT, Warszawa 1983 (in Polish).
[43] Strömer H., *Mathematische theorie der Zuverlassigkeit. Einführung und Anwendungen*, Akademie Verlag, Berlin 1970 (in German).
[44] Stuckey A.T., *Methods used for the analysis of pipe networks*, WWE No. 877/1969.

TECHNICAL TRANSACTIONS 3/2018
CZASOPISMO TECHNICZNE 3/2018
ENVIRONMENTAL ENGINEERING
DOI: 10.4467/2353737XCT.18.044.8339
SUBMISSION OF THE FINAL VERSION: 5/2/2018

Krzysztof Gaska (krzysztof.gaska@polsl.pl)
Department of Technologies and Installations for Waste Management, Silesian University of Technology

Agnieszka Generowicz (agenerowicz@pk.edu.pl)
institute of Water Supply and Environmental Protection, Faculty of Environmental Engineering, Cracow University of Technology

Józef Ciuła (jozef.ciula@swns.pl)
Nowy Sącz Water Networks Company

A HPC (HIGH PERFORMANCE COMPUTING) BASED MODEL FOR OPTIMISATION OF THE WASTE MANAGEMENT SYSTEMS IN METROPOLITAN REGIONS

ZAAWANSOWANY MODEL OPTYMALIZACJI SYSTEMÓW GOSPODARKI ODPADAMI W METROPOLIACH WYLORZYSTUJĄCY RÓWNOLEGŁĄ ARCHITEKTURĘ OBLICZENIOWĄ

Abstract

The article presents the conception of an intelligent system for monitoring and managing the municipal waste disposal in metropolises. Applying advanced IT solutions using intelligent computational techniques enables the passage from the passive position of self-government units (JST) in managing the waste disposal to the active position, especially in decision making during the problem solving of planning systems associated with the organisation management of the complex infrastructure of the waste disposal. The aim of using ICT systems is an increase in the reliability of the economy of systemic waste, monitoring in real time, the stabilization of the work of the system and the optimization of logistic and technological processes in the context of the raw material, energy application and simultaneously limiting the influence on all components of the environment.

Keywords: SMART City, waste management, predictive models, artificial intelligence, GIS, parallel computing architecture (HPC)

Streszczenie

W artykule przedstawiono koncepcję inteligentnego systemu monitorowania i zarządzania gospodarką odpadami komunalnymi w metropoliach. Zastosowanie zaawansowanych rozwiązań informatycznych wykorzystujących inteligentne techniki obliczeniowe umożliwia na przejście z biernej pozycji jednostek samorządowych (JST) w zarządzaniu gospodarką odpadami do aktywnego działania, w tym szczególnie, podejmowanie decyzji podczas rozwiązywania problemów planistycznych związanych z organizacją systemu zbiórki i systemu transportu odpadów i kompleksowym zarządzaniem złożoną infrastrukturą systemów gospodarki odpadami. Celem wykorzystania systemów ICT jest zwiększenie niezawodności systemów gospodarki odpadami, monitoring w czasie rzeczywistym, stabilizacja pracy systemu oraz optymalizacja procesów logistycznych i technologicznych w kontekście wykorzystania surowcowego, energetycznego przy jednoczesnym ograniczeniu wpływu na wszystkie komponenty środowiska (woda, powietrze, gleba).

Słowa kluczowe: SMART City, gospodarka odpadami, modele predykcyjne, sztuczna inteligencja, GIS, równoległa architektura obliczeniowa (HPC)

1. Introduction – Diagnosis of the market – identification of the needs of local government units in managing waste disposal systems

Transformation of global economies towards effective using natural resources and lowering the release of pollution emissions, incl. greenhouse gases, is becoming one of civilization's key challenges. The negative consequences of climate change or irreparable loss of natural resources, and the lack of integrated actions are manifest already in the near future and are forcing us to take advanced of procedural, technological research on drawing new solutions assisted with ICT systems.

Waste disposal in connection with energy management and low-carbon is establishing the integrated system of action in using all advanced technologies. Implementing effective, technologically advanced solutions in all areas, in the context of reducing the power consumption and materials, is of key importance in increasing the use of renewable energy and introducing environment-friendly technological innovations. The actions to be taken are determined by an adopted strategy for limiting the climate change in global warming. The national economy is being shaped by interpenetrating sectors of the scrap, water-sewage, transport and municipal energy management. The solution suggested by the authors is a response to contemporary challenges in the context of the rapid development of intelligent cities (SMART Cities) in the area of balanced managing the municipal infrastructure in metropolises, including the balance of resources, in the context of increasing the effectiveness of resource use.

The decision support system constituting the central element of the target computer integrated solution fulfils the strategic objectives of waste disposal established in the provisions of law that allows for the identification of scattered sources of waste, on quantitative and qualitative characteristics of waste generated in dispersed sources, to stocktaking and managing tied objects from waste disposal, as well as for assisting decision-making action on the different territorial reach and for simulations of multi-variants of waste management.

The suggested answer takes into account standards and requirements determined in Directives of the EU and refined trends of the ICT development of systems in strategic priority areas.

The advantage of the proposed system above others operating so far is the element of the intelligence including predicting future events and states of components of the system of the waste disposal in particular. This feature in combination with the self-studying concluding intelligent system is not only used to react to current changes in parameters causing the destabilization of the system, but also to anticipate these changes and with the appropriate advance to correct the functioning of the system preparing it for the predicted development of the situation, which is outweighing specifically in folded infrastructure distributed systems about the great inertia.

The technologically advanced solutions suggested by the authors, i.e. the ICT system of the HPC class (High Performance Computing), with the concluding module included, allows (through analysis of the decision space and of events in real time) for predicting future behaviours of individual system components of the waste disposal and effectively to respond

to unknown cases, assisting in taking strategic decisions. The system in question includes advanced tools, mainly on the algorithmic level, for applying new computational technologies (including the possibility of GPGPU/FPGA using for the acceleration calculations, of stream-oriented data processing and multi-variants simulation) in precisely defined models of the space of states and events, constituting components of the database of objects and technological processes of the system of the waste disposal.

In case of the waste market, mass balance is an important aspect of the work of every system, in the context of the demand in raw materials, including the means of conveying the energy in conditions of the appearance of factors causing the destabilization of its correct functioning.

The rapid development of computer and IT technologies is determining the economic development in various sectors of the national economy, and thus close cooperation with scientific bodies is required from entrepreneurs, including small and medium enterprises [3, 5, 8].

Applying advanced computer solutions that exploit intelligent computational techniques enables a transformation from the passive position of local government units (JST) in managing the waste disposal to active action, in particular decision making during the solving of planning systems associated with the organization of the collection system and the system of the transport of waste and monitoring, and comprehensive management of the complex infrastructure of the waste disposal. In order to use ICT systems increasing the reliability of economy systems in waste, monitoring in real time, the stabilization of the work of the system and the optimization of logistical and technological processes in the context of the raw material is needed, energy application while simultaneously limiting the influence on all components of the environment (water, air, soil).

2. Concept of the expert system supporting the processes optimizing management of waste disposal

The Expert system (Fig.1) is based on a predefined component environment – ProWaste Objects .NET, which is one of the ProWasteEnterpriseSolutionFramework .NET environment (author's solution). The real-time HPC (nVidia CUDA) system is a component-oriented environment with a specialized database of knowledge representing the elements of technological processes in the area of water and sewage management, considered as control objects. All variables related to supported devices are updated online through the OPC UA servers of the monitoring and control system. This system creates a platform for advanced algorithm operation, acting as the interface between the control and monitoring systems and the PLC. System configuration ensures permanent monitoring of communication connections and automatic system response in case of a failure, sudden parameter changes, etc. Key performance indicators may also be included in the system, and all calculations required for optimal system performance are performed in real-time using a parallel computing architecture. This advanced multi-parameter control system consists of a main MPC controller and individual MPC controllers for each component device.

The Expert system is based on a process state space model. It is implemented on a higher supervision level in order to manipulate the multiple control loops and to optimize control systems and track changes in process variable values. The state space model is used to forecast the influence of time independent process variables – both manipulated and anticipated – on dependent output variables of the process – so regulated like unregulated predicted variables. The model allows for consideration of the dynamics of the process between the change of independent variables and the expected changes of the dependent variables. The optimization algorithm predicts the future course of the process and compares it to the operating objectives. Internally, the algorithm also calculates the strategy of future decisions, saving only the current changes in the settings of the lower level regulators. This multi-stage computing process, repeated at each execution, enables the driver to plan ahead in order to ensure optimum dynamic control of regulated process variables. EnviroLab.ProWater ensures flexibility in the combination of control targets and control of limits of variable for multiple variables. Each adjustable process variable may contain a target value and/or limiting values limits can be defined as absolute minimum/maximum values or minimum/ maximum deviations from set point values).

One of the main tasks will be to develop advanced prediction control (MPC), adaptive (APC) and inferential control algorithms that will provide:

- Significant reduction in operating costs and higher stability of operating parameters than with classic PID controllers.
- Reduced time for diagnosis and analysis of processes and thus rapid response to interferences and undesired working conditions - through state domain analysis.
- Rapid response to unknown events (interfering signals) destabilizing the technological process. The EnviroLab Enterprise Solution incorporates dedicated HPC (High Performance Computing) tools, advanced prediction models for MPC and adaptive and inferential control and thus provides stable and optimal working performance of the equipment.
- Significant shortening of the design process or modifications of the visualization and process control systems through the use of the RAD.NEnvironT component.
- Significant reduction of process modelling errors, thanks to the author's innovative solutions (predefined NEnviroT intelligent real-time components – processes, objects, events), meaning that all process units can be precisely modelled and optimized to support planning and profit maximization.
- quick adaption of the drivers' operating strategies to economic and environmental goals by using innovative technology in the ProWaste Enterprise Solution. It allows users to adapt quickly to changing economic scenarios, making it easy to design and maintain the driver.
- Significant increase in system performance due to the use of 64-bit architecture
- Real-time knowledge base optimization based on process, computational, and laboratory data
- API to C ++, Java and .NET for individual adapters and direct integration
- Use of Complex Event Processing technology (complex event processing).

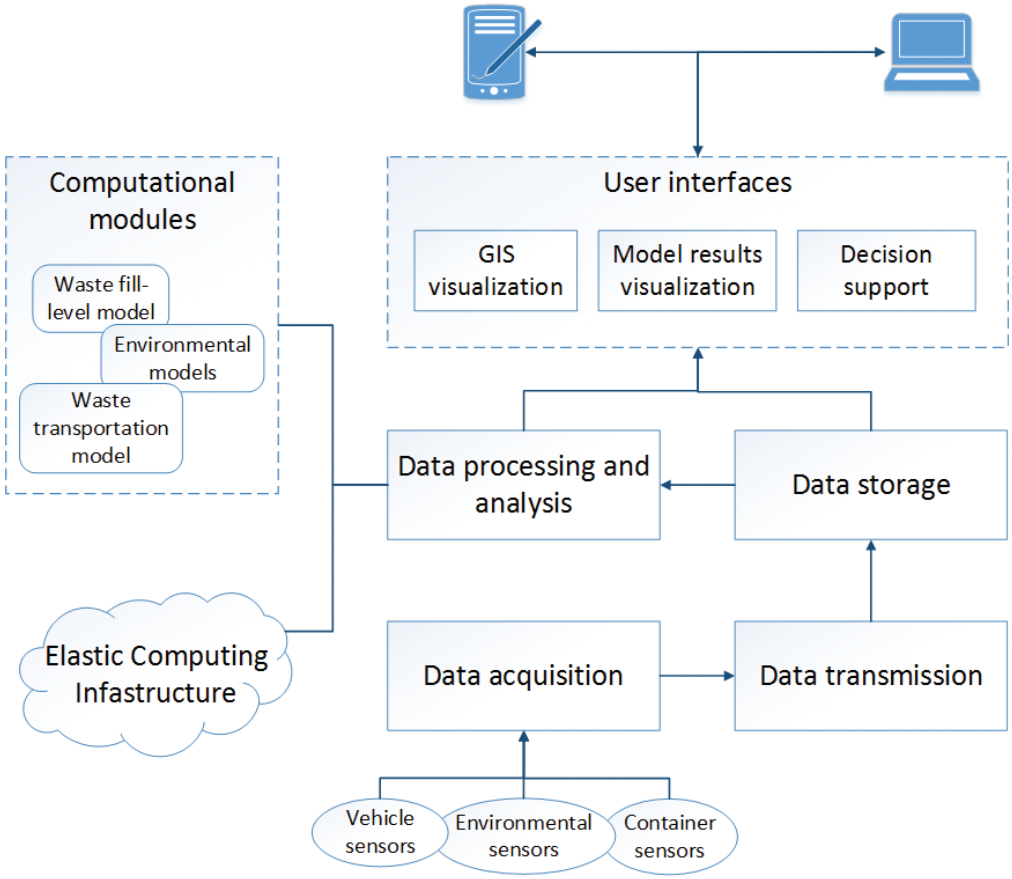


Fig. 1. The Topology of the Expert System

3. The Topology of the Expert System

Control in a hierarchical structure is an effective way to control complex control systems (Fig. 1). In the hierarchical system, the system decomposes into separate subsystems or cells, with autonomous information processing and decision-making. The main reasons for the use of control structures are:

1. The ability to decompose a complex decision problem with a large number of coordinates, the most difficult solution, a few minor simpler decision problems, a smaller number of state coordinates (smaller space of states) solved sequentially or in parallel.
2. Increased flexibility and clarity with modules easier to modify and detect any errors,
3. Shorter computation time (quicker algorithms), thanks to the use of parallel computing architecture
4. Increased reliability of the operation of the control system.

The structure of the hierarchical system (Fig. 2) consists of the main layers (Supervisor, Optimizer, and Follower) and subsystems (slow, medium and fast). Higher levels generate

controls and decisions that help the lower layer perform its tasks. Layer separation is the result of functional decomposition of the control system, and subdivision is the result of decomposition of system dynamics in the time scale. In addition, it is uses the available information about the system such as dynamics and interactions between its individual elements. (knowledge base on control objects).

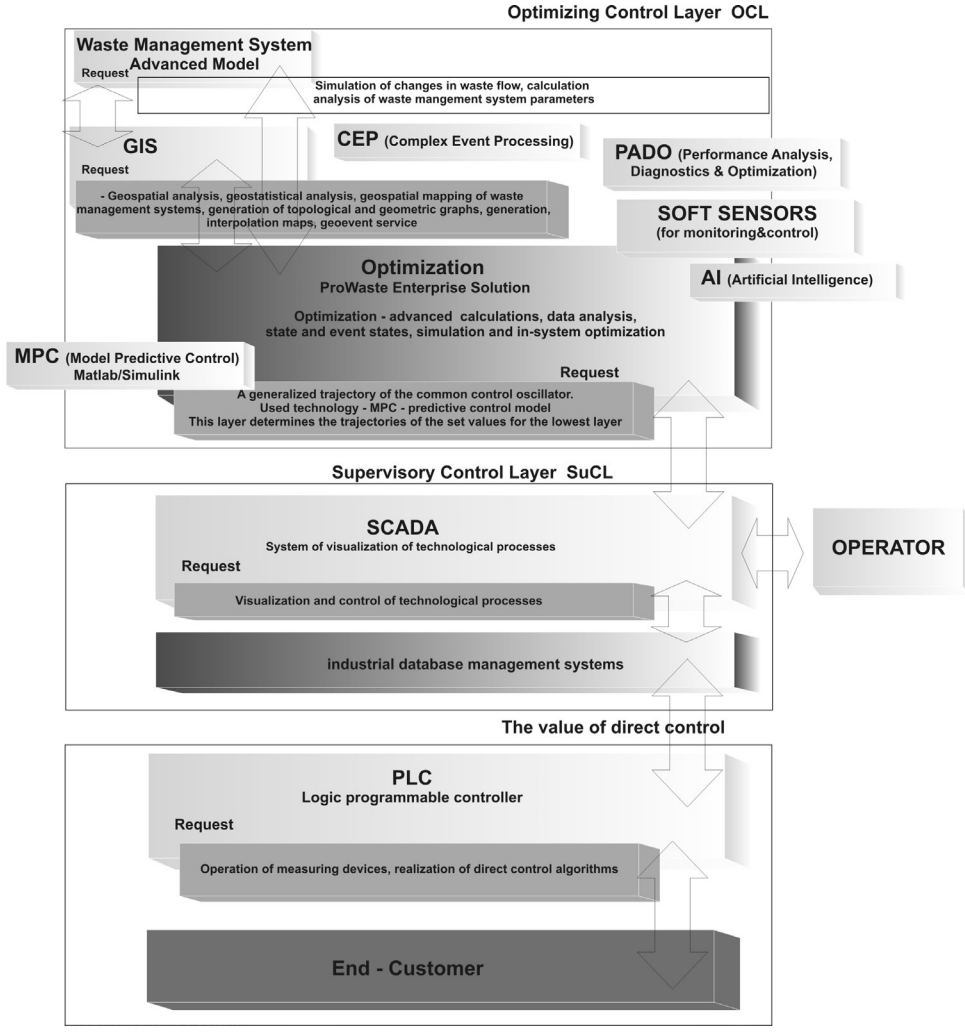


Fig. 2. Architecture/Topology of the System

The Supervisory Control Layer (SuCL) coordinates the individual components of the Waste management structure, evaluates the operational status of the system, and selects the best control trajectory. The observation of the actual control results and its comparison with the forecast allows decision on whether to change the control strategy. Hence, one of the important tasks of this layer is to switch between control strategies [6].

The Optimizing Control Layer (OCL) is responsible for the generation of control trajectories for individual control variables. The MPC – Predictive Control Model is used here. This layer determines the trajectories of the set points for the lowest layer [6, 7, 10].

The Follow Up Control Layer (FCL) is the lowest layer and is responsible for the safety of the processes at the facility, in accordance with trajectories defined in the higher control layer. It has direct access to the controlled object and can implement advanced control strategies developed in higher layers. These tasks implement straight-line algorithms in the form of classical PID algorithms or algorithms written in logic memory of programmable PLC [6, 7, 9].

OCL is divided into three sublayers (slow, medium and fast) corresponding to different dynamics of technological processes in the water supply system and dynamics of interfering inputs. The basic tasks of the slow control sub layer are based on the determination of trajectories of pressures and flows.

The medium control sub layer generates trajectories for output variables controlling the control process, using the MPC model in the full range of the water supply (sewerage) system.

The task of the fast control sublayer is to force the control trajectory defined by the average control sub-set on the object. This sublayer is responsible for meeting the system input control requirements while minimizing power consumption [3, 4].

The operation of the MPC predictive algorithms is based on the knowledge about the future behaviour of the regulated variable in order to determine the value of the control variables. In order to predict the future values of the controlled parameters, the mathematical model of the control object (control object model) uses the previous control signal values and disturbance values (past, present and future) [1- 3, 10]. The basic advantages of predictive control are the following:

- possibility of application for both linear and nonlinear objects / processes,
- the construction of SISO (Single Input Single Output) and MIMO (Multiple Input Multiple Output) control systems,
- consideration of the limitations of process variables,
- incorporation of the internal interaction in the control object by using the object model (the regulator "knows" the object by the control object model),
- incorporation of time delays variable of control objects,
- optimization of the economic indicators related to control.

The MPC algorithm takes into account different types of restrictions:

- limitation of the values of control variables,
- limitation of the increment values of control variable,
- limitation of the output variables,
- limitation of technological variables that are formulated analogously to constraints on output variables.

4. The structure of a virtualized knowledge database on object and process control (MPC)

The expert system supporting the monitoring, control and management of technological processes in water and wastewater management systems is based on the EnviroLab Enterprise Solution Platform component environment. The NVviTT Framework is an advanced technology platform for integrated industrial software (with built-in dedicated modules). The system is scalable in a virtualized modular structure (IFL+CASK+CLOUD). It sets new directions and standards in process simulation, process research and algorithmic control of technological processes, utilizing the latest ICT solutions and advanced predictive and adaptive control models, including laboratory (ACD / Labs, LIMS). The author's innovative solution – the .NavviTT platform, in addition to known solutions, applies real-time knowledge database updated from different, often dispersed sources (process data from SCADA, homo, heterogeneous, structural and non-structural). It uses a number of innovative data processing tools, including validation and statistical analysis.

One of the key elements used in the proposed EnviroLab_Enterprise_Solution is the application of virtual sensors in PLCs to increase the reliability of automation systems by providing rapid response to damage. Virtual sensors are especially useful in situations where it is not possible to measure the process variables (measurement is too expensive; no sensor can be installed in the plant or no suitable measurement method is available).

The main innovations were conducted in the structure of control algorithms, their temporal complexity and the speed of technological process stabilization (followed by optimization). It is crucial to have specialized domain knowledge, dedicated and structured base of knowledge about the expert system, which is logically and physically divided into three independent modules with different functionalities (Fig. 2, 3):

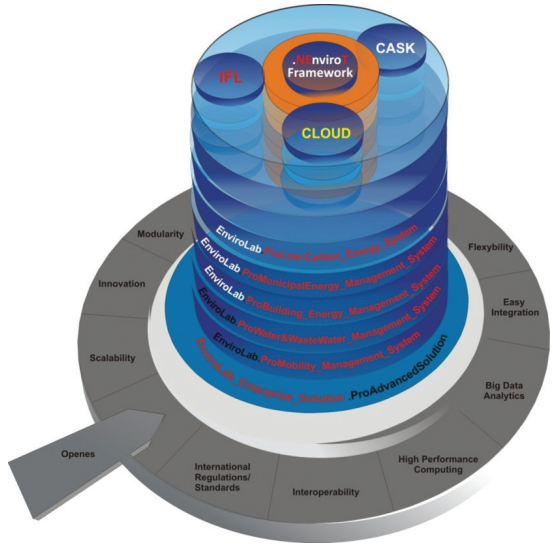


Fig. 3. EnviroLab Enterprise Solution Logical Structure

IFL – Integrated Field Laboratories

The advanced model of virtualized laboratory space supports the process of building and updating the knowledge base of expert system (CASK layer), calibration, validation, testing of MPC/APC control models, event models – by incorporating laboratory test (process, technological, analytical). The IFL environment, through the introduction of the IFL. NEnviroT module is a dedicated tool, including LabResearch_Architect.NEnviroT, which is an expert support system (using LabView, LIMS, ACD / Labs) to conduct research and interpret analysis results. The built-in IFL.NEnviroT advanced research wizard tool allows lab resources to be integrated and a virtual environment to be created – the infrastructure of dispersed labs – by preceding elements of diagnosis, identification, and formulation of a research problem, including the test program required to build a knowledge base for control objects for drivers predictive MPC, adaptive APC, including neural model learning and calibration, validation of object models, events, and real-time updates of knowledge base.

CASK – Cyberinfrastructure, Analytics, Simulation and Knowledge (Expert system in the CASK layer) includes the following functional modules:

- EnviroLab.AdvancedProcessControl (integrated diagnostics and optimization toolkit and advanced APC process control, predictive control of MPC processes)
- EnviroLab.EnvironmentalMonitoring (an advanced HPC solution that simulates pollution propagation and optimizes process parameters to reduce negative environmental impacts).
- EnviroLab.EnterpriseSolutions (open and scalable NET-based platform that utilizes the latest information technology and modular plug-in solutions) dedicated standard. EnviroLab.NEnviroT – provides SDK, API, and many other dedicated / domain tools –which allows applications and standard models to be modified or extended by editing the source code of the class representing the given object, process, etc.

Cloud-GRID – technology

Highly integrated companies are constantly searching for ways to get faster and more efficient data retrieval from industrial installations. One of these is the use of data-driven solutions in the so-called cloud computing, using remote communication through HMI operator interfaces. In the concept of modern industrial networks Industry 4.0 employees and managers have free and remote access to object-level system data. This allows for more efficient resource management, monitoring and control of the machines.

The .NEnviroT platform is based on the environment of intelligent components of process models, events (Complex Event Processing), objects, test models. The knowledge base has crucial significance for the precision of mapping of real processes, objects and events – updated, analysed and validated in real time based on data from various sources (homo/heterogeneous, p/non-structural etc.) computational (e,g. CFD modelling).

The expert system is based on precisely defined models of control objects, where the knowledge base of processes, state spaces, and device performance is built on the basis of expert knowledge and advanced ICT systems. Unlike other solutions, it provides a complete

and integrated environment which incorporates advanced IFL, modelling and simulation (CASK) components by moving some of the CLOUD computing solutions.

The adopted technology of system construction provides maximum flexibility (the ability to expand the system with new modules and thematic databases) and scalability (the ability of the designed system to function efficiently with a growing number of users and increasing volumes of data being processed).

5. Conclusions

The rapid development of computer and IT technologies also determines economic development in various sectors of the national economy, including for waste disposal.

Applying advanced computer solutions exploiting intelligent computational techniques enables the passage from the passive position of local government units (JST) in managing the waste disposal to active action, incl. decision making during the problem solving of planning systems associated with the organization of the collection system and the system of the transport of waste and monitoring and comprehensive managing the complex infrastructure of the waste disposal. In order to use ICT systems to increase the reliability of economy systems is waste monitoring in real time, the stabilization of the work of the system and the optimization of logistic and technological processes in the context of the raw material, energy application while simultaneously limiting the influence on all components of the environment (water, air, soil).

This concept of intelligent system of monitoring and managing the management of the municipal waste in metropolises is a response to contemporary challenges in the context of the rapid development of intelligent cities (SMART Cities) in balanced managing the council infrastructure of the management in metropolises, incl. balance of storage, in the context of increasing the effectiveness of using resources.

References

[1] Dyson B., Chang N.B., *Forecasting municipal solid waste generation in a fast-growing urban region with system dynamics modelling*, Waste Management 25(7), 2005, 669–679.

[2] http://webarchive.nationalarchives.gov.uk/20140328084622/http://www.environment-agency.gov.uk/research/commercial/102922.aspx (access: 22. 01. 2015).

[3] Gaska K., Wandrasz A., *Mathematical modelling of biomass fuels formation process*, Waste Management, 28(6), 2008, 973–985.

[4] Gaska K., Generowicz A., *Modelling of the integrated waste management systems using grapfh-oriented data structures*, XII International Waste Management and Landfill Symposium, S. Margherita di Pula (Cagliari), Sardinia 2009.

[5] Generowicz A., Kowalski Z., Kulczycka J., Banach M., *Ocena rozwiązań technologicznych w gospodarce odpadami komunalnymi z wykorzystaniem wskaźników jakości technologicznej i analizy wielokryterialnej*, Przemysł Chemiczny 90(5), 2011, 747–753.

[6] Generowicz A., Kowalski Z., Makara A., Banach M., *The application of multi-criteria analysis in the management of waste in Cracow, Poland*, Waste Management 32, 2012, 349–351.
[7] Goedkoop M., Spriensma R., *The Eco-indicator 99: A damage oriented method for Life Cycle Impact Assessment. Methodology Report,* Amersfoort, 2001.
[8] Lozano M., Valles J., *An analysis of the implementation of an environmental management system in a local public administration*, Journal of Environmental Management, 82(4), 2007, 495–511.
[9] McDougall F., Hruska J.P., *The use of Life Cycle Inventory tools to support an integrated approach to solid waste management*, Waste Management & Research, 18(6), 2000, 590–594.
[10] White P.R., Franke M., Hindle P., *Integrated Solid Waste Management. A Life Cycle Inventory*, McGraw-Hill, London 1996.

TECHNICAL TRANSACTIONS 3/2018
CZASOPISMO TECHNICZNE 3/2018
ENVIRONMENTAL ENGINEERING
DOI: 10.4467/2353737XCT.18.045.8340
SUBMISSION OF THE FINAL VERSION: 26/2/2018

Justyna Mazur (justyna.mazur93@gmail.com)
student of Faculty of Architecture, Cracow University of Technology

Plants as natural anti-dust filters – preliminary research

Rośliny jako naturalne filtry przeciwpyłowe – badania wstępne

Abstract

In urban areas, particulate matter (PM) are the leading cause of air pollution. They can adversely affect human health. One of the solutions to reduce pollution and improve air quality is the use of plants as natural, biological filters which trap the particles on the leaves. Studies on the assimilation capacity of species (*Hedera helix*, *Parthenocissus quinquefolia, Fagus sylvatica*) were analyzed. In addition, based on two examples of structures in Krakow, the storage capacity of vines such as *Hedera helix* and *Parthenocissus quinquefolia* and *Fagus sylvatica* tree, for fraction of PM (0.2–2.5 μm; 2.5–10 μm; 10–100 μm) was estimated. Analyses have shown the particular efficacy of creeping vines in PM capture.

Keywords: air pollution, PM, green walls, phytoremediation, ecology

Streszczenie

W obszarach miejskich substancje pyłowe są najczęstszą przyczyną zanieczyszczenia powietrza. Cząstki pyłu zawieszonego (PM) mogą szkodliwie oddziaływać na zdrowie człowieka. Jednym z rozwiązań zmniejszających zanieczyszczenie i poprawiających jakość powietrza jest stosowanie roślin jako naturalne, biologiczne filtry, zatrzymujące cząstki stałe na liściach. W artykule przeanalizowano wyniki badań dotyczące zdolności asymilacji pyłów wybranych gatunków: (*Hedera helix, Parthenocissus quinquefolia, Fagus sylvatica*). Dodatkowo oszacowano zdolność akumulacji pnączy *Hedera helix* i *Parthenocissus quinquefolia* oraz drzewa *Fagus sylvatica* dla frakcji PM (0,2–2,5 μm; 2,5–10 μm; 10–100 μm) na przykładzie dwóch budowli w Krakowie. Analizy wykazały szczególną skuteczność pnączy w wychwytywaniu PM.

Słowa kluczowe: zanieczyszczenie powietrza, PM, zielone ściany, fitoremediacja, ekologia

1. Introduction

One of the most harmful components of polluted air in cities are particulate matter PM [1], commonly known until the nineties as Total Suspended Particulate (TSP) [2]. They are a mixture of solid and liquid organic and inorganic substances of various sizes, origin and chemical composition. PM are classified in three fractions with diameters: 0.2–2.5 μm; 2.5–10 μm; 10–100 μm [3–6].

Particles in the atmosphere can also be classified according to their source: primary and secondary [7] or anthropogenic (combustion of solid fuels, transport, etc.) and natural (volcanic eruptions, forest fires, etc.) [1, 3, 6, 8, 9].

Prolonged PM activity can lead to chronic illness and life expectancy shortened by up to three years [3, 8]. PM particles, especially <10 μm may penetrate the lungs, deep into the alveolas and harm human health [1, 4–6, 10] causing bronchitis, asthma, cardiovascular disease or developmental defects [9, 11].

PM particles show many differences in the way they settle on leaves. The dry accumulation of PM on the leaf occurs through gravity sedimentation, capture, deposition and drilling. It depends on the particle size and method of land development [1, 7]. The number of deposited PMs is affected by: remobilization, amount of wax on leaf surface, climate, environmental factors, weather, leaf size and porosity, micromorphology, and plant height [1, 7, 12–14]. It has been shown that PM is most efficiently absorbed by broad, rough leaves, with a large number of hairs, although smaller leaves with complex morphology, can also be effective [3, 8, 14]. One of the studies showed no correlation between PM accumulation and leaf roughness and size [14]. Another study indicates that it is not the amount of wax, but its chemical composition and structure, that affects the absorption capacity of the leaf [3]. The absorption of PM is negatively influenced by: leaf softness, larger gaps between leaves, or smooth surface. It has been observed that the upper part of the leaf is susceptible to higher gravity sedimentation of larger PMs, where twice as many PM are located as compared to the lower part of the leaf, and PM greater than 10 μm are located only on the upper part of the leaf [4, 12].

Plants in cities regulate air temperature and can mitigate the urban heat island effect. [4, 15]. They can be used as a biological filter to capture PM [1, 7]. Fitoremediation is the only easy, affordable and environmentally friendly way to clean the air [3, 8, 13]. Deposition of PM on leaves can affect their physiological processes, such as photosynthesis. Some of them may be absorbed, but they are more likely to stay on the leaf surface for a short time and then return to the atmosphere or to fall to the ground around the tree with the rain, thereby increasing soil contamination [1, 3, 4, 7, 16].

Trees, due to their size, are more effective in capturing atmospheric PM than other types of vegetation, but planting them densely on a small surface will not increase the accumulation capacity of pollutants [7]. The location and size of the trees in the street section may impede the flow of air and reduce the ventilation of the street [11, 16]. Conifer needles produce a thicker wax layer and more efficiently accumulate PM than leaves, but are less resistant to high levels of contamination [1, 3, 8, 14].

Shrubs are less effective than trees, but they play a large role in the lower layers of plantings. In combination with trees, they form an effective structure for PM interception. However, the influence of the planting structure on air circulation and deposition of pollutants should be taken into account [13, 14]. Green walls can cover the vertical planes of walls and buildings [12, 15]. There are two types of vertical green systems: a living wall consisting of modular panels attached to the support structure and a second type, which takes advantage of the natural ability of the vines to climb directly onto the elevation or by use of cables and ties. The superiority of green walls may be due to a large. biologically-active surface (1 m^2 of vines = about 5 m^2 of leaves) and the vertical-to-horizontal ratio [17]. Installing a vertical green, you have to keep in mind the selection of the species, its gravity, the ability to capture the particles, etc. [4, 15].

2. Aim of the work

The aim of the work was to analyze the accumulation capacity of individual PM fractions (0.2–2.5 μm, 2.5–10 μm, 10–100 μm) of three plant species and to calculate the possible amount of PM accumulation by these plant species on the example of two buildings in Krakow. The goal was also to find the best scenario for the harvest of PM by the leaves of the climber, which could occupy the surface of the object possible for greening and by a tree that could grow in the space occupied by the building.

3. Methods and research tools

According to [1, 3, 14, 18], an analysis of results available in the subject literature was carried out pertaining to the ability to accumulate PM fractions for species, that were investigated at least 3 times, except for *Fagus sylvatica* (investigated twice). The analyses were presented in the form of arithmetic means. Extreme results used in data analysis were rejected.

Subsequently, based on the leaf index value available in literature, three species were selected: *Hedera helix, Parthenocissus quinquefolia* and *Fagus sylvatica*, which were used in further analyzes. Leaf area index (LAI) shows the total leaf area of the plant relative to the surface area above which it is located [19]. Two buildings in Krakow were selected, the walls of which can be greened with the vines of the *Hedera helix* or *Parthenocissus quinquefolia*. The objects were selected on the basis of the following criteria: the functional nature of the facility (public functions) and the size of the area for greening. First building (A) is a tenement house "Pod Pająkiem" at Karmelicka Street 35. The second building was (B), Krakow City Hall at Powstania Warszawskiego Avenue 10 (Fig. 1, 2). The base of the building was calculated using measurement tools built into [21]. The elevation of buildings was calculated using LIDAR data available at: [22].

The ability to assimilate PM (0.2–2.5 μm, 2.5–10 μm, 10–100 μm) by *Hedera helix* and *Parthenocissus quinquefolia*, which could grow on the walls of buildings which are available

for greening was calculated. It has also been calculated how many *Fagus sylvatica* individuals could grow on the area occupied by the building and what would their efficiency in absorbing of PM be.

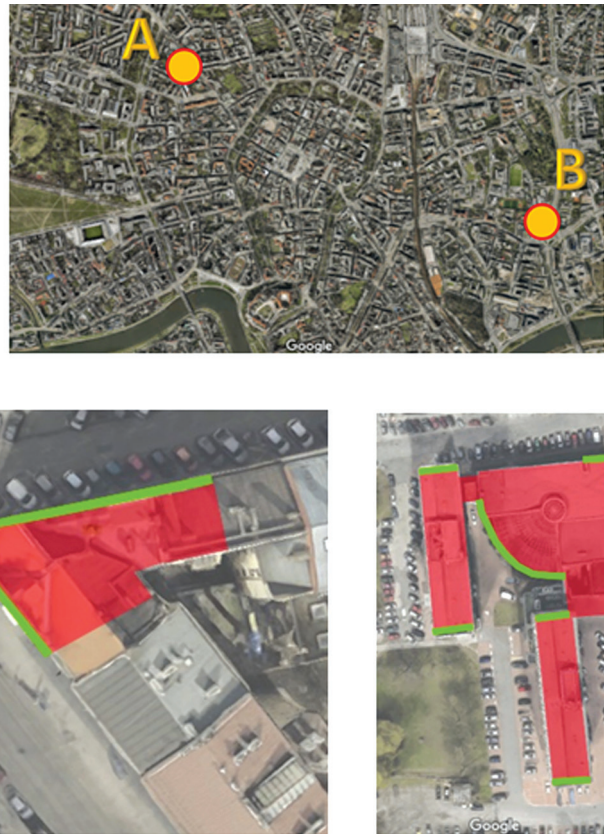


Fig. 1. From the top: location of the analyzed buildings, presentation of the area occupied by the building (red) and location of the elevation available for greening (green). A – tenement house "Pod Pająkiem", B – Municipality of Krakow, own work based on photo from [21]

The following assumptions were taken into account during the analysis: Leaf Area Index (LAI) for the *Hedera helix* for the minimum leaf area = 2.6 m^2/1 m^2 of the wall and for the maximum leaf area = 7.7 m^2/1 m^2 of the wall [4]; LAI for *Parthenocissus quinquefolia* = 5 m^2/1 m^2 wall [17]; leaf area of 1 individual *Fagus sylvatica* (age: 100 years, height 25 m, width 25 m, crown area = 491 m^2 = approx. 1,200 m^2 leaves) [20]. Four experiments have been analyzed, exploring the ability for PM accumulation of different plant species, using a gravimetric method. The averaged results are shown in Table 1.

For vines, the area of the elevation was multiplied by the LAI of a given species and by the average PM accumulation capacity rate according to Table 1. For beech trees, the area of the building's base was divided by the surface of the tree's crown. The result corresponds to the number of trees that could grow on the surface occupied by the building. Then, the

obtained number of trees was multiplied by the leaf area of one individual (1,200 m^2) and by the average PM accumulation value according to Table 1. In the calculations above, the influence of many environmental factors, such as sulphate or nitric oxide concentration, and many others, was omitted.

4. Results analysis

Table 1. Juxtaposition of average PM accumulation values for analyzed species

Species	0.2–2.5 µm [µg/cm^2]	2.5–10 µm [µg/cm^2]	10–100 µm [µg/cm^2]	Sum of PM [µg/cm^2]
Hedera helix	1.08	3.10	16.70	20.88
Fagus sylvatica	0.90	2.30	14.00	17.20
Parthenocissus quinquefolia	1.00	2.60	12.00	15.60

Explanation: The table shows the average analyses values from four studies [1, 3, 14, 18], own work

Based on the data obtained, it can be concluded that *Hedera helix* most effectively accumulates a wide spectrum of PM fraction – 20.88 µg/cm^2. Fagus sylvatica performs slightly worse, capturing an average of about 17.20 µg/cm^2. The weakest assimilation capacity in the wide range was displayed by *Parthenocissus quinquefolia* that accumulates only 15.6 µg/cm^2. *Hedera helix* has a relatively low value of deposited fine PM, which may stem from the small amount of hairs on the leaf surface. It compensates this by the most effective interception of PM 10–100 µm. The reason for the low effectiveness of accumulation of the 0.2–2.5 µm and 2.5–10 µm fraction for *Parthenocissus quinquefolia* may be a small amount of hair on the leaf surface. *Fagus sylvatica*, like *Parthenocissus quinquefolia*, does not have a large number of hairs on the leaf, with only the main nerve on the underside.

Based on the analysis of the subject literature, the number of particles of PM that can be captured by vines (*Hedera helix* and *Parthenocissus quinquefolia*) greening the facades of the surveyed buildings: tenement "Pod Pająkiem" and the Municipality of Krakow were analyzed/estimated, as well as how many PM could be intercepted by tree species (*Fagus sylvatica*) growing on the surface occupied by these buildings. The results are shown in Table 2.

Based on the analyses conducted, no *Fagus sylvatica* (25 m high, 25 m wide) will fit on the surface occupied by the „Pod Pająkiem" tenement house (A) and the covered facade of this building will retain a total of 706.97 g PM for *Hedera helix* and 540.62 g PM for *Parthenocissus quinquefolia*. It follows that – in this case – only the creeping vine will be able to reduce the amount of PM (Fig. 3).

In the same way,the ability to capture PM on the second analyzed building of the Krakow City Hall (B) was investigated. It has been shown that the area currently occupied by the office (2,900 m^2) will fit 5 *Fagus sylvestris* trees (25 m high, 25 m wide crown), which will capture 1,032 g of PM. The area of all leaves of the vines that can cover the undeveloped walls of the office building is 1,300 m^2, which for the *Hedera helix* will allow for the interception

of 1,397.92 g of PM and 1,018.55 g for *Parthenocissus quinquefolia*. It follows that only the *Hedera helix* covering the facade of the building can capture more PM than five *Fagus sylvatica* trees which could grow on the surface occupied by the building. It also follows that five *Fagus sylvatica* trees planted in this area will capture more PM than the coverage of *Parthenocissus quinquefolia* with the area available for greening on the facade of the building.

Table 2. Juxtaposition of analyzed species

Species	Building/ Elevation – Base	0.2–2.5 µm [g]	2.5–10 µm [g]	10–100 µm [g]	Sum [g]
Hedera helix	A/E	38.38	110.16	5,588.43	706.97
	B / E	72.31	207.55	1,118.07	1,397.92
	A/P	25.03	71.84	387.02	483.89
	B/P	161.30	462.99	2,494.15	3,118.43
Fagus sylvatica	A/P	0.00	0.00	0.00	0.00
	B/P	54.00	138.00	840.00	1,032.00
Parthenocissus quinquefolia	A/E	34.50	92.12	414.00	540.62
	B/E	65.00	173.55	780.00	1,018.55
	A/P	22.50	60.08	270.00	352.58
	B/P	145.00	387.15	1,740.00	2,272.15

Explanation: A – "Pod Pająkiem" tenement house (base of the building – 450 m^2, elevation for the creeping vine – 690 m^2), B – Municipality of Krakow (base of the object – 2,900 m^2, elevation for the creeping vine – 1,300 m^2) E – elevation available for greening, P – base surface of the building, own work

Fig. 2. Visualization of green walls on the facades of the analyzed buildings. The area of green walls excluding window and door openings is marked in green. A – tenement building "Pod Pająkiem", B – Municipality of Krakow, own work based on photo from [21]

Fig. 3. Illustration of potential tree plantings (green circles) in the areas occupied by the analyzed buildings (white color). A – tenement building "Pod Pająkiem", B – Municipality of Krakow, own work based on photo from [21]

5. Conclusions

Due to the limited area of greenery in the cities, vines that have a high ability to accumulate PM and a very good individual of horizontal-to-vertical surface ratio should be planted. During increasing urban development, it is recommended to introduce a requirement for greening of elevations, especially on public buildings. When introducing trees and bushes into the city, those species that have the best ability to capture PM, such as *Hedera helix*, should be selected.

Most of the studies were carried out at the end of the season, so the obtained data on PM accumulation did not reflect the weight deposited throughout the growing season. The obtained results, developed on the basis of literature data, are influenced by the number of studies, the date of harvesting of the plant material (autumn), the distance from the source of pollutants, the place of research, the weather or the selection of the measurement method. In order to be able to draw more precise conclusions about the capture of PM by plant foliage, more research is needed with more species and in more diverse locations.

References

[1] Przybysz A., Sæbø A., Hanslin H.M., Gawroński S. W., *Accumulation of particulate matter and trace elements on vegetation as affected by pollution level, rainfall and the passage of time*, Science of the Total Environment 481/2014, 360–369.

[2] Jędrak J., Konduracka E., Badyda J., Dąbrowiecki P., *Wpływ zanieczyszczeń powietrza na zdrowie*, Krakowski Alarm Smogowy, Kraków 2017.

[3] Dzierżanowski K., Popek R., Gawrońska H., Sæbø A., Gawroński W.S., *Deposition of particulate matter of different size fractions on leaf surfaces and in waxes of urban forest species*, International Journal of Phytoremediation 13/2011, 1037–1046.

[4] Ottelé M., van Bohemen H.D., Fraaij A.L.A., *Quantifying the deposition of particulate matter on climber vegetation on living walls*, Ecological Engineering 36/2010, 154–162.

[5] Freer-Smith P.H., Beckett K.P., Taylor G., *Deposition velocities to Sorbus aria, Acer campestre, Populus deltoides X trichocarpa 'Beaupre'', Pinus nigra and X Cupressocyparis leylandii for coarse, fine and ultra-fine particles in the urban environment,* Environmental Pollution 133/2005, 157–167.
[6] Kowalska M., Krzych Ł., *The impact of particulate matter and sulphur dioxide on blood pressure – current knowledge,* Arterial Hypertension, 11/2007, 435–442.
[7] McDonald A.G., Bealey W.J., Fowler D., Dragosits U., Skiba U., Smith R.I., Donovan G.R., Brett H.E., Hewitt C.N., Nemitz E., *Quantifying the effect of urban tree planting on concentrations and depositions of PM10 in two UK conurbations,* Atmospheric Environment 41/2007, 8455–8467.
[8] Popek R., Gawrońska H., Wrochna M., Gawroński S.W., Sæbø A., *Particulate matter on foliage of 13 woody species: deposition on surfaces and phytostabilisation in waxes – a 3-year study,* International Journal of Phytoremediation, 15/2013, 245–256.
[9] Jakubiak M., Urbański K., *Urban planning solutions in the context of dispersion of road pollution,* Journal of Water and Land Development, 30/2016, 71–80.
[10] Jayasooriya V.M., Ng A.W.M., Muthukumaran S., Perera C.J.B., *Green infrastructure practices for improvement of urban air quality,* Urban Forestry & Urban Greening 21/2017, 34–47.
[11] Xue F., Li X., *The impact of roadside trees on traffic released PM10in urban street canyon: Aerodynamic and deposition effects,* Sustainable Cities and Society 30/2017, 195–204.
[12] Weerakkody U., Dover J.W., Mitchell P., Reiling K., *Particulate matter pollution capture by leaves of seventeen living wall, species with special reference to rail-traffic at a metropolitan station, Urban* Forestry & Urban Greening 27/2017, 173–186.
[13] Sæbø A., Hanslin H.M., Torp T., Lierhagen S., Gawronska H., Dzierzanowski K., Gawronski S., *Chemical composition of vegetation along urbanisation gradients in two European cities,* Environmental Pollution 198/2015, 116–125.
[14] Sæbø A., Popek R., Nawrot B., Hanslin H.M., Gawronska H., Gawronski S.W., *Plant species differences in particulate matter accumulation on leaf surfaces,* Science of the Total Environment 427–428/2012, 347–354.
[15] Price A., Jones C.E., Jefferson F., *Vertical Greenery Systems as a Strategy in Urban Heat Island Mitigation,* Water Air & Soil Pollution 226:247/2015.
[16] Nowak J.D., Crane D.E., Stevens J.C., *Air pollution removal by urban trees and shrubs in the United States,* Urban Forestry & Urban Greening 4/2006, 115–123.
[17] Łakomy K., Bobek W., *The modern systems of construction for climbers – technologies, solution and project problems, plants selection,* Technical Transactions, 11/2011, 135–144.
[18] Popek R., Gawrońska H., Gawroński S.W., *Zdolność krzewów do akumulacji mikropyłów z powietrza,* Nauka Przyroda Technologie, 5:6/2011.
[19] Leśny J., Szoszkiewicz K., Juszczak R., Olejnik J., Serba T., *Wskaźnik ulistowienia LAI roślinności drzewiastej I krzewiastej terenów podmokłych,* Acta Agrophysica, 9:3/2007, 673–684.
[20] Malczyk T., *Zieleń w krajobrazie terenów inwestycyjnych,* Oficyna Wydawnicza PWSZ w Nysie, Nysa 2012.
[21] www.maps.google.com (access: 16.02.2018).
[22] http://www.geoxy.pl/pl.lidar_online.html (access: 16.02.2018).

TECHNICAL TRANSACTIONS 3/2018
CZASOPISMO TECHNICZNE 3/2018
ENVIRONMENTAL ENGINEERING
DOI: 10.4467/2353737XCT.18.046.8341
SUBMISSION OF THE FINAL VERSION: 5/3/2018

Tomasz Siuta (tsiuta@iigw.pl)
Institute of Hydraulic Engineering and Water Management, Department of Hydraulics and Hydrology, Cracow University of Technology

The impact of deepening the stilling basin on the characteristics of hydraulic jump

Wpływ zagłębienia niecki wypadowej na zmiany charakterystyki odskoku hydraulicznego

Abstract

In this article, the results of computational fluid dynamic (CFD) modelling of the hydraulic jump conditions occurring in the experimental prismatic rectangular stilling basin with sudden crosswise expansion are presented. The FLOW 3D software program was used to numerically solve Reynolds Navier-Stokes (RANS) equations with the application of the k-ε turbulence model. The influence of the depth magnitude of the stilling basin and the ending sill installation on the hydraulic jump turbulent characteristics and submergence condition changes was investigated. Based on the results of the numerical modelling, it was found that various spatial flow processes contribute to the submergence condition of the hydraulic jump. These processes include: crosswise flow expansion within the stilling basin; local tail water surface level increase and total head loss due to vertical flow contraction; installation of the additional terminal sill. This contribution to the submergence condition allows a reduction in the required depth of the stilling basin, calculated on the basis of a one-dimensional simplified approach without consideration to the spatial characteristic of the hydraulic jump.

Keywords: the stilling basin, the hydraulic jump, CFD modelling, conjugated depth, energy dissipation

Streszczenie

W artykule przedstawiono przykład zastosowania techniki CFD (Numeryczna Mechanika Płynów) do modelowania warunków wystąpienia odskoku hydraulicznego w eksperymentalnej pryzmatycznej niecce wypadowej z nagłym poszerzeniem. W celu rozwiązania równań Reynolds Navier-Stokes (RANS) wraz z modelem turbulencji typu k-ε użyto programu FLOW 3D. Przedmiotem badania był wpływ głębokości niecki wypadowej i wysokości progu wylotowego na zmianę charakterystyki przepływu turbulentnego i warunki zatopienia odskoku. Na podstawie uzyskanych wyników modelowania stwierdzono, iż przestrzenne procesy przepływu, takie jak: poprzeczna ekspansja strumienia w niecce wypadowej, lokalne podniesienie poziomu zwierciadła wody dolnej na skutek kontrakcji pionowej przepływu w przekroju wyjściowym z niecki, instalacja progu wylotowego mogą przyczynić się w istotny sposób do zatopienia odskoku hydraulicznego, a tym samym pozwalają na redukcję wymaganej głębokości niecki wypadowej obliczanej na podstawie uproszczonej jednowymiarowej analizy przepływu bez uwzględnienia przestrzennego charakteru badanego odskoku hydraulicznego.

Słowa kluczowe:

1. Introduction

Hydraulic structures like dams, weirs and culverts release an elevated mass of water within upstream river reach – this leads to the formation of a hydraulic jump just downstream of the structure within the river channel. It is a very important challenge for engineers to investigate different techniques which allow to change the characteristics of hydraulic jump in that way which is preferable to keep the process of energy dissipation under control. One of these techniques is appropriate designing of the stilling basin. Studies of stilling basin performance are of great importance for the construction, maintenance and safety of hydraulic structures. Increases to the efficiency of energy dissipation reduce the risk of downstream erosion of the structure which could constitute a threat to its stability. Designing of stilling basin requires the calculation of its depth and length to satisfy hydraulic jump submergence condition for given design discharge and corresponded tail water level range. In addition to these specifications, structures such as baffle blocks, sills, corrugated beds may be implemented within the stilling basin to enhance the processes of energy dissipation. CFD numerical modelling of the complex spatial flow may be implemented as an important tool to improve the efficiency of the design and exploitation procedures of the stilling basin.

In the article, results of numerical modelling testing of the hydraulic jump within the stilling basin both with and without the sill are presented – this testing was performed with the application of the FLOW 3D software program [1].

The main objective of this numerical test was the investigation of the hydraulic jump stabilisation and submergence within the stilling basin with rectangular cross section (Fig. 1). Flow is discharged into the stilling basin via a chute of the width which is half the width of the stilling basin. This allows for flow expansion in a crosswise direction. Previous literature concerns the experimental [2, 3] and theoretical [4] investigation of the impact of sudden channel expansion on the hydraulic jump performance; however, there is a shortage of numerical modelling of such cases. The hydraulic jump submergence condition for different depths of stilling basin, both with and without the ending sill, was tested. Additionally, the shear stress distribution at the bottom of the stilling basin and the downstream channel reach was determined under condition of the design discharge.

The CFD method allows the simulation of flow processes by discretisation and numerical solution of Navier-Stokes equations for each computational cell [5]. Many researchers apply this method to investigate hydraulic jump performance within a stilling basin or river channel with different turbulent energy dissipation techniques and devices. The majority of works focus on experimental data fitting [6], testing different turbulent models like k-ε, the RNG, hybrid models [7] and air entrainment influence on hydraulic jump characteristics [8, 9]. Others works are focused on hydraulic jump design issues relating to hydraulic structures. There are some techniques inducing and controlling hydraulic jump through the use of application structures like macro-roughness, sudden drop, terminal sill and stilling basin. One of these technique is the application of corrugated beds [10] which increase Reynolds stresses and decrease the second conjugated depth of jump by causing strong turbulence in flow. Numerical simulation of hydraulic jump on a corrugated bed

was evaluated using standard *k*-ε turbulent model [11] and achieved strong agreement with experimental data.

In one study, a FLOW-3D model was applied to evaluate the potential of abrasion in the stilling basin of Colombia Power Plant and to aid the optimisation of the recommended modification of the structure. The model reproduced the recirculation due to the formation of the roller associated to hydraulic jump and small recirculation just downstream of the terminal sill [12].

The USBR (U.S. Bureau of Reclamation) type III stilling basin [13], for which design procedures are well established and based on experimental laboratory data, is commonly used in the US. It is generally applied on canal structures, small outlet works, and small spillways dedicated for cases of Froude numbers above four. Recently some literature has appeared concerning CFD numerical investigation of influence of adverse hydraulic conditions on the hydraulic jump characteristics. Researchers obtain a detailed insight of the role of each basin element and their adapting roles when insufficient tail water conditions exist [14]. Furthermore stepped spillway hydraulic flow condition influence on the size of the basin dimensions was explored. An interesting geometric modification of the USBR basin type is the application of convergence walls. The hydraulic jump performance change due to this geometrical modification which was tested by CFD model [15] and researchers find that in all the discharges cases, the hydraulic jump in the basin with converged walls has larger efficiency of kinetic energy dissipation in comparison to the case of the stilling basin with parallel walls.

2. Numerical experiment

2.1. Geometry of the hydraulic structure

The geometric domain of the model includes: discharging tank, chute-weir, stilling basin and short channel reach. The chute-weir invert is 4 m above the bottom of the channel (Fig. 1). The width of the chute is 5 m and the width of the stilling basin is 10 m. The length of the basin was assumed based on the second conjugated height (h_2) approximation (Table 1) and is 13 m. The magnitude of the design discharge is 50 m^3/s, the corresponding upstream water head is 6.6 m and tail water head is 2 m above the bottom of the channel (Fig. 1). The different depth magnitudes of the stilling basin were investigated to control the hydraulic jump submergence condition.

The following variants of the stilling basin were investigated:

*S*2 – the stilling basin of the depth d = 2 m without the sill and with ending wall inclination 1:1 (Fig. 1)
*S*1 – the stilling basin of the depth d = 1 m without the sill and with vertical ending wall
*S*05 – the stilling basin of the depth d = 0.5 m without the sill and with vertical ending wall
*S*05*p* – the stilling basin of the depth d = 0.5 m with the terminal sill (h_s = 0.5 m)
*S*02*p* – the stilling basin of the depth d = 0.2 m with the terminal sill (h_s = 0.5 m)

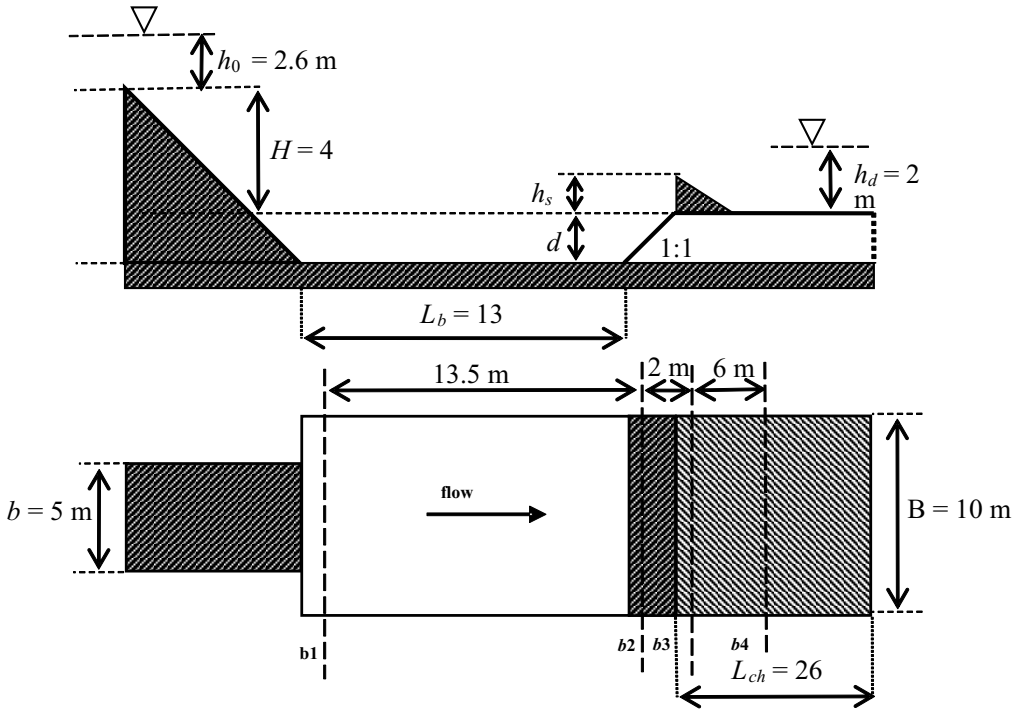


Fig. 1. The stilling basin geometry with the control cross sections ($b1$–$b4$)

2.2. The simplified hydraulic design of the stilling basin

Analysis of the M-y Diagram indicates that the larger value of the unitary flow discharge (q), the smaller the difference of the conjugated depths (h_2–h_1) and the smaller the amount of dissipated energy by the hydraulic jump [16]. The crosswise flow expansion results in a decrease of the unitary flow discharge value at the downstream cross section of the stilling basin – this contributes to more efficient energy dissipation ($\Delta H/H_1$) within the stilling basin (Table 1).

The theoretical depth (Table 1) of the stilling basin (which allow for the hydraulic jump submergence for given flow conditions) was calculated based on one-dimensional hydraulic analysis of the hydraulic jump within the prismatic rectangular channel [16] for the given design *see above note* discharge (50 m^3/s). Such simplified hydraulic analysis in the case of the crosswise flow expansion may lead to significant overestimation of the depth required for hydraulic jump submergence within the stilling basin; therefore, CFD numerical modelling was used to verify this condition.

In order to achieve a reference framework for the tested variants of the stilling basin with flow expansion, the hydraulic jump conjugated depths (eq. 1 and eq. 2), total head loss (ΔH) and corresponding depth of the theoretical stilling basin (required for submergence of the hydraulic jump) were calculated (Table 1) based on two cases of unitary flow rate magnitudes uniformly distributed in the cross section of the theoretical stilling basin without crosswise

flow expansion. The first flow rate magnitude q_1 = 10 m²/s is a value associated with the chute width (5 m) and second magnitude q_2 = 5 m²/s with the width of the tested stilling basin (10 m) for the given design discharge (Q = 50 m³/s) and tail water head (h_d = 2 m). The approximation of the first conjugated depth was calculated based on the Bernoullie equation (eq. 1) assuming the value of the velocity coefficient φ = 0.9. The sequent conjugated, depth was calculated based on the conservation of momentum flux applied for one-dimensional flow in a prismatic rectangular channel (eq. 2). Finally, the reference depth values (d) of the theoretical stilling basin required to satisfy submergence of the hydraulic jump under condition of the investigated unitary flow values (q_1 and q_2) were fixed (Table 1).

$$h_1 = \frac{Q}{\varphi \times b \times \sqrt{2g\left(H_{up} - h_1\right)}} \tag{1}$$

where:

h_1 – first conjugated depth, m,
H_{up} – total upstream head above the bottom of the stilling basin, m,
Q – discharge magnitude, m³/s,
b – the width of the stilling basin, m,
φ – velocity coefficient.

$$h_2 = 0.5h_1\sqrt{1+8\frac{\left(\frac{Q}{b}\right)^2}{gh_1^3}} - 1 \tag{2}$$

where:

h_2 – second conjugated depth, m,

$$\Delta H = H_1 - H_2 = \frac{\left(h_1 - h_2\right)^3}{4 \cdot h_1 \cdot h_2} \tag{3}$$

where:

H_1 – total energy head at the cross section of h_1, m,
H_2 – total energy head at the cross section of h_2, m,
ΔH – total energy head loss within the hydraulic jump, m.

Table 1. The calculated depths (d) of the theoretical stilling basin of different widths (5 m and 10 m) and hydraulic jump parameters: conjugated depths – h_1, h_2, energy head loss – ΔH and relative energy head loss – $\Delta H/H_1$ (eq. 3)

q [m²/s]	d [m]	Fr_1 [–]	h_1 [m]	h_2 [m]	ΔH [m]	$\Delta H/H_1$ [%]
5	0	4.8	0.48	3	2.8	47
10	0	3.2	1	4	1.7	29
5	1.2	5.1	0.45	3.2	3.6	52
10	2.5	3.4	0.85	4.5	2.1	40

2.3. Mathematical model of the turbulent flow

The Flow-3D computational fluid dynamics program was used for the numerical solution of the Reynolds-averaged Navier–Stokes (RANS) momentum transport equations (eq. 5) and continuity equations (eq. 4). This program uses the finite volume approach with a staggered grid for discrete representation of the governing equations. All equations are formulated with area and volume porosity functions. This formulation, called the 'fractional area/volume obstacle representation' (FAVOR) method is used to model complex geometric regions and solid boundaries [2]. The volume of fluid (VOF) method is employed in FLOW-3D to properly capture the free-surface of liquid; the additional transport equation is added (eq. 6).

The general governing equations for incompressible flow, including the FAVOR variables, are given by:

$$\nabla \circ \left(\mathbf{u}\mathbf{A}_{\mathbf{f}} \right) = 0 \tag{4}$$

$$\frac{\partial \mathbf{u}}{\partial \mathbf{t}} + \frac{1}{\mathbf{V}_{\mathbf{f}}} \left(\mathbf{u}\mathbf{A}_{\mathbf{f}} \circ \nabla \right) \mathbf{u} = -\frac{1}{\rho} \nabla \mathbf{P} + \mathbf{G} + \mathbf{f} \tag{5}$$

$$\frac{\partial \mathbf{F}}{\partial \mathbf{t}} + \frac{\mathbf{F}}{\mathbf{V}_{\mathbf{f}}} \nabla \circ \left(\mathbf{u}\mathbf{A}_{\mathbf{f}} \right) = 0 \tag{6}$$

where:

- $\mathbf{u}$ – fluid velocity vector, m/s,
- $\mathbf{f}$ – viscous accelerations vector, m/s^2,
- $\mathbf{P}$ – pressure, N/m^2,
- $\mathbf{G}$ – body accelerations vector, m/s^2,
- $\mathbf{V}_{\mathbf{f}}$ – volume fraction, m^3,
- $\mathbf{A}_{\mathbf{f}}$ – area fraction vector, m^2,
- $\mathbf{F}$ – fluid fraction.

The RANS model includes a formulation for the turbulent kinetic energy (k) and the rate of turbulence dissipation (ε) to obtain Reynolds stress (τ_{Rij}) tensor and turbulent viscosity as follows:

$$\tau_{Rij} = \upsilon_t \left[\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right] - \frac{2}{3} \delta_{ij} k \tag{7}$$

$$\upsilon_t = 0.9 \frac{k^2}{\varepsilon} \tag{8}$$

where:

- u_i – velocity component $i = 1, 2, 3$, m/s,
- v_t – turbulent viscosity, N · s/m^2,
- k – turbulent kinetic energy, J/kg,
- ε – rate of turbulence dissipation, J/(kg · s),
- τ_{Rij} – Reynolds stress tensor, N/ m^2.

In this work, the k-ε turbulence model was used [1, 2].

2.4. Boundary condition

The upstream face boundary condition is in the form of constant in time flow rate ($Q = 50$ m^3/s) discharged into the chute (Fig. 1) and the downstream face boundary is in the form of constant in time hydrostatic pressure distribution. Tail water head (h_d) at the ending cross section of the channel is equal to 2 m. A constant value of surface roughness coefficient ($k_r = 0.001$ m) was assumed for the solid boundary.

The flux control cross sections (baffle) were defined within the numerical model (Fig. 1) in order to calculate energy loss as result of: the hydraulic jump ($b1$–$b2$); sudden vertical contraction at the ending cross section of the stilling basin ($b2$–$b3$); flow along the short channel reach (6 m length) just downstream of the stilling basin ($b3$–$b4$). Thus, three control volumes were defined. In each baffle, the total hydraulic energy rate is calculated and the flux-averaged total head is then fixed (eq. 8).

$$H_e = \frac{\int_A H \cdot \mathbf{u} \circ \mathbf{n} \cdot dA}{Q} \tag{9}$$

where:

H_e – flux-averaged hydraulic head at control cross section, m,
H – total hydraulic head of streamline at control cross section, m,
$\mathbf{u}$ – resultant velocity vector, m/s,
$\mathbf{n}$ – normal unitary vector,
A – area of control cross section, m^2,
Q – volumetric flow rate, m^3/s.

2.5. Results of the numerical modelling analysis

The results of the numerical modelling for different variants of the stilling basin enabled the evaluation of the hydraulic jump submergence conditions for given values of the design discharge ($Q_d = 50$ m^3/s) and tail water head ($h_d = 2$ m). In the case of S2, the hydraulic jump is fully submerged (Fig. 2), the total head loss magnitude by this hydraulic jump is: $\Delta H_e = 2.65$ m (Table 2) and this value is 0.55 m larger than the calculated value (Table 1) in the case of the hypothetical flow without crosswise expansion ($q = 10$ m^2/s and $d = 2.5$ m). The flow layer under the roller is in the shape of a thin prism (Fig. 2a) – this is consistent with the experimental data relating to velocity distribution characteristics [4]. The velocity magnitude reaches 8.5 m/s in the middle of the stilling basin and decreases at the ending cross section of the stilling basin to a value of 4.5 m/s where the vertical component of the velocity (about 2.1 *m/s*) is dominant.

This large value of the vertical component of the velocity is associated with abrupt flow vertical contraction and a water surface level rise of about 0.25 m above the level of the tail water surface. This surface level increase magnitude is comparable with the magnitude of the energy head loss due to vertical flow contraction (Table 2).

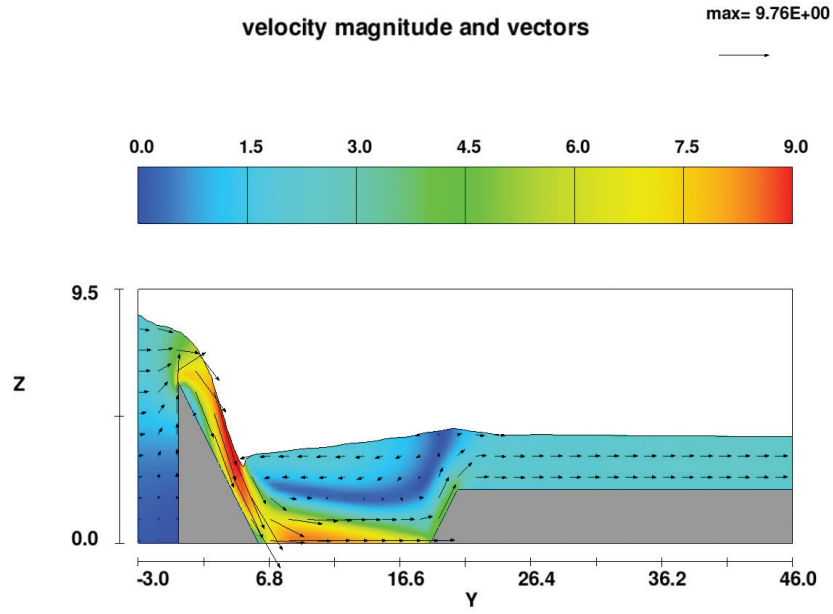


Fig. 2a. Velocity distribution at the central vertical cross section of the stilling basin for the design flow rate condition (case S2)

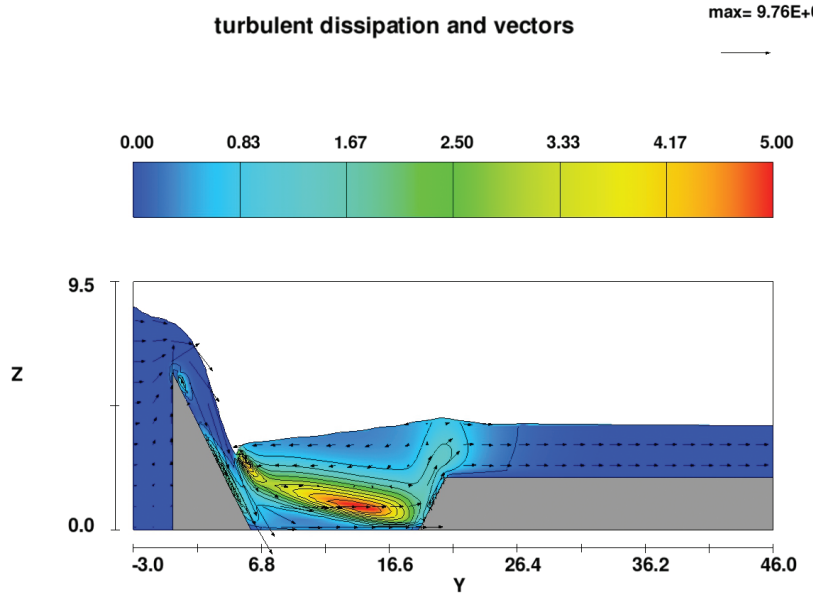


Fig. 2b. Turbulent dissipation rate distribution at central vertical cross section of the stilling basin for the design flow rate condition (case S2)

The turbulent energy dissipation rate magnitude reaches maximum value (5 J/kg/s) within the shear layer about 1 m above the bottom in the middle of the stilling basin (Fig. 2b) and is associated with the maximum magnitude of velocity gradient.

Table 2. Table 2. Flux-averaged hydraulic head (H_e) and head losses (ΔH_e) at flux baffles for different basin cases and for given design discharge (Q_d = 50 m^3/s)

Basin case/ baffle	S2		S1		S05		S05p		S02p	
	H_e [m]	ΔH_e [m]	H_e [m]	ΔH_e [m]	H_e [m]	ΔH_e [m]	H_e [m]	ΔH_e [m]	H_e [m]	ΔH_e [m]
b1	7.57	-	6.56	-	6.15	-	6.15	-	5.96	-
b2	4.92	2.65	4.03	2.53	3.97	2.18	3.35	2.8	3.49	2.47
b3	4.71	0.21	3.54	0.49	3.10	0.87	2.95	0.40	2.99	0.50
b4	4.4	0.31	3.46	0.08	2.95	0.15	2.85	0.10	2.69	0.30

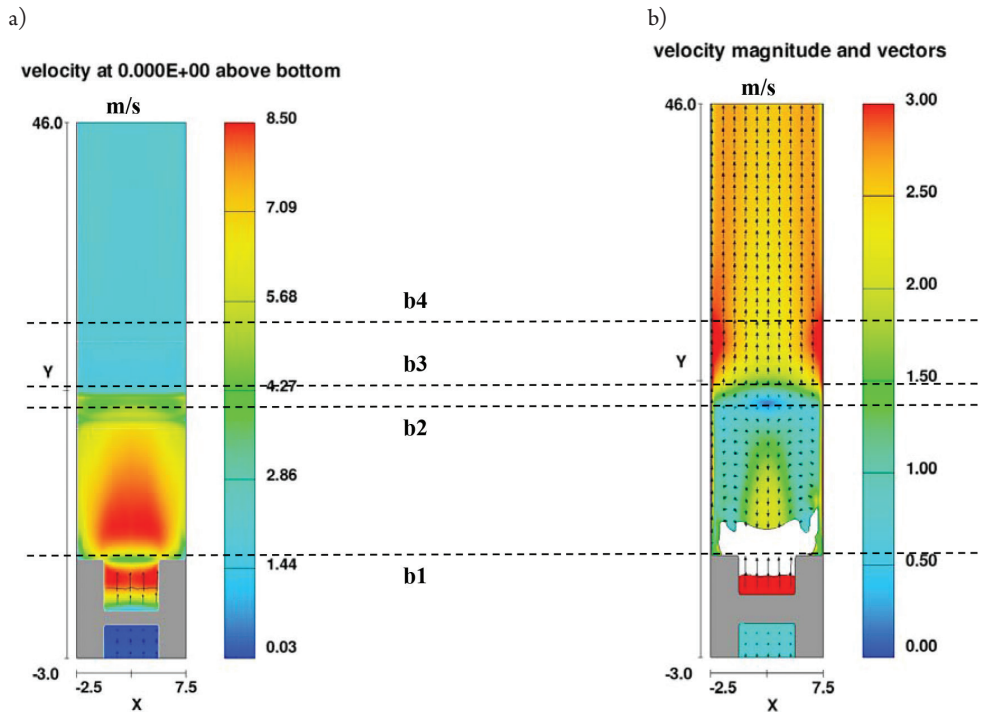


Fig. 3. Velocity distribution: a) at the bottom; b) at a level of 1.5 m above the channel bottom for the design see above note flow rate condition (case S2)

Within the control zone b3–b4 of the channel reach, the value of the turbulent energy dissipation rate is less than 0.5 J/kg/s and velocity is almost uniformly distributed (within the range: 2.8–2.9 m/s) at the downstream cross section of the channel (Fig. 3).

The hydraulic conditions described above are highly acceptable from the point of view of the safe exploitation of the stilling basin *S*2. In spite of this, the hydraulic jump within *S*2 is submerged under the condition of the design discharge and tail water head magnitude, the reduction of depth of the stilling basin will often allow reduced costs of investment; therefore, a stilling basin of lower depths (1 m, 0.5 m and 0.2 m with or without the ending sill) were tested for the determination of the hydraulic condition.

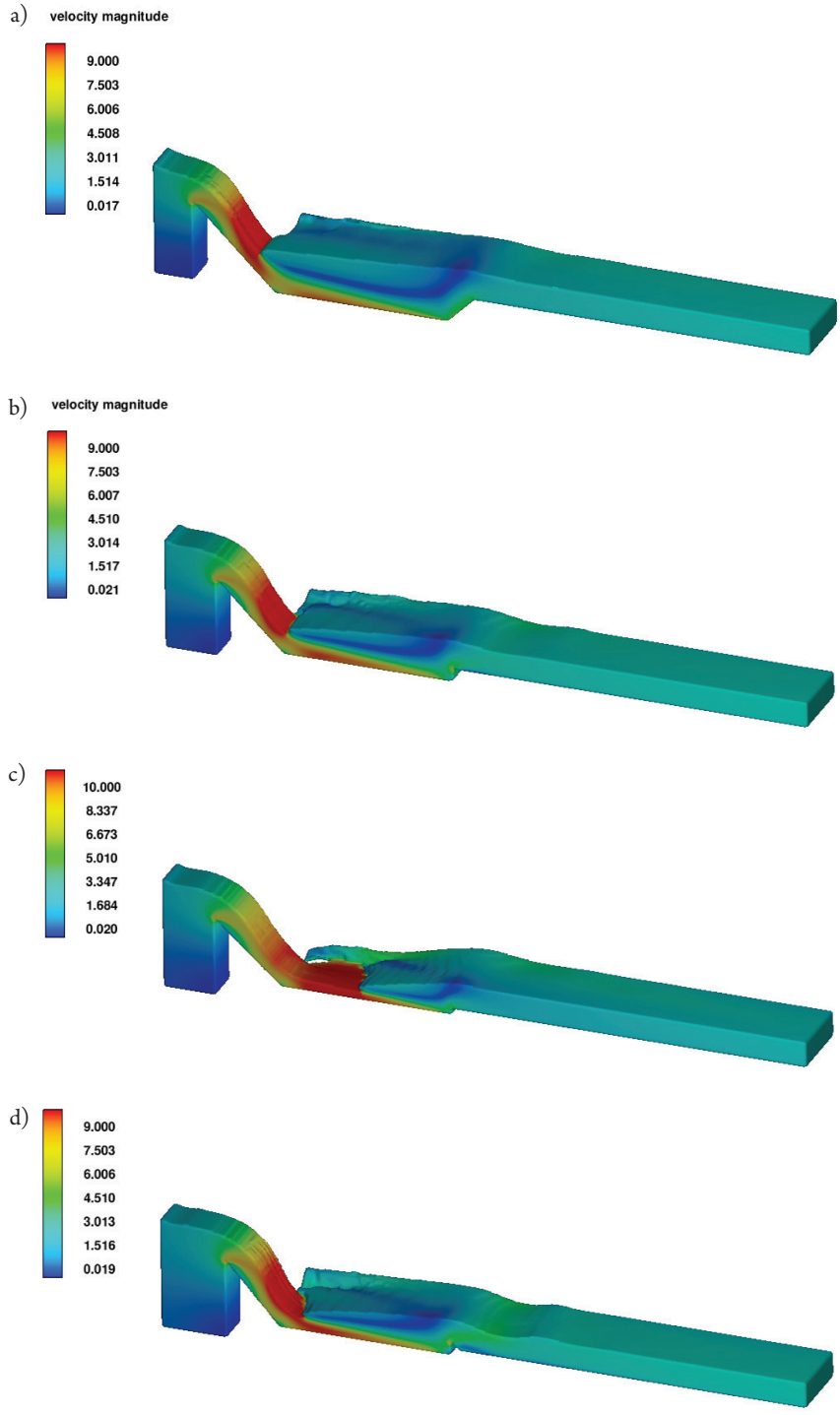


Fig. 4. Spatial velocity distribution (longitudinal symmetrical half of the picture) of the flow domain: a) case S2, b) case S1, c) case S05, d) case S05p

In the case of *S*1, the hydraulic jump is still submerged (Fig. 4b) as is also implied by the simplified calculations for the uniform inflow (q = 5 m^2/s and d = 1.2 m) in a rectangular prismatic channel without crosswise expansion (Table 1). The stream flow expansion within the stilling basin may contribute to the reduction of the sequent conjugated depth of the hydraulic jump, which in the case of *S*1 is equal to 3.2 m based on the results of numerical modelling. In all tested cases, a sudden significant rise in the water surface level close to the sidewalls of the stilling basin is identified (Fig. 4). The significant transversal velocity component magnitude (Fig. 5) contributes to the kinetic energy exchange into the potential energy within sidewall zones.

This energy exchange is the source of additional energy loss and indicates the spatial character of the hydraulic jump.

The significant contribution to kinetic energy dissipation, and thus the sequent conjugated depth of the hydraulic jump reduction, is the effect of the vertical flow contraction due to the vertical ending wall of the *S*1, and the total head loss over the control zone b2–b3 is ΔH_e = 0.5 m (Table 2).

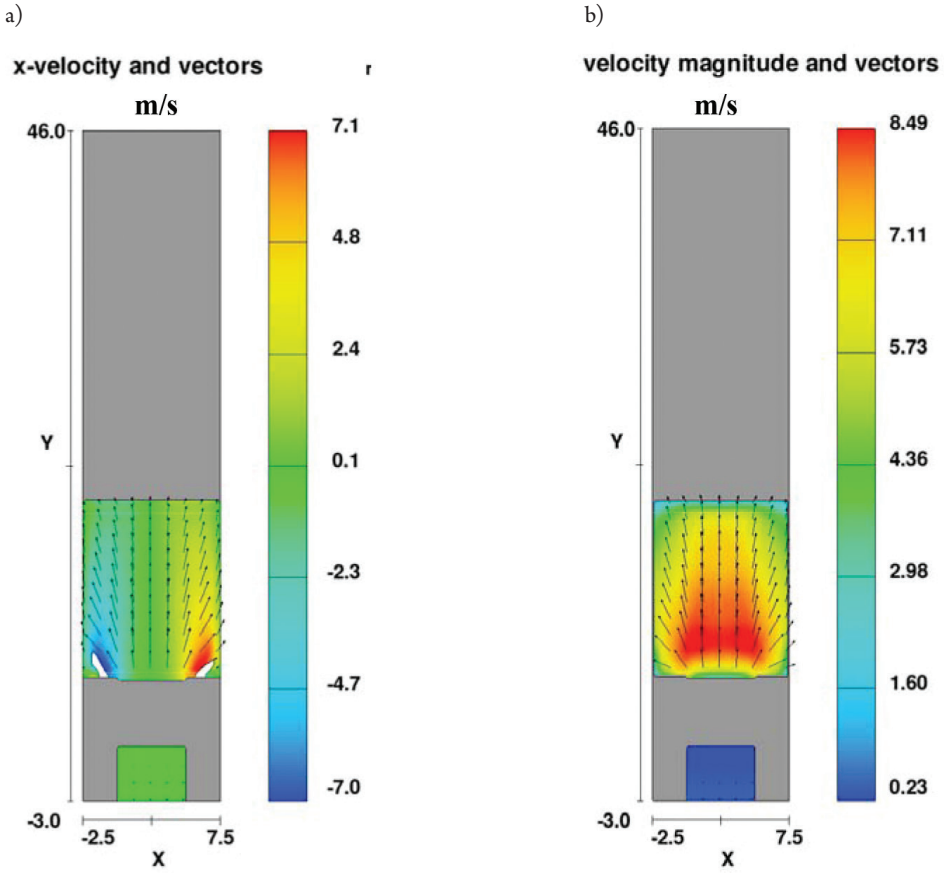


Fig. 5. Velocity distribution at the bottom of the stilling basin: a) x-velocity component; b) resultant velocity magnitude (case *S*1)

As a result of the depth of the stilling basin reducing to 0.5 m, the hydraulic jump is not yet submerged but remains stabilised within the stilling basin (Fig. 4c). In this case, kinetic energy dissipation by the hydraulic jump is ΔH_e = 2.2 m and is smaller than those modelled for other cases, but in the zone of the vertical flow contraction (section b2-b3), the total head loss (ΔH_e = 0.87 m) is the largest in magnitude (Table 2).

Within the stilling basin where supercritical flow condition occurs (Fr = 3), the largest value of bottom shear stress (τ_b = 350 Pa) is identified. Maximum values of the shear stress occur close to the chute outlet and symmetrically on each side (Fig. 6). The shear stress value at the bottom of the channel is not larger than 20 [Pa] and is comparable with other tested cases.

In spite that the hydraulic jump within S05 is not submerged the major energy dissipation occur inside of the stilling basin (Table 2) and at its exit zone therefore total head loss within control section b3-b4 of the channel is small and equal about 0.15 m. It is likely that the unsubmerged hydraulic jump would move downstream due to a small decrease in the tail water surface level; therefore, the impact of the ending sill (h_s = 0.5 m) application within the stilling basin on hydraulic jump submergence was tested (variant S05*p*). In this case, the magnitude of total head loss within control section b1-b2 increases (ΔH_e = 2.8 m) and the hydraulic jump becomes submerged (Fig. 6d).

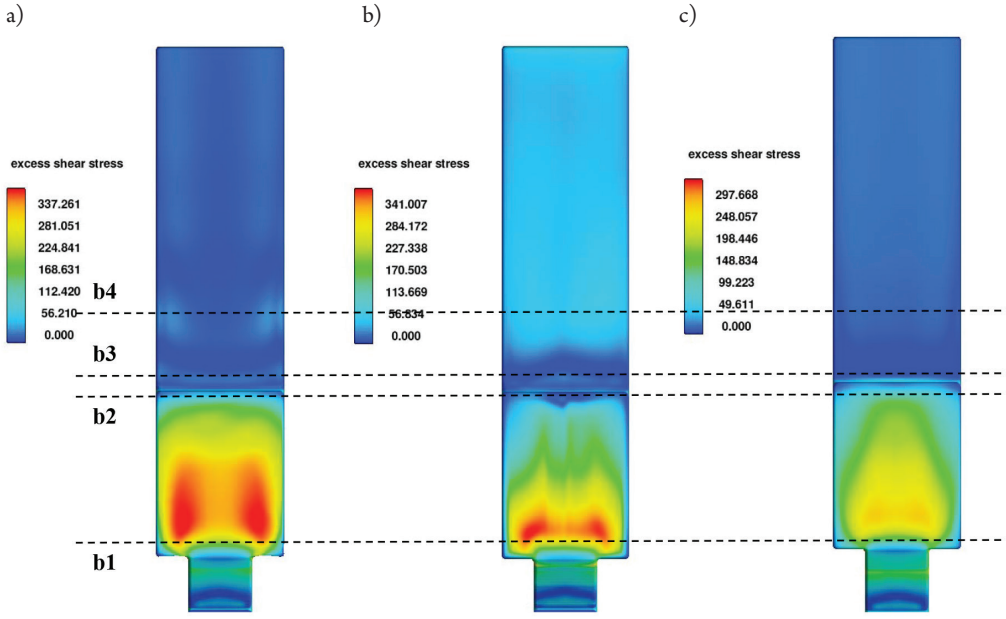


Fig. 6. Shear stress distribution [Pa] at the bottom of the stilling basin and channel: a) case S02*p*; b) case S05*p*; c) case S1

Taking into account cost reduction due to the application of the lower depth stilling basin and fulfilment execution of the hydraulic jump submergence condition, variant S05*p* seems to be optimal. Nevertheless, significant decreasing of the tail water surface level (in this case by 1 m or more) results in the occurrence of unwanted pulsation of the weak supercritical flow (*Fr*<1.6) which could be avoided through the application of solution S1 or S2.

If the depth of the basin was decreased to a value of d = 0.2 and the ending sill was used (variant *S02p*), the hydraulic jump would be not submerged and strong waves on the water table within the channel would be observed (Fig. 7). In this case, a small decrease in the tail water table level may cause motion in the hydraulic jump downstream in the channel leading to potential erosion downstream in the channel. Just downstream of the stilling basin, there are regions with a large magnitude of shear stress gradient (Fig. 6a) which may contribute to the erosion of the channel bottom in regions close to the sidewalls of the channel.

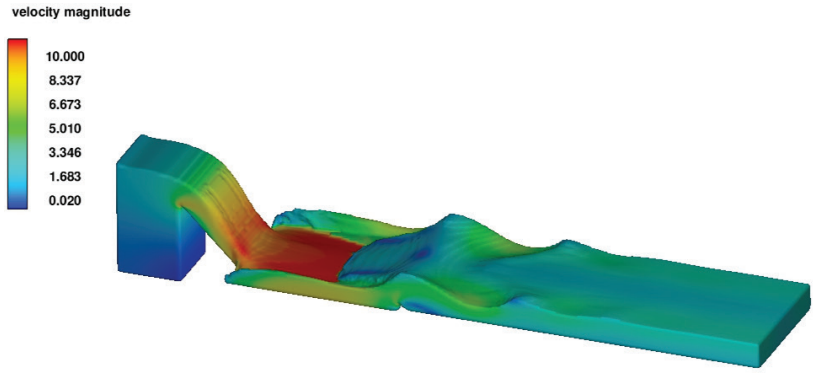


Fig. 7. Spatial velocity distribution and hydraulic jump (case *S02p*)

3. Conclusion

The three-dimensional numerical CFD modelling of the turbulent flow in the stilling basin with sudden crosswise expansion allows the capture of the spatial characteristics of the hydraulic jump and enables testing the influence of the stilling basin depth magnitude changes on the submergence condition of the hydraulic jump. The results of the numerical modelling indicate that such factors as crosswise flow expansion, vertical flow contraction resulting from a stilling basin deepening, and installation of the terminal sill lead to decreases in the sequent conjugated height of the hydraulic jump and to local increases in the level of the tail water table which significantly contribute to the hydraulic jump submergence condition. In the practice it means that the stilling basin of smaller depth than one implies from the one dimensional simplified analysis (in this studied case, depth reduction occurs from 2 m to 1 m or to 0.5 m under condition of the ending sill installation) could be applied to disperse turbulent kinetic energy of the flow from the spillway efficiently. The CFD numerical modelling may contribute to the more realistic designing of the stilling basin and thus may improve the operational safety profile and reduce costs of an investment.

References

[1] Flow Science, Inc., 2014, *FLOW-3D User Manual, Release* 11.0.3, USA 2014.

[2] Matin M.A., Hasan M, Islam M.R., *Experiment on hydraulic jump in sudden expansion in a sloping rectangular channel*, Journal of Civil Engineering (IEB), 36(2)/2008, 65–77.

[3] Urbański J., Siwicki P., *Zastosowanie programu CFD fluent do obliczeń charakterystyk turbulencji strumienia w dolnym stanowisku jazu*, Infrastruktura i Ekologia Terenów Wiejskich, No. 2007/4(2).

[4] Gandhi S., *Characteristics of Hydraulic Jump*, International Journal of Mathematical, Computational, Physical, Electrical and Computer Engineering, Vol. 8, No. 4/2014.

[5] Tannehill J.C., Anderson D.A., Pletcher R.H., *Computational Fluid Mechanics and Heat Transfer*, 2nd Ed., Taylor & Francis, USA 1997.

[6] Bayon-Barrachina A., Amparo Lopez-Jimenez P., *Numerical analysis of hydraulic jumps using OpenFOAM*, Journal of Hydroinformatics, 17(4)/2015, 662–678.

[7] Carvalho R.F., Lemos C.M., Ramos C.M., *Numerical computation of the flow in hydraulic jump stilling basin*, Journal of Hydraulic Research 46(6)/2008, 739–752.

[8] Chanson H., Gualtieri C., *Similitude and scale effects of air entrainment in hydraulic jump*, Journal of Hydraulic Research 46(1)/2008, 35–44.

[9] Mortensen J.D., Barfuss S.L., Johnson M.C., *Scale effects of air entrainment by hydraulic jumps within closed conduits*, Journal of Hydraulic Research, Vol. 49/2011, 90–95.

[10] Ead S. A., Rajaratnam N., *Hydraulic jumps on corrugated beds*, Journal of Hydraulic Engineering, ASCE, Vol. 128, No. 7/2002, 656–663.

[11] Abbaspour A., Farsadizadeh D., Dalir A H., Sadraddini A.A., *Numerical study of hydraulic jumps on corrugated beds using turbulence models*, Turk. J. Eng. Environ. Sci., 33(1)/2009, 61–72.

[12] Amorim Jose Carlos C., Rodrigues Cavalcanti R., Marques Marcelo G.A., *Numerical and Experimental Study of Hydraulic Jump Stilling Basin*, Advances in Hydro-Science and Engineering, Vol. VI/2007.

[13] Peterka A.J., *Hydraulic design of stilling basins and energy dissipators*, Engineering Monograph 25, U.S. Bureau of Reclamation 1963.

[14] Valero D., Bung D., Crookston B., Matos J., *Numerical investigation of USBR type III stilling basin performance downstream of smooth and stepped spillways*, [in:] B. Crookston & B. Tullis (eds.), *Hydraulic Structures and Water System Management*, 6th IAHR International Symposium on Hydraulic Structures, Portland, June 2016, 652–663.

[15] Babaali H., Shamsai A., Vosoughifar H., *Computational modeling of the hydraulic jump in the stilling basin with convergence walls using CFD codes*, Arabian Journal for Science and Engineering., 40(2)/2015, 381–395.

[16] Ven Te Chow, *Open-Channel Hydraulics*, McGraw-Hill, New York 1959.

TECHNICAL TRANSACTIONS 3/2018
CZASOPISMO TECHNICZNE 3/2018
ENVIRONMENTAL ENGINEERING
DOI: 10.4467/2353737XCT.18.047.8342
SUBMISSION OF THE FINAL VERSION: 5/2/2018

Barbara Tchórzewska-Cieślak (cbarbara@prz.edu.pl)
The Faculty of Civil and Environmental Engineering and Architecture, Rzeszow University of Technology

Characterization of risk function in the analysis and assessment of water supply systems safety

Charakterystyka funkcji ryzyka w analizach i ocenach bezpieczeństwa systemów zaoaptrzenia w wodę

Abstract

Safety is a term referring to lack of threat, guarantee of liquidation or minimizing threats. With regard to drinking water, consumer safety is understood as the probability of avoiding threat resulting from consuming water with quality that is incompatible with applicable normative or lack of water. However, an assumption that the measure of loss of safety is a risk function has become the paradigm. The classical definition of risk implies that it means the possibility of loss, harm with a certain probability. At the same time, there is a term of uncertainty which means lack of certainty and often refers to lack of safety. The main purpose of this paper is to present methods of analysis and risk assessment and the characterization of risk function in probabilistic and fuzzy (possibilistic) aspect.

Keywords: risk, safety, water supply system

Streszczenie

Bezpieczeństwo jest pojęciem odnoszącym się do braku zagrożenia, gwarancji likwidacji lub minimalizacji zagrożeń. W nawiązaniu do konsumentów wody do spożycia rozumiane jest jako prawdopodobieństwo uniknięcia zagrożenia, wynikającego ze spożycia wody o jakości niezgodnej z obowiązującym normatywem lub jej brakiem. Paradygmatem stało się natomiast przyjęcie, że miarą utraty bezpieczeństwa jest funkcja ryzyka. Z klasycznej definicji ryzyka wynika, że oznacza ono możliwość poniesienia straty, szkody z określonym prawdopodobieństwem. Równolegle funkcjonuje termin niepewność, który oznacza brak pewności i często odnosi się do braku bezpieczeństwa. Głównym celem pracy jest przedstawienie metod analizy i oceny ryzyka oraz charakterystyki funkcji ryzyka w aspekcie probabilistycznym i rozmytym (posybilistycznym).

Słowa kluczowe: ryzyko, bezpieczeństwo, system zaopatrzenia w wodę

1. Introduction

Safety is a term referring to lack of threat, guarantee of liquidation or minimizing threats, certainty that nothing bad will happen [1, 2, 5, 9, 11, 12]. With regard to drinking water consumer safety is understood as the probability of avoiding threat resulting from consuming water with quality that is incompatible with applicable normative or lack of water [5]. However, an assumption that the measure of safety loss is the risk function has become the paradigm [5–7, 10, 18]. A holistic approach to the analysis and assessment of the safety of collective water supply systems (CWSS) implies three basic safety control strategies:

- empirical safety control based on the results of statistical research of threats,
- evolutionary safety control based on an individual response to any threat,
- integrated safety control based on risk analysis of possible scenarios of developing dangerous situations.

The PN-ISO 31000 standard "Risk management. Principles and Guidelines" defines risk as a combination of sequences of events and the associated probability of risk occurrence. However, according to the standard ISO 9001: 2015, "Quality management systems" risk is defined as an influence of uncertainty on the functioning of systems or undertakings [16, 12].

In the social sense, risk is associated with threat. At the same time the term uncertainty functions which means lack of certainty and often refers to lack of safety [11, 12]. The decision-making process in terms of ensuring an acceptable level of operational safety for collective water supply systems (CWSS) and thus the effective choice of safety and protection measures for occurring threats is called risk management. At present, for water supply systems the European standards PN-EN 15975-1:2016: Water supply safety – Guidelines for crisis management and risk- Part 1: Crisis management and the PN-EN 15975-2: 2013-12: Water supply safety – Guidelines for crisis management and risk -Part 2: Risk management are applied.

The main purpose of this paper is to present the characteristics of the risk function in probabilistic and fuzzy (possibilistic) aspect.

2. Methods of analysis and risk assessment

Risk analysis is conducted in order to determine risk by estimating the probability of undesirable events and their consequences. The risk analysis should use historical knowledge of the operation of the system, analytical methods and experience. The purpose of risk analysis is to determine its value using the appropriate method [7, 10, 14]. Options of risk analysis approach [7, 17, 16, 18]:

- standard approach – assumes the introduction of standard safety procedures, regardless of the outcome of the risk analysis,
- expert approach – assumes the use of expert knowledge to protect those components of the system that are at high risk, detailed risk analysis – includes risk identification, assessment of threats and their causes, probable consequences and susceptibility to threat, for all the components of the system,

- mixed approach – depending on the complexity of the system one or more of the abovementioned methods is used.

Risk assessment is a comparison of determined values with risk acceptance criteria, which is the basis for safety analysis. At this stage it is very important to set the risk acceptability criteria so that they can be used in the decision making process regarding the operation of the system (e.g. repair or modernization). Such criteria should take into account the reliability requirements of the system (both quantitatively and qualitatively, in accordance with applicable legal regulations, and social and economic conditions) [15, 17, 18].

The risk assessment for CWSS with built-in safety systems requires the use of probabilistic safety assessment (PSA) predictive methods. The PSA modification is a quantitative method for risk analysis (QRA).

The choice of method depends primarily on the purpose of the risk analysis, the level of detail and the size of the data on system operation [6, 7, 9, 16, 19, 20]. Frequently, data are acquired from experts so knowledge of expert systems and in many cases of probabilistic methods is also necessary in risk analysis [8]. Starting risk analysis and assessment, you should know the "nature" of the risk and its basic properties.

Depending on the complexity of the problem, as well as the detailed purpose of risk analysis in the CWSS, the following methods of risk analysis and assessment are used [12–16, 18]:

- probabilistic methods belonging to quantitative methods for risk analysis (QRA) or probabilistic risk analysis (PRA), methods based on mathematical statistics and probability calculus,
- qualitative methods of risk analysis (QLRA),
- quantitative-qualitative methods for risk analysis, including;
  - matrix methods (two or multi-parameter),
  - fault tree analysis (FTA),
  - event tree analysis (ETA),
  - methods using the Markov and semi-Markov processes (Markov risk analysis -MRA),
- possibility methods – possibility risk analysis (PRA);
  - Bayesian methods (Bayes risk analysis – BRA),
  - fuzzy risk analysis (FRA),
  - neuro-fuzzy risk analysis (N-FRA),
- simulation methods using computerized hydraulic models and control, processing and data recording systems (e.g. SCADA), computer databases (e.g. Geographic Information System – GIS), and also the Monte Carlo simulation method, as well as genetic algorithms. They are a tool to support the risk analysis process [16, 18].

The QRA method allows all the significant risk management elements to be taken into account, such as identifying threat sources, determining representative emergency scenarios, estimating the probability of their occurrence, and associated losses and assessing risk based on current criteria.

An important problem is the choice of a risk reduction method that requires introducing some actions to reduce the likelihood of undesirable events (measures that reduce the likelihood of a dangerous event initiating a series of emergency events, measures that reduce

the likelihood of developing a dangerous situation into an emergency event) as well as measures reducing consequences.

Criteria for risk acceptability are especially important in defining the objectives of the safety management system in waterworks companies. The level of safety is usually acceptable if the CWSS meets the requirements set out in legal regulations and standards. The decision-making process in terms of ensuring an acceptable safety level for the operation of the CWSS, i.e. the effective choice of safety and protection measures for occurring threats, is called risk management [9–12, 17, 20].

3. Characteristics of the risk function

In risk analysis, priority is given to identifying potential threats and estimating the probable consequences [1, 10, 18]. Analysing the safety of the CWSS, we consider only those threats (undesirable events) that may have the most serious consequences. Those events are most often characterized by small or even very small probabilities of occurrence. Frequently, the lack of a sufficiently large statistical database of such events prevents the value of these probabilities from being correctly determined. In such cases it is necessary to rely on experts' knowledge. It should be borne in mind that in the theory of safety the principle "dealing with all sorts of threats, the ones with the least probability are characterized by the most serious consequences", is applied. You cannot omit unlikely events for which we do not have a database. Another approach, however, is applied in classical systems reliability analyses where all undesirable events (failures) are taken into account to determine the basic reliability characteristics that are a base for planning system modernization and repairs. The value of risk is described by the so-called risk function, representing the dependence of parameters: the probability of occurrence of emergency events posing a threat to the CWSS safety (P), consequences (losses, effects) of emergency (undesirable) event (C), under certain system operating conditions determining the so-called system vulnerability (V) to the development of the so-called representative scenario of emergency events $r = f(P, C, V)$ [4, 10, 15, 18].

The result of the risk analysis should be the expected value of certain losses (e.g. threat for health or lives of water consumers).

From a mathematical point of view, the expected value is determined by the dependence [3, 4]:

- for a continuous random variable:

$$E(C) = \int_0^{\infty} Cp(C)dc \tag{1}$$

- for a discrete variable:

$$E(C) = \sum_{i=1}^{N} c_i \cdot p_i = r(\mathrm{RES}) = r \tag{2}$$

where:

$R\,(\mathrm{RES}) = r$ – risk of a representative emergency scenario,
C – set of possible losses, $C = \{c_i\} = \{c_1, c_2, c_2 \ldots n\}$,
i – i-th value of loss or assumed loss range, n – maximum size or loss range.

The probability that the loss C_1 will occur as a result of an undesirable event is:

$$P(C = c_i) = p_i \tag{3}$$

In order to assess the risk acceptability levels (tolerable, controlled, unacceptable), the limit value of losses C_{gr} must be accepted.

Figure 1 shows the risk curve [3, 4] determined by the acceptable level of losses C_1 and the limit level of losses C_2.

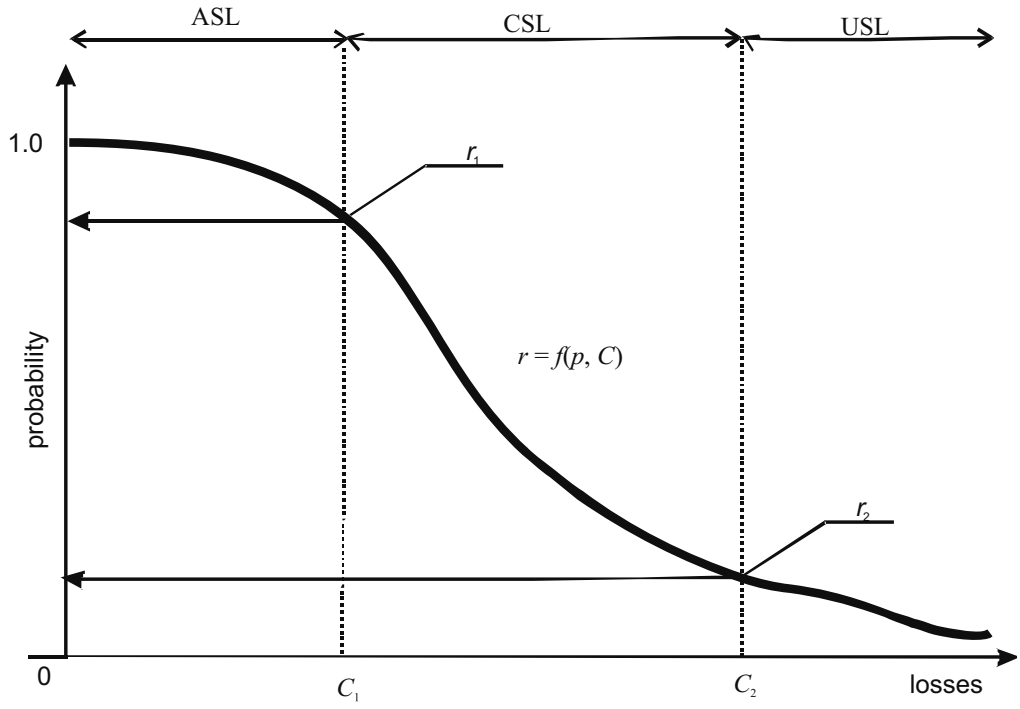


Fig. 1. The risk curve

The characteristic points of the curve are defined as follows:

- r_1 – producer's risk – it is the limit value of risk, under the given operating conditions (determining the vulnerability to threat parameter V), below which there is no real threat to water consumers, the risk is tolerable for water consumers – an acceptable safety level (ASL).
- r_2 – consumer's risk – it is the limit value of the risk, under the given operating conditions (determining the vulnerability to threat parameter V), above which there is a real risk to water consumers, the risk is unacceptable for water consumers – an unacceptable safety level (USL).
- the range from r_1 to r_2 means a controlled risk – a controlled safety level (CSL).

According to the risk curve shown in Figure 1, for the discrete random variable of losses $C = c_i$: the conditional loss expectation value (risk function) in the form [3, 4], is defined:

$$r = E(C|C \le C_{gr}) = \int_0^{\infty} Cp(C)dc \tag{4}$$

For particular levels of risk the formula (4) takes the form [3, 4]:

- tolerable risk r_T

$$r_T = E(C|(0 < C \le C_1) = \frac{\int_0^{C_1} C \cdot p(C)dc}{\int_0^{C_1} C \cdot p(C)dc} = \frac{\sum_{i=0}^{C_1} c_i \cdot p_i}{\sum p_i} \tag{5}$$

- controlled risk r_C

$$r_C = E(C|C_1 \le C \le C_2) \frac{\int_{C_1}^{C_2} C \cdot p(C)dc}{\int_{C_1}^{C_2} p(C)dc} = \frac{\sum_{C_1}^{C_2} c_i \cdot p_i}{\sum p_i} \tag{6}$$

- unacceptable risk r_U

$$r_U = E(C|C_2 \ge C_2) \frac{\int_{C_2}^{\infty} C \cdot p(C)dc}{\int_{C_2}^{\infty} p(C)dc} = \frac{\sum_{C_2} c_i \cdot p_i}{\sum p_i} \tag{7}$$

In this way, in the risk analysis we can determine the risk values for threats with a predetermined probability. Such an analysis of threats will allow for their hierarchization and proper assessment depending on the aim of risk analysis and assessment. The risk function defined in this way, based on the probability distribution of undesirable events and expected loss values, requires a clear definition of statistical characteristics, which involves the need to have a sufficient amount of reliable data (certainty).When data are burdened with the so-called "high uncertainty", a method that can be useful is fuzzy risk analysis (FRA) [8, 18]. Unlike in the classical set, the fuzzy set boundary is not precisely defined, but there is a smooth transition from a total lack of membership of the element to the set through its partial membership to its total membership. This smooth transition is defined by the so-called membership function μ_R, where R is a set of fuzzy numbers. In this way you can define a fuzzy risk function for which [8, 18]:

- the membership function assigns to each element x from the considered area (space) the value in the interval [0, 1] μ_R: $X \rightarrow [0, 1]$, which means that every element x from the space X belongs to the fuzzy set R with a certain degree of membership
- the risk fuzzy set R is defined as: $R = \{r_i, \mu_R(r_i)\}$, $i = \{i = T$ – tolerated risk, $i = C$ – controlled risk, $i = U$ – unacceptable risk.
- the values μ_R of the membership function μ_R are real numbers in the interval [0, 1],

The individual risk parameters are described by n linguistic variables (tolerated, controlled, unacceptable). Then fuzzy numbers are assigned to the individual linguistic assessments, which for the triangular membership function are defined as triple $x_j = (l_j, m_j, h_j)$, where: $j = 1, 2, \ldots n$.

The values of fuzzy numbers for the triangular membership function assigned to the individual linguistic variables are determined according to the following relations [8, 18]:

- for $j = 1$

$$x_j = \left(0;0;\frac{1}{n-1}\right) \tag{8}$$

- for $2 \leq j \leq n-1$

$$x_j = \left(\frac{j-2}{n-1};\frac{j-1}{n-1};\frac{j}{n-1}\right) \tag{9}$$

- for $j = n$

$$x_j = \left(\frac{n-2}{n-1};1;1\right) \tag{10}$$

where:

x_j – form of j-th fuzzy number ,
n – number of the linguistic variables describing a given parameter (risk),
j – consecutive number of linguistic variable, $j = 1, 2, \ldots n$.

For the linguistic variables that characterize the value of risk in a three-step scale, fuzzy numbers are determined according to the formulas (8), (9), (10), Tab. 1 [18].

Table 1. Fuzzy numbers for the linguistic variables describing risk

The linguistic variables	Fuzzy numbers
tolerated	(0,0; 0,0; 0,5)
controlled	(0,0; 0,5; 1,0)
unacceptable	(0, 5; 1,0; 1,0)

Exemplary fuzzy set of risk $R = \{r_i, \mu_R(r_i)\} = \{(r_T, 0{,}7), (r_K, 0{,}3), \{(r_N, 0{,}0\}$.

4. The application example

An analysis and assessment of the risk of failure for a water supply network for a city with 200 000 inhabitants was carried out. The analysis was performed in the following stages:

- to determine the type of water supply network. It was assumed that from the safety point of view, the analysis should be performed primarily for the main water network,
- to determine the failure intensity indicator based on statistical failure data,
- to determine the limit for the failure intensity indicator for water supply networks,

- to assess the risk of failure by traditional eq. 2.,
- to assess the risk of failure using a fuzzy model (using Matlab Fuzzy Toolbox) [18].

Table 2 presents the basic, statistical characteristics of the failure intensity indicator [18].

Table 2. Basic, statistical characteristics of failure intensity indicator for main water network determined from eight years of operation

Type of water network	Average network length [km]	Average frequency failure	Average failure intensity indicator λ [average frequency failure/km·a]	Standard deviation
main	49.5	13.1	0.26	0.05

Criteria for assessing the value of failure intensity indicator were based on [18]. For main networks, it is assumed that:

- low for $\lambda \leq 0.3$ [average frequency failure/km·a],
- medium for $\lambda = 0.3 \div 0.7$ [average frequency failure/km·a],
- high for $\lambda > 0.7$ [average frequency failure/km·a].

Based on these criteria, the analysed main water network has a low average intensity of failure throughout the year. The criteria for the loss parameter *C* were adopted in accordance with the principle that for a main network the potential value of losses due to a failure is high. According to formula (4) and curve 1, the value of expected loss $E(C)$ and risk value for the example was estimated as the tolerated risk. In order to verify the conducted analysis, a fuzzy risk analysis was also carried out. The input parameters for the developed model are [18]:

- x_1 – a variable characterizing the probability of occurring an adverse event (failure),
- x_2 – a variable characterizing losses incurred as a result of an adverse event (failure).

The output parameter is the risk value.

The fuzzy risk model is based on Mamdani's inference [18]. The general form of the rules is as follows: *If* x_1 *is* P_i *and* x_2 *is* C_j, *then* $y = r_{ij}$.

Modelling was performed using the Matlab (Fuzzy Toolbox) program. A detailed description of the model can be found in [18]. The probability values of failure occurrence are presented as a data set, whose components are fuzzy subsets described with linguistic variables: $P = \{P_1, P_2, P_3, P_4, P_5\}$. These variables are presented using the adopted triangular membership functions [18]. The values for the loss parameter are also shown as set ($j = 1, 2, 3, 4, 5$), whose elements are fuzzy subsets described using linguistic variables ($C = \{C_1, C_2, C_3, C_4, C_5\}$). These variables are presented by means of the adopted triangular membership functions [18]. Data entered for the analysed example for fuzzy risk assessment (Table 1) showed that the risk is at the controlled level.

5. Conclusions

- Ensuring safe use of public water supply requires the use of management methods that include risk analysis and assessment;
- It includes an assessment of the relationship between occurring threats and used safety and protection barriers;
- Comparison of the risk acceptance criteria should normally be made in the context of "the best judgement" from risk analysis not in the context of optimistic or pessimistic results of analysis;
- A risk assessment also depends on the knowledge and information available to the expert and the decision maker, hence the assessment of the uncertainty in the results will vary depending on the assessment of the people involved in it;
- The risk acceptance criteria should be updated as the knowledge about risk changes;
- Fuzzy risk modelling can be applied to an uncertain database;
- The safety of the CWSS should be considered from the perspective of the whole system, its impact on the environment and not from the perspective of the selected sub-systems or objects;
- A holistic safety management should be built into the system through constant analysis throughout its life cycle.

References

[1] Bajer J., *Zagrożenia dla ilości i jakości wody dostarczanej odbiorcom*, Wodociągi–Kanalizacja, 6(52), 2008, 40–43.

[2] Cichoń T., Królikowska J., *The impact of some external factors on the metrological properties of a water meter*, Technical Transactions, 1-Ś/2015, 3–11.

[3] Haimes Y.Y., *Risk analysis of fracture and failure*, Mat Reserch Innvation, 1998, 16–21.

[4] Haimes Y.Y., *Risk Modeling, Assessment, and Management*, Wiley, London 2015.

[5] Hrudey S.E., Hrudey E.J., *Safe drinking water. Lessons from recent outbreaks in affluent nations*, IWA Publishing, New York 2004.

[6] Iwanejko R., Lubowiecka T., *Analiza ryzyka w systemie zaopatrzenia w wodę studium zagrożeń*, Czasopismo Techniczne, 7-Ś/2003, 153–168.

[7] Kaplan S., Garrick B.J., *On the quantitative definition of risk. Risk Analysis*, 1(1), 1981, 11–27.

[8] Kleiner Y., Rajani B.B., Sadiq R., *Failure risk management of buried infrastructure using fuzzy-based techniques*. Journal of Water Supply Research and Technology, Aqua, 55(2), 2006, 81–94.

[9] Królikowska J., Królikowski A., *Niezawodność bezpieczeństwa sieci kanalizacyjnej*, Czasopismo Techniczne Środowisko, 1-Ś/2011, 101–111.

[10] Rak J., *Istota ryzyka w funkcjonowaniu systemu zaopatrzenia w wodę*, Oficyna Wydawnicza Politechniki Rzeszowskiej, Rzeszów 2000.

[11] Rak J., *Podstawy bezpieczeństwa systemów zaopatrzenia w wodę*, Wydawnictwo PAN – Komitet Inżynierii Środowiska, Lublin 2005.
[12] Rak J., *Bezpieczeństwo systemów zaopatrzenia w wodę*, Instytut Badań Systemowych PAN, Warszawa 2009.
[13] Rak J., Tchórzewska-Cieślak B., *Czynniki ryzyka w eksploatacji systemów zaopatrzenia w wodę*, Oficyna Wydawnicza Politechniki Rzeszowskiej, Rzeszów 2007.
[14] Rak J., Tchórzewska-Cieślak B., *Five – parametric matrix to estimate risk connected with water supply system operating*, Environment Protection Engineering, 2, 2006, 37–47.
[15] Rak J., Tchórzewska-Cieślak B., *Metody analizy i oceny ryzyka w systemie zaopatrzenia w wodę*, Oficyna Wydawnicza Politechniki Rzeszowskiej, Rzeszów 2005.
[16] Rak J, Tchórzewska-Cieślak B., *Ryzyko w eksploatacji systemów zbiorowego zaopatrzenia w wodę*, Wydawnictwo Saidel-Przywecki, 2013.
[17] Tchórzewska-Cieślak B., *Zarządzanie ryzykiem w ramach planów bezpieczeństwa wody*. Ochrona Środowiska, l4, 2009, 57-60.
[18] Tchórzewska-Cieślak B., *Metody analizy i oceny ryzyka awarii podsystemu dystrybucji wody*, Oficyna Wydawnicza Politechniki Rzeszowskiej, Rzeszów 2011.
[19] Wieczysty A. et al., *Podnoszenie niezawodności komunalnych systemów zaopatrzenia w wodę. „Metody oceny i podnoszenia niezawodności działania komunalnych systemów zaopatrzenia w wodę"*, ed. A. Wieczysty, Monografia Komitetu Inżynierii Środowiska PAN, 2, Kraków 2001.
[20] Zimoch I., *Niezawodność bezpieczeństwa jako priorytet współczesnego zarządzania eksploatacją sieci wodociągowej*, Międzynarodowa Konferencja Naukowo-Techniczna „Przedsiębiorstwa wodociągowo-kanalizacyjne w dobie współczesnych problemów", AQUA, Szczyrk 2009, 97–104.

TECHNICAL TRANSACTIONS 3/2018
CZASOPISMO TECHNICZNE 3/2018
MECHANICS
DOI: 10.4467/2353737XCT.18.048.8343
SUBMISSION OF THE FINAL VERSION: 15/2/2018

Sylwia Łagan
Aneta Liber-Kneć (aliber@pk.edu.pl)
Institute of Applied Mechanics, Faculty of Mechanical Engineering, Cracow University of Technology

The determination of mechanical properties of prosthetic liners through experimental and constitutive modelling approaches

Określenie właściwości mechanicznych linerów ortopedycznych z wykorzystaniem eksperymentu i modelowania konstytutywnego

Abstract

The aim of the study was the estimation of the ability of hyperelastic material models for the fitting of experimental data obtained in the tensile testing of silicone liners used in lower-limb prosthetics. Three groups of liners were analysed: I – silicone liner, II – part of the liner in which the silicone has a fabric reinforcement, III – silicone liner with an outer covering material. Both longitudinal and circumferential samples were taken. The Neo-Hookean, Mooney-Rivlin and Ogden parameters of constitutive models of hyperelastic materials were calculated.

Keywords: prosthetic liners, constitutive models, hyperelastic material, tensile test

Streszczenie

Celem badań była ocena przydatności modeli materiałów hipersprężystych do dopasowania danych doświadczalnych uzyskanych w próbie rozciągania dla silikonowych linerów ortopedycznych stosowanych w protezach dolnych. Przeanalizowano trzy grupy: I – liner silikonowy, II – liner silikonowy z wewnętrznym wzmocnieniem, III – liner silikonowy z zewnętrznym wzmocnieniem. Wyróżniono dwa kierunki pobrania próbek: podłużny i obwodowy. Zidentyfikowano parametry określonych funkcji modeli konstytutywnych materiałów hipersprężystych: Neo-Hookean'a, Mooney-Rivlin'a i Ogden'a.

Słowa kluczowe: linery protetyczne, modele konstytutywne, materiał hipersprężysty, próba rozciągania

1. Introduction

When observing the development of prosthetic technology, it can be seen that for a long time, a significant problem has been the biomechanical compatibility of a lower limb prosthetic with the residual limb - together, these constitute a new appendage after an amputation [11]. The direct contact of the patient's living tissues with the materials used for the comprising the prosthesis remains an important issue with regard to creating optimal solutions. A limb, as a new appendage which meets the load-bearing function requires characteristics aimed at eliminating, or at least limiting, the concentration of loads localised in sensitive areas. A prosthetic socket is one of the most important elements of the lower limb prosthesis. It transfers the load to the limb and therefore must meet the strength requirements with respect to the loads for which it is designed to transfer [3].

Today, the closest that is achieved to the ideal solution is to provide prosthetics with innovative sockets and silicone liners. Modifications to the liners has led to the development of a new type of suspension system of the prosthesis which provides comfortable attachment to the person's residual limb, and with care and soothing additions, this has a positively influence on the skin of the limb [2, 14]. Due to the good see above note adhesion of the silicone liner to the residual limb, this causes the reduction of friction presented on the skin; thus, it is often called the 'second skin'. Silicone liners may be used alone in the prosthesis as comfortable inserts or as an element of a total-surface bearing fit suspension for patients with a higher degree of mobility. There are three commonly used solutions using silicone liners in the suspension systems: shuttle lock, suction and vacuum [4, 6]. These solutions provide: comfort; the reduction of vertical movements of the limb in the socket, which stabilizes the walk; increases in the mobility of the person. The person's comfort depends on the precision of the socket and selection of an appropriate fit [2, 13]. Therefore, the modelling of problems of strength as an engineering tool is an important issue that enables the simulation of the behaviour of materials and structures in the selected load conditions. One of the methods of analysis for the selection of parameters is a static tensile test which provides the experimentally obtained material constants. This data provides the possibility to predict the stress characteristics at different levels of deformation. This can be a measure of the usage comfort evaluation of the product and the criterion of durability with regard to repeated attachment and removal. It can also be potentially useful for assessing prosthetic compatibility by providing quantitative information on the similarities and differences of these products.

The aim of this study was to evaluate selected mechanical characteristics obtained during the static tensile testing of longitudinal and circumferential samples taken from two types of the prosthetic liners (M and W). The results of tests were used to identify the parameters of the constitutive equations of selected hyperelastic models (the Neo-Hookean, the Mooney-Rivlin and the Ogden) to determine the material constants of the tested liners. The tested materials exhibit non-linearity and anisotropy; however, the assumption was made that its tensile behaviour can be predicted with the use of hyperelastic material models. From the point of view of clinical applications, the tested liners should show similar mechanical behaviour to skin. For skin,

which is an anisotropic material, hyperelastic material models are used in the modelling of the mechanical behaviour with assumption of its isotropy [16]. We can hypothesise that data from uniaxial stretching tests can be used in the simulation and modelling of the material parameters of orthopaedic liners – this is important from the point of view of obtaining a better match between the properties of the liners and the properties of the skin on the residual limbs. The usefulness of the hyperelastic material models for silicone liners was thus investigated in this study.

2. Material and methods

Samples were prepared from two models of orthopaedic liners produced by the leading manufacturers of prosthetic supplies, MediLiner FIRST® (M) and Willow Wood® Express AKLiner (W) (Fig. 1). The circumference dimension at a distance of 4 cm from the distal point is in the range of 32–47 cm, and the proximal dimension at the proximal end at a distance of 30 cm from distal point is in the range of 44–80 cm. The total length of the liners was approx. 370 mm.

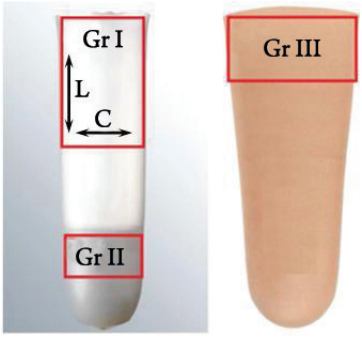


Fig. 1. The locations from where samples were taken: groups I and II from the liner M, group III from the liner (W); the directions of sampling have been marked (for all groups the same)

The first group (I) of samples was cut from the proximal part of the liner (M). The second group (II) of samples was cut from the distal part of the liner, where the silicone had a fabric reinforcement (M+). The third group (III) of samples was cut from the second type of liner (W). For all three groups (I–III), longitudinal (L) and the circumferential (C) samples were taken. The locations and orientations of samples taken are shown in Fig.1. All samples had the same dimensions: length 100 ± 1 mm and the width 10 ± 0.1 mm; however, these were of different thicknesses. The average thickness, the average cross-sectional area and the number of samples are presented as average values with standard deviations in Table 1.

For the prepared samples, the mechanical properties under static tension were determined with the use of the Instron 4465 testing machine with a force sensor of 5 kN. The samples were mounted using flat clamps and they were extended at a speed of 25 mm/min. The test was carried out in room conditions at a temperature of 22 ± 1°C. The measurement base of the samples was 60 ± 1 mm. The calculated values of maximum stress, stretch and energy of deformation are shown as the arithmetic average values with a standard deviation of (X ± SD).

Table 1. The characteristic parameters of samples

Liner type	M		M+		W	
Group of samples	G I		G II		G III	
Direction of taken	C $n = 4$	L $n = 4$	C $n = 3$	L $n = 3$	C $n = 4$	L $n = 4$
Average thickness (SD) [mm]	2.45 (0.11)	2.39 (0.04)	4.49 (0.13)	5.60 (0.05)	3.99 (0.02)	3.94 (0.03)
Average cross-sectional area [mm^2]	24.5	23.9	44.9	56.0	39.9	39.4

For all tested groups of samples, the analysis of experimental data was conducted by plotting the stress (σ) against the stretch (λ). The average experimental curves were then determined and on their basis, model analyses were carried out. In the process of fitting the tension experimental data with the use of constitutive hyperelastic laws, OriginPro 7.5 software was applied.

3. Constitutive material models

To describe the non-linear stress versus stretch relationship of hyperelastic materials, several constitutive models can be employed [1, 2]. These models assume that hyperelastic material behaviour may be characterised by the strain energy density function (W) expressed in terms of the three invariants of the strain Cauchy-Green tensor I_1, I_2, I_3, given as (1):

$$W = f(I_1, I_2, I_3) \tag{1}$$

Hyperelastic materials are special kinds of elastic material. For many engineering materials, including especially vulcanized elastomers, linear elastic models have poor accuracy in predicting the non-linear behaviour of the material.

The dependence of stress-strain for the material is referred to as isotropic non-linear elastic – it is independent on the strain rate. It is assumed that the material is also non-compressible. The response to the load applied on an unfilled silicone often shows the behaviour of a perfectly hyperelastic material. For filled elastomers, as well as for soft tissues, hyperelastic models are also used with the assumption of homogeneity of the material [7, 16].

The three invariants of the stretch tensor I_1, I_2, I_3, are given as (2), (3), (4):

$$I_1 = \lambda_1^2 + \lambda_2^2 + \lambda_3^2 \tag{2}$$

$$I_2 = \lambda_1^2\lambda_2^2 + \lambda_2^2\lambda_3^2 + \lambda_3^2\lambda_1^2 \tag{3}$$

$$I_3 = \lambda_1^2\lambda_2^2\lambda_3^2 \tag{4}$$

Thus, equation (1) can be given by formula (5), where C_{ijk} are the material constants:

$$W = \sum_{(i+j+k=1)}^{\infty} C_{ijk} (I_1 - 3)^i (I_2 - 3)^j (I_3 - 1)^k \tag{5}$$

Considering the conditions of the uniaxial tension of incompressible materials, the principal stretches λ_i are describe by the formula (6) and (7):

$$\lambda_1^2 \lambda_2^2 \lambda_3^2 = 1 \tag{6}$$

$$\lambda_1 = \lambda, \lambda_2 = \lambda_3 = \frac{1}{\sqrt{\lambda}} \tag{7}$$

therefore, the expression of the invariants of the stretch tensor become simplified and it is written as the principal stretch (8):

$$I_1 = \lambda^2 + 2\frac{1}{\lambda}, I_2 = 2\lambda + \frac{1}{\lambda^2}, I_3 = 1 \tag{8}$$

In the literature, hyperelastic models (Ogden, Money-Rivlin and Neo-Hookean) are widely described in order to determine the relationship of the stress-stretch of hyperelastic materials such as polymers, rubber-like materials or biological tissues [1, 7, 8, 12, 16].

Neo-Hookean material model

For the Neo-Hookean model, the strain energy density function is expressed as the principal stretch function (9), and the stresses are given by equation (10):

$$W = C_1 (I_1 - 3) \tag{9}$$

$$\sigma_{11} - \sigma_{33} = 2C_1 \left(\lambda^2 - \frac{1}{\lambda} \right) \tag{10}$$

Mooney-Rivlin material model

The strain energy density function for the Mooney-Rivlin model is expressed as the principal stretch function in the form (11):

$$W(\lambda_1, \lambda_2, \lambda_3) = C_1 (I_1 - 3) + C_2 (I_2 - 3) \tag{11}$$

therefore, for the uniaxial tensile, the stresses can be expressed as (12):

$$\sigma_{11} - \sigma_{33} = 2C_1 \left(\lambda^2 - \frac{1}{\lambda} \right) - 2C_2 \left(\frac{1}{\lambda_2} - \lambda \right) \tag{12}$$

The necessary and sufficient conditions for to be positive are given by the inequalities $C_1 \geq 0$; $C_2 \geq 0$ [16].

Ogden material model

The strain energy density function for the Ogden model is given as (13):

$$W(\lambda_1,\lambda_2,\lambda_3)=\frac{2\mu}{\alpha^2}\left(\lambda_1^{\alpha}+\lambda_2^{\alpha}+\lambda_3^{\alpha}-3\right) \tag{13}$$

where α and μ are the material constants: the shear modulus and the strain hardening exponent [2].

For the uniaxial tension conducted by the load in the direction of the long axis, the nominal stress can be given as (14):

$$\sigma_{11}-\sigma_{33}=\frac{2\mu}{\alpha}\left(\lambda^{\alpha-1}-\lambda^{-1(\alpha/2)}\right) \tag{14}$$

In order to compare the quality of the fit of a theoretical model to experimental data, the coefficient of determination R^2 is defined by formula (15):

$$R^2=\frac{\sum_{i=1}^{n}(\hat{y}_i-\overline{y})^2}{\sum_{i=1}^{n}(y_i-\overline{y})^2} \tag{15}$$

where y_i is the actual value of the variable, is the theoretical value of the variable on the basis of models, is the arithmetic mean average of the experimental value of the variable.

4. Results

On the basis of the tensile test, the stress-stretch curves for both circumferential and longitudinal samples for all three groups were obtained. For each group, the average stress-stretch curves were determined and the analysis of the results was performed. Circumferential samples taken from liners had lower values of maximum stress (Table 2). These values were: for group I (liner M), 0.15 MPa at a stretch of 3.08; for group II (liner M+ with reinforcement in the distal part), about 0.2 MPa at a stretch of 2.88; for group III (liner W) 0.16 MPa at stretch of 3.11. In the case of the longitudinal samples, the values were: for group I (M), the maximum stress was 0.16 MPa at a stretch of 3.23; for group II (M+) it was 0.23 MPa at a stretch of 2.36; in group III (W), the stress reached almost 1.54 MPa and the stretch was 1.57. The higher stress values were obtained for the knitted fabric reinforced silicone in comparison to the silicone without reinforcement for the longitudinal samples (Fig. 2).

Group I showed no significant differences in the stress values between the circumferential and longitudinal samples confirming the homogeneous composition of the liner without reinforcement in the proximal part. The addition of a knitted fabric reinforcement in the distal part of the liner had an influence on the tensile test results, raising the tensile stress and varying the slope of the stress-stretch curve. The orientation of the samples (whether circumferential or longitudinal) in groups with the fabric-reinforced silicon (groups II and III) also significantly influenced the stress-strain curves (Fig. 3). For the circumferential samples, the curve had a lower slope and a larger stretch. In group II, this was about 25% of the value relating to the longitudinal samples; in group III, it was over 40%.

Table 2. The characteristic mechanical parameters of tested liners

Liner type	**M**		**M+**		**W**	
Group of samples	G I		G II		G III	
Direction of taken	C	L	C	L	**C**	**L**
Maximum stress [MPa]	0.15 (0.02)	0.16 (0.01)	0.20 (0.003)	0.23 (0.01)	0.16 (0.04)	1.54 (0.54)
Stretch at maximum stress [–]	3.08 (0.17)	3.23 (0.05)	2.88 (0.11)	2.36 (0.02)	3.11 (0.14)	1.57 (0.54)
Energy of deformation [10^{-3}J] (SD)	252.61 (33.72)	297.62 (37.48)	547.81 (21.98)	400.46 (24.26)	380.93 (80.05)	711.34 (57.05)

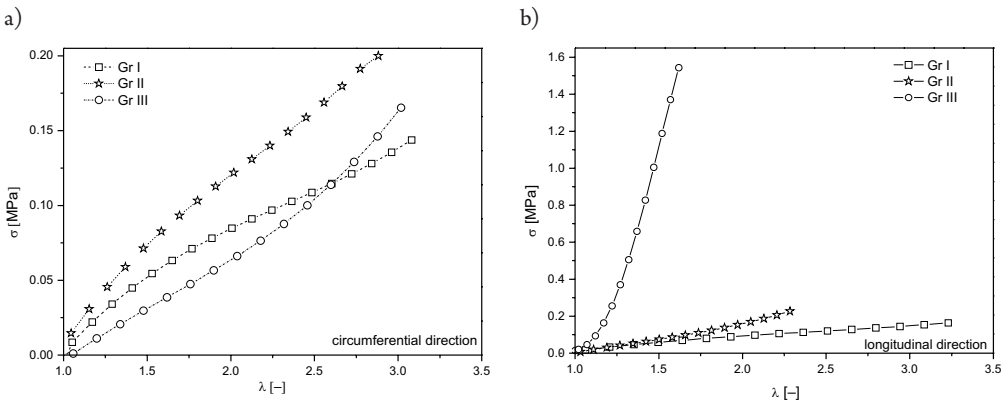


Fig. 2. Stress (σ) – stretch (λ) representative curves: a) circumferential, b) longitudinal

This confirmed the importance of the weaving direction in a fabric reinforcement for a silicone liner.

A highly accurate fit can be observed between the models and experimental stress when analysing the constitutive modelling (Fig. 4).

Analysis of the determined material constants in the constitutive modelling revealed strong agreement of values of the shear modulus (μ) obtained by the Mooney-Rivlin and the Ogden models in groups I and II for both circumferential and longitudinal samples and

in group III for circumferential samples (Tables 4 and 5). A strong fit between experimental data and these models is also confirmed in Figure 4. The exception was the shear modulus (μ) for group III in the longitudinal samples. For the Mooney-Rivlin model, it was a value of 0.67 MPa, and for the Ogden model, it was 0.397 MPa. The Neo-Hookean model was the least compatible with the experimental data – the constant μ was almost 2.5 times lower in the first group for both circumferential and longitudinal samples, and in the second group, for circumferential samples (Table 3). Knowing the requirements of orthopaedic surgeons and therapists, an orthopaedic device can be designed in such a way that will support the residual limbs, e.g. by controlled pressure.

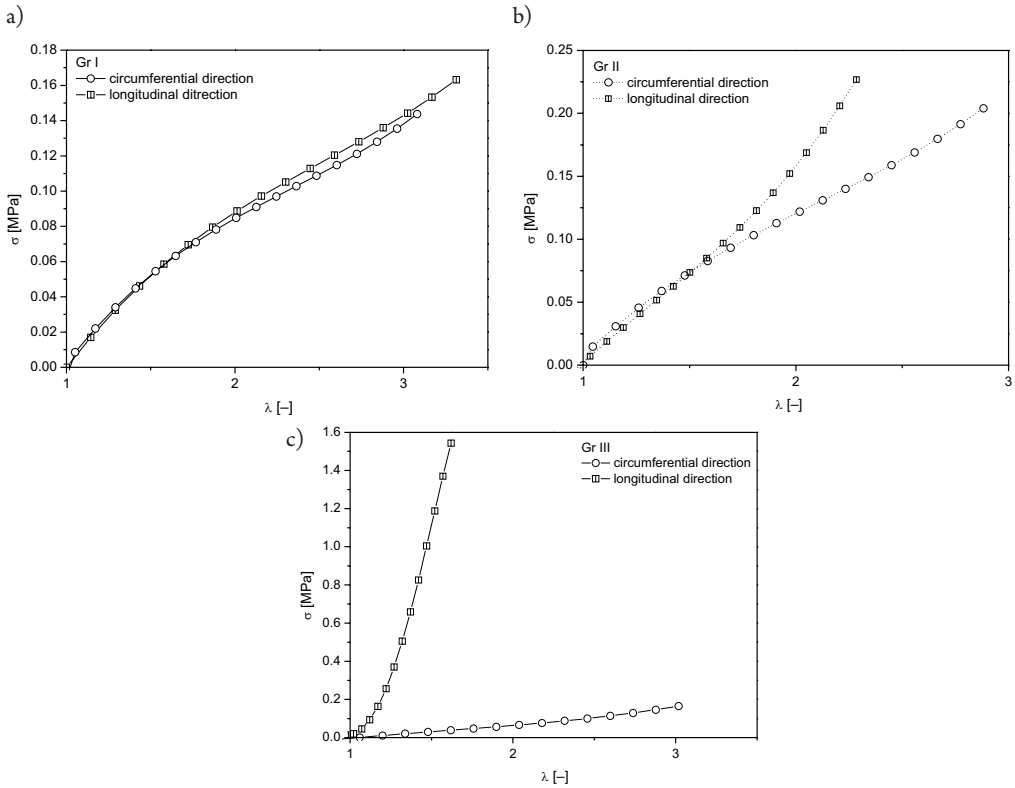


Fig. 3. Comparison of stress-stretch (σ-λ) curves for three groups of samples for both circumferential and longitudinal samples

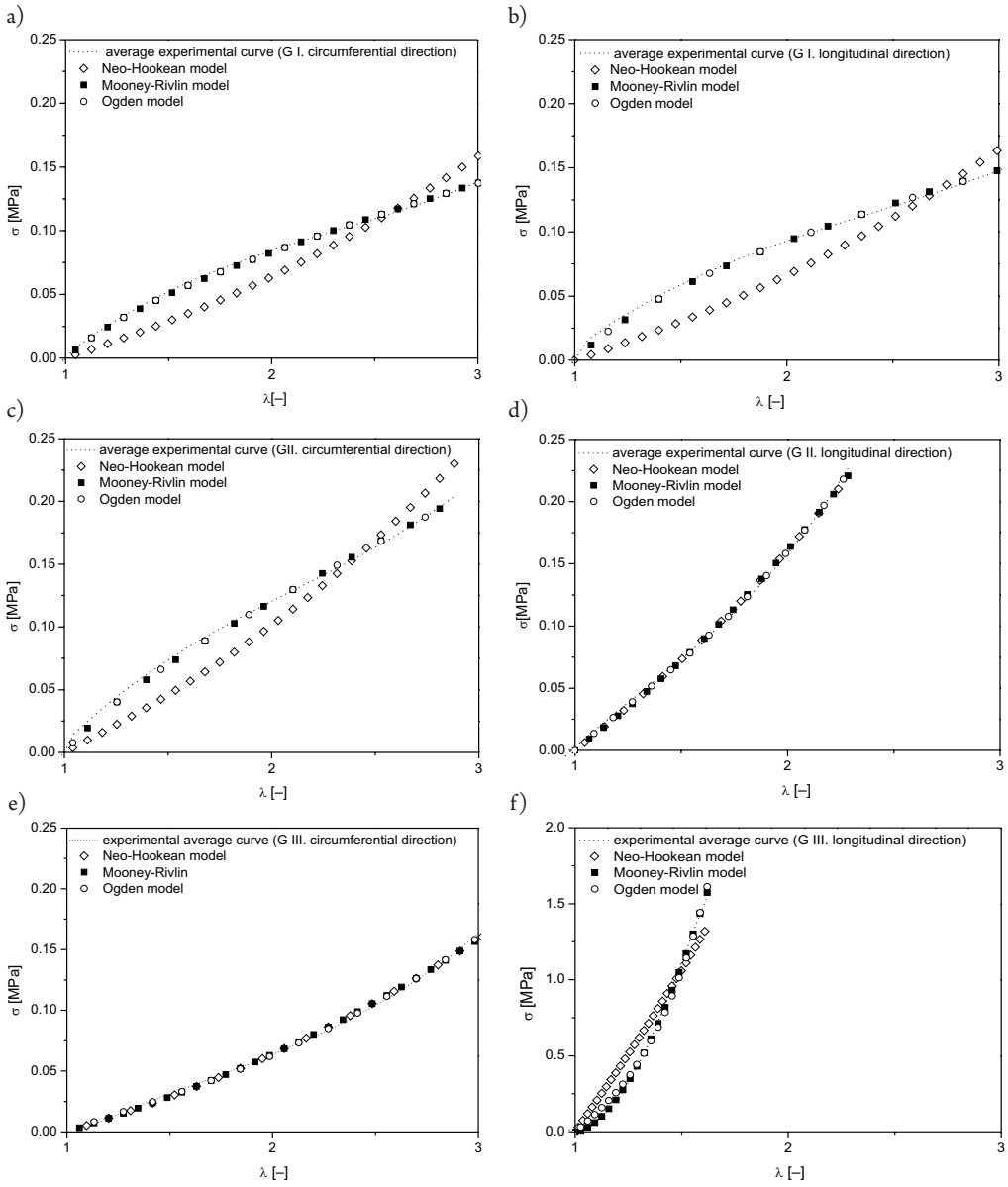


Fig. 4. Fitting to the tension data using three hyperelastic material models

Table 3. Parameter values for the Neo-Hookean model

Group of samples	Gr I (M)		Gr II (M+)		Gr III (W)	
Direction of samples taken	C	L	C	L	C	L
C_1 [MPa]	0.0091 ±0.0004	0.0095 ±0.0005	0.0145 ±0.0006	0.0230 ±0.0001	0.0091 ±0.0001	0.3354 ±0.1637
R^2	0.8394	0.8017	0.9133	0.9981	0.9975	0.9319
μ [MPa]	0.018	0.019	0.029	0.046	0.018	0.671

Table 4. Parameter values for the Mooney-Rivlin model

Group of samples	Gr I (M)		Gr II (M+)		Gr III (W)	
Direction of samples taken	C	L	C	L	C	L
C_1 [MPa]	0	0	0.0014 ±0.0004	0.0230 ±0.001	0.0081 ±0.0004	0.3354 ±0.1878
C_2 [MPa]	0.0238 ±0.0006	0.0259 ±0.0011	0.0320 ±0.0011	0	0.0021 ±0.0001	0
R^2	0.9986	0.9955	0.9979	0.9981	0.9981	0.9318
μ [MPa]	0.048	0.052	0.067	0.046	0.020	0.670

Table 5. Parameter values for the Ogden model

Group of samples	Gr I (M)		Gr II (M+)		Gr III (W)	
Direction of samples taken	C	L	C	L	C	L
μ [MPa]	0.049 ±0.006	0.056 ±0.001	0.067 ±0.001	0.053 ±0.001	0.024 ±0.001	0.397 ±0.030
α	1.9392 ±0.0245	1.8691 ±0.0378	2.1128 ±0.0382	3.4016 ±0.0346	3.1253 ±0.0466	6.3136 ±0.2704
R^2	0.9986	0.9975	0.9978	0.9993	0.9986	0.9923

5. Discussion

Eshragi et al. [4] investigated the pressure distribution in orthopaedic liners for the lower leg and demonstrated that the highest values occur from the anterior, followed by the posterior and the lateral and the medial (0.02–0.12 MPa). Our research indicates that the pressure generated remains in the safe range of loads for the strength of the liners (0.15–1.6 MPa).

Analysis of the results of the experiment demonstrates the non-linear nature of the load-deformation curves. As with the work of Ali et al. [1], similar values of the model constants for the Ogden and the Mooney-Rivlin models were obtained. Comparing the levels of model fit with the experimental data, a strong agreement of both the Mooney-Rivlin and the

Ogden models can be observed for all analysed cases across the whole range of considered stretch [5, 7]. By contrast, the Neo-Hookean model showed a strong fit in only two cases –for liner (M+) for longitudinal samples, and for liner (W) for circumferential samples. These samples of silicone were reinforced with fabric, for which the weaving direction was in line with the direction of the applied load. However, in the case of samples taken from liner (W) for longitudinal samples, a weaker fit was evident for the experimental curve for the Ogden model in the range of stretch between 1 and 1.25. Comparing the obtained values of the model parameters with the study conducted by Martins et al. [7], similar values of C_1 for the Noe-Hookean and C_1 and C_2 for the Mooney-Rivlin model for the group III in the longitudinal direction can be seen. The Ogden model with six parameters (C_1–C_6) makes it difficult to compare values. In the investigation of Ali et al. [1], the application of the constitutive models of hyperelastic materials for rubber was also made. However, they obtained nominal stress values of 25 MPa and nominal strain above 6; furthermore, a visible mismatch in the graph of the Neo-Hookean and the Mooney-Rivlin models can be seen. This makes the results difficult to discuss. However, the weakest fit of the Neo-Hookean model can be confirmed.

Sasso et al. [15] made a comparison and numerical simulations using a multi-parameter model for different compressions and the uniaxial and biaxial tensile load cases (the Ogden with six parameters, the Mooney-Rivlin with five parameters); again, observing the weakest fit with the Noe-Hookean and a better fit with the multi-parameter models.

Shergold et al. [16] used the Mooney-Rivlin and the Ogden models to analyse silicon Sil8800 and B452 as well as pig skin tissue by studying uniaxial stretching and compression at different strain rates (0.004, 0.4, 40. 4000/s^{-1}). They reported: μ = 0.4–7.5 and α of 1.2 for pig skin; μ = 0.4–2.8, α = 3, C_1 = 0.5 and C_2 = 0 for B452; μ = 2.1–8.0, α = 2.5, C_1 = 1.0 and C_2 = 0.9 for Sil8800; C_1 = 0.3 and C_2 = 0 for human skin. They indicated that the value of constant μ depends on the strain rate. It is considered that the value of α is related to the geometric evolution of the collagen network, and is independent on strain rate. In contrast, the resistance to rearrangement by deformation of the collagen fibers influence on value of μ and these deformation mechanisms are sensitive to the strain rate. The results were analysed in a context comparable to human skin parameters [16]. The substitute materials of human skin (rubber-like materials and pig skin tissues) are often used in medical engineering; therefore, knowing the basic mechanical properties of these substitutes is important for the comfort of the person and for strength requirements.

6. Conclusions

The mechanical properties of the rubber-like material can be described by constitutive models of hyperelastic materials, such as the Mooney-Rivlin and the Ogden models, at low, medium and large deformations, despite the fact that they are represented by the energy density function based on the principal strain invariants (the Mooney-Rivlin) and on three principal stretches (the Ogden). For two of the three analysed constitutive models

of hyperelastic materials, a strong fit of tension experimental data (correlation coefficient $R^2 > 0.99$) was obtained. For the Noe-Hookean model, a weaker fit was observed.

The procedure used to optimise the material parameters for three models has been successfully used for silicone materials designed for orthopaedic liners. Silicone liners are made from materials which exhibit a hyperelasticity with elastic stress-strain dependence for high deformation levels – they also exhibit a strong non-linear relationship. Hyperelastic material models can be used to determine the non-linear properties of the RTV-silicone and TPE liners with the assumption that material is homogenous, incompressible and isotropic.

References

[1] Ali A., Hosseini M., Sahari B.B., *A review and comparison on some rubber elasticity models*, Journal of Scientific and Industrial Research, 2010, 69, 495–500.

[2] Baars E.C.T., Geertzen J.H.B., *Literature review of the possible advantages of silicon liner socket use in trans-tibial prostheses*, Prosthetics and Orthotics International, 2005, 29(1), 27–37.

[3] Cavaco A., Ramalho A., Pais S., Duraes L., *Mechanical and structural characterization of tibial prosthetic interfaces before and after aging under simulated service conditions*, Journal of the Mechanical Behavior of Biomedical Materials, 2015, 43, 78–90.

[4] Eshraghi A., Osman N.A.A., Gholizadeh H., Ali S., Sævarsson S.K., Abas W.A.B.W., *An experimental study of the interface pressure profile during level walking of a new suspension system for lower limb amputees*, Clinical Biomechanics, 2013, 28, 55–60.

[5] Kim B., Lee S.B., Lee J.,Cho S., Park H., Yeom S., Park S.H., *A comparison among Neo-Hookean model. Mooney Rivlin model. and Ogden model for chloroprene rubber*, International Journal of Precision Engineering and Manufacturing, 2012, 13 (5), 759–764.

[6] Klute G.K., Glaister B.C., Berge J.S., *Prosthetic liners for lower limb amputees: A review of the literature*, Prosthetics and Orthotics International, 2010, 34(2), 146–153.

[7] Martins P., Natal Jorge R., Ferreira A., *A comparative study of several material models for prediction of hyperelastic properties: application to silicone-rubber and soft tissues*, Strain, 2006, 42, 135–147.

[8] Mooney M., *A theory of large elastic deformation,* Journal of Applied Physics, 1940, 11, 582–592.

[9] Ogden R.W., *Non-linear elastic deformations*, Dover Publications Inc., 1984.

[10] Ogden R.W., Saccomandi G., Sgura I., *Fitting hyperelastic models to experimental data*, Computational Mechanics, Springer-Verlag, 2004.

[11] Rajťúková. V., Michalíková. M., Bednarčíková. L., Balogová. A., Živčák. J,, *Biomechanics of lower limb prostheses*, Procedia Engineering, 2014, 96, 382–391.

[12] Rivlin R.S., *Large elastic deformations of isotropic materials IV. Further developments of the general theory*, Philosophical Transactions of the Royal Society London, 1948, A241, 379–397.

[13] Safari M.R., Meier M.R., *Systematic review of effects of current transtibial prosthetic socket designs—Part 1: Qualitative outcomes,* Journal of Rehabiltation Research and Development, 2015, 52(5), 491–508.
[14] Sanders J.E., Nicholson B.S., Zachariah S.G., Cassisi D.V., Karchin A., Fergason J.R., *Testing of elastomeric liners used in limb prosthetics: Classification of 15 products by mechanical performance,* Journal of Rehabilitation Research&Development, 2004, 41(2), 175–186.
[15] Sasso M., Palmieri G., Chiappini G., Amodio D., *Characterization of hyperelastic rubber-like materials by biaxial and uniaxial stretching tests based on optical methods,* Polymer Testing, 2008, 27, 995–1004.
[16] Shergold O.A., Fleck N.A., Radford D., *The uniaxial stress versus strain response of pig skin and silicone rubber at low and high strain rates,* International Journal of Impact Engineering, 2006, 32, 1384–1402.

TECHNICAL TRANSACTIONS 3/2018
CZASOPISMO TECHNICZNE 3/2018
MECHANICS
DOI: 10.4467/2353737XCT.18.049.8344
SUBMISSION OF THE FINAL VERSION: 19/2/2018

Andrzej Skrzyniowski
Andrzej Mruk (jendrek@mech.pk.edu.pl)
Faculty of Mechanical Engineering, Cracow University of Technology
Dorota Skrzyniowska (skdorota@pk.edu.pl)
Faculty of Environmental Engineering, Cracow University of Technology

Rotary smart car parking system

Samochodowy parking obrotowy

Abstract

Problems with parking individual cars, particularly in urban areas, were already experienced by the inhabitants of developed countries in the interwar period. These problems continue to appear, as the number of cars increases and there are fewer and fewer parking spaces. Therefore, solutions are sought after to ensure the largest possible number of parking spaces with the smallest occupied area, especially in large agglomerations of highly developed countries. The article presents one of the possible solutions – rotary car parks (The Rotary Automated Car Parking System; RACPS). The research results reffered to in the article are the effect of works conducted at the Faculty of Mechanical Engineering of the Cracow University of Technology in the field of modelling and optimisation of a rotary car park [5, 6, 8].

Keywords: rotary parking system, smart parking system, parking spot, Taguchi method, typology of parking lots, reducing model of rotary parking lots, RACPS, parking system solutions

Streszczenie

Kłopoty z parkowaniem samochodów indywidualnych szczególnie w obszarach miejskich mieli już w okresie międzywojennym mieszkańcy krajach rozwiniętych. I te problemy w dalszym ciągu się pojawiają, ponieważ ilość samochodów wzrasta, a miejsc do parkowania mamy coraz mniej. Dlatego poszukuje się rozwiązań, które mają zapewnić jak największą liczbę stanowisk parkingowych przy jak najmniejszej powierzchni zajmowanej, szczególnie w wielkich aglomeracjach krajów wysokorozwiniętych. W artykule przedstawiono jedno z możliwych rozwiązań, jakim są parkingi obrotowe (The Rotary Automated Car Parking System; RACPS). Wyniki badań przytoczone w artykule są efektem prowadzonych na Wydziale Mechanicznym Politechnice Krakowskiej prac z zakresu modelowania i optymalizacji obrotowego parkingu samochodowego [5, 6, 8].

Słowa kluczowe: obrotowy parking samochodowy, miejsca parkingowe, optymalny parking obrotowy, metoda Taguchi, typologia parkingów samochodowych, model redukcyjny parkingu obrotowego

1. Introduction

The first multi-level automated car park (APS) was built in Paris in 1905 at Ponthieu Street, located near Champs Elysées, with an elevator that transports cars to the appropriate level, where the staff parked their cars. Automated car parks appeared in the USA (New York) in the 1920s and remained popular there until today [24].

The most famous car parks of this type in Europe are the so-called car towers. An example is a car park in the Volkswagen plant in Wolfsburg, Germany, which was built from two concrete silos originally used for storing grain, creating the TurmFahrt glass towers - car parks intended for cars leaving the production line. Each can accommodate 400 cars. On their way to the silo, they go through an underground tunnel connecting the factory and the towers. Both silos can also be visited. Visitors sit in a glass gondola, i.e. an elevator, moving up and down through the centre of each building. The view of tens of cars appearing as if they were hanging in the air is on of the town's major attractions.

Currently, the largest automated car park in Europe is a car park for 1000 cars located in Denmark in Århus (DOKK1) at Hack Kampmanns Plads 6 [25].

2. Typology, a brief review of car park solutions and justification for choosing a rotary car park for analysis

Rotary car parks are one of the most advanced systems among all parking systems used so far [23].

For the purposes of the analysis, the following types of car parks have been considered:

PA – automatic car park (generally),
PP – automatic car park with inter-storey progressive movement of cars,
PO – automatic car park with inter-storey rotary movement of cars (rotary car park),
PTW – traditional multi–level car park, which is a closed construction facility only,
OPO – optimal rotary car park, as the preferred technical development.

The selection of OPO car parks for the above purposes (for the model: PO 10, i.e. 10 cars as the smallest for this type of unit) is justified by the following features resulting especially from a parametric comparison with the car park types listed above:

1) possibility of modular coaxial increase of parking capacity (including up to 120 and more, allowing comparative operations) owing to the adopted car entry/exit direction perpendicular to the drum axis,
2) no exhaust emissions to the environment, which, according to the calculation results, makes both OPO and all PA car parks **"an ecological profit"** in relation to the PTW car park,
3) approx. 50% energy saving in relation to the PP car park used most often due to the lower resistance in the drum rotational movement than in the platform progressive movements, which allows for approx. a twofold decrease of installed power,

4) approx. 60% shorter car access time compared to PP car park (90 instead of 230 seconds) due to shorter internal transport routes and a larger number of entry and exit stations,
5) building coverage ratio of 2.8 m^2/ number of parking spaces representing approx. 22% and approx. 86% space savings in relation to PP and PTW car parks,
6) building capacity ratio of 48 m^3/ number of parking spaces representing approx. 17% and approx. 47% space savings in relation to PP and PTW car parks,
7) unit investment cost of PLN 30 thousand/number of parking spaces representing approx. 60% and approx. 75% savings in expenditures concerning serial production in relation to PP and PTW car parks,
8) adoption of the following design solutions, which are advantages over other car parks listed:
   - levelling platforms with the use of two prismatic guides with the shape of a circular sector, which gives the entire openwork structure the appropriate lightness and the possibility of proper support of the drum on its both sides,
   - transfer of the drive on the drum side perimeter instead of on the axis with the use of a finger cycloidal chain gear, which, in addition to the dimensional minimisation of the drum bearing shafts, enables full self-locking of its movement,
   - lowered trench depth to approx. 1/3 of the ground height, i.e. to the parameter conditioning rational adoption of the non-returnable automatic direction (because of gravity) of parking and unparking (entry from one side and exit from the other side of the drum),
   - method of car entry and exit eliminating access roads and ways, thus eliminating land and space losses,
   - modular multiplication of the car park size, reducing the effects of possible system failures (then, one failure affects only one module).

3. Description of a rotary car park

The technical solution of a multi-level, automatic car park presented below is substantially ahead of the existing solutions and responds to the problems presented in the literature on car parks. Therefore, it [5]:

- lowers the investment cost of a parking space several times;
- reduces the value of its total area indicator;
- reduces the average time of parking and unparking vehicles;
- significantly increases the reliability of its operation
- reduces the amount of operating costs of the car park.

The suggested car park (Fig. 1) is a car park rotating around its horizontal axis. The car goes directly to the parking space; therefore, no internal access roads or ways are needed. This car park has no ceilings, load-bearing walls or pillars, and the gravitational system of moving vehicles (using the inclination of the transit area) does not require any mechanical devices or drives. It is simple, cheap and reliable [8, 26].

Depending on the car park size, the aforementioned platforms, together with hangers, rotatably mounted in the rotary drum sides, may be single or several-level, with two cars on each level. The size of the drum is also determined by the number of arms with platform hangers [8].

The drum, mounted on two supports, rotates, and the entry and exit of cars takes place near half of its height, in a plane perpendicular to its axis of rotation. Therefore, the car park is partly let into the ground, and partly protrudes above it [8].

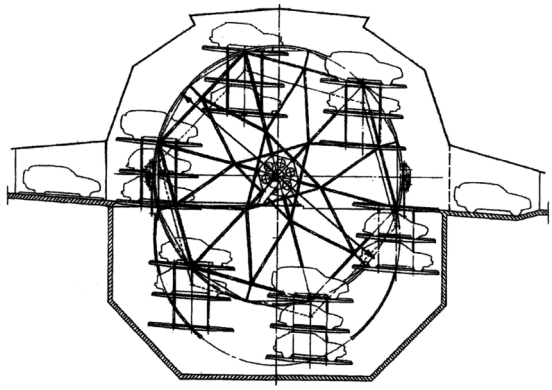


Fig. 1. Rotary car park – an example of a solution [26]

To verification of working proposed parking physical design in the scale of 1:33 (Fig. 2–5) was performed.

Fig. 2. Rotary car park – model, side view

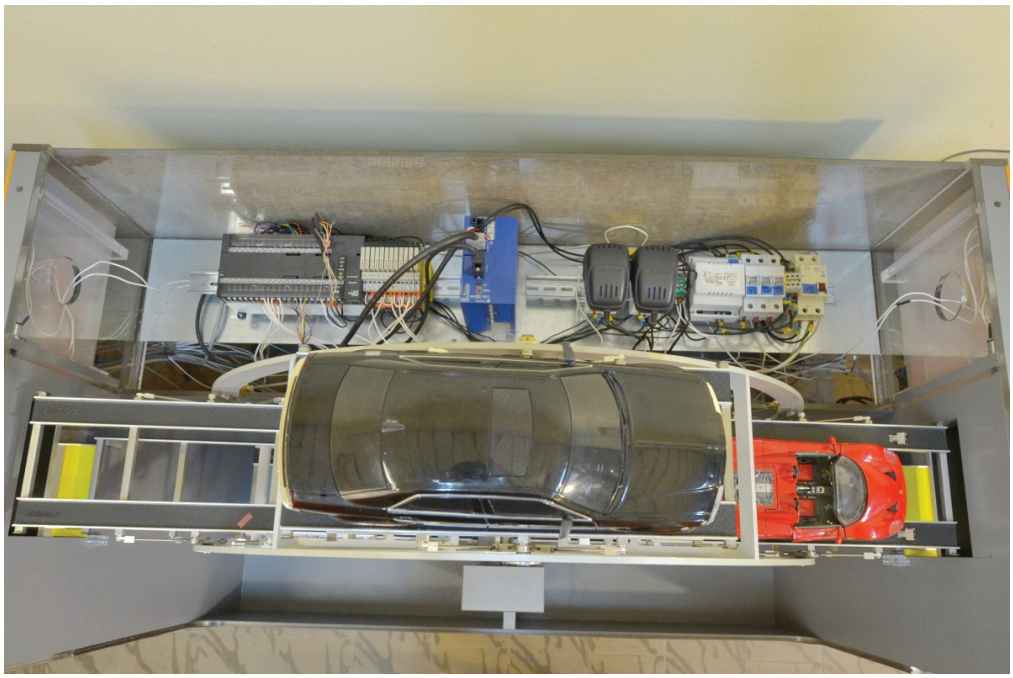


Fig. 3. Rotary car park – model, overhead view

Fig. 4. Rotary car park – model, car view on the platform

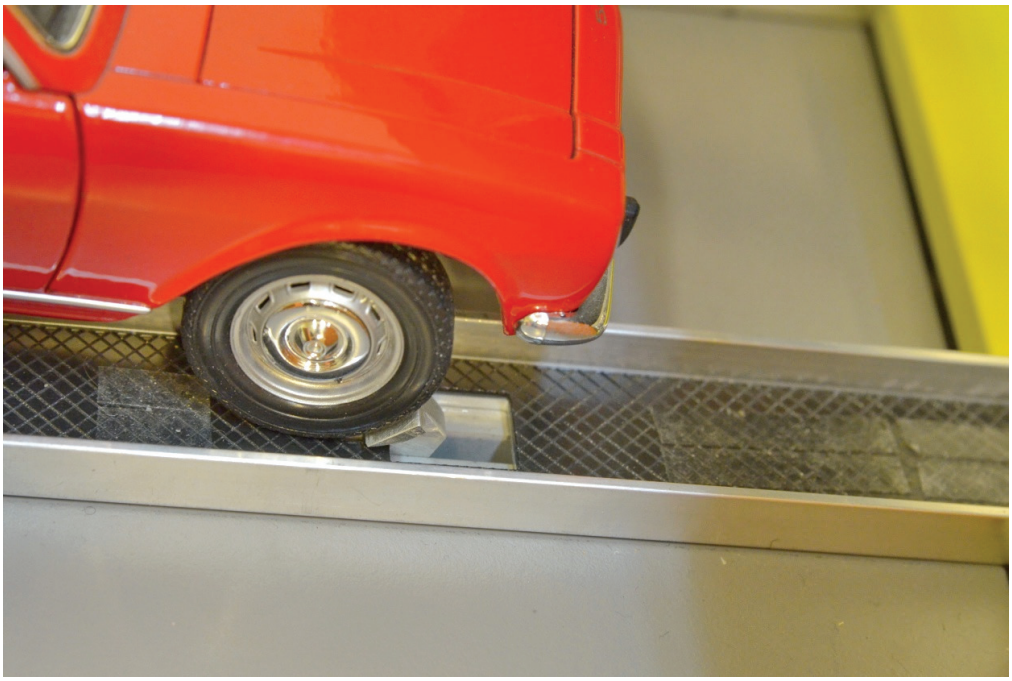


Fig. 5. Rotary car park – model, platform view

The operation of the model was navigated by a complied computer program. The carried-out study on the model allowed for dates to be accepted for the optimisation of the parking solution.

4. Design of a rotary car park model - computer simulation

The model design of the car park was conducted in two parts: control automation and automation of the mechanical part.

For the automation of the car park operation, entry – parking – exit of cars, the control diagram is adopted and modelled in Fig. 6.

To perform the automation tasks, a controller was selected, for example, for the task being developed: Omron sensor type CJ1, which gives the possibility of 2560 digital input and output signals, which in the case of the "extensive" car park model will allow to expand the controller capabilities.

The Omron-CX-Programmer V4.0 software package was used to write the programme to the controller.

CX-Programmer has all the tools for creating programmes in the ladder language.

The programme was divided into sections, which facilitated programming and made "movement" in the written programme easier.

Each section is responsible for individual elements of the car park movement control in accordance with the following list:

Section 1	Entry to platform 1
Section 2	Entry to platform 2
Section 3	Exit from platform 1
Section 4	Exit from platform 2
Section 5	Initial conditions
Section 6	Main programme
Section 7	Subprogramme calls
Section 8	Verification if there's a free space in the car park
Section 9	Calculation of weights
Section 10	Selecting a position on platform 1
Section 11	Selecting a position on platform 2
Section 12	Calculation of the car park rotation angle at the entry
Section 13	Car park rotation implementation
Section 14	Start of counting car parking time
Section 15	Selection of the entry position
Section 16	Calculation of the car park rotation angle at the exit
Section 17	Car park rotation implementation
Section 18	End of counting car parking time [5].

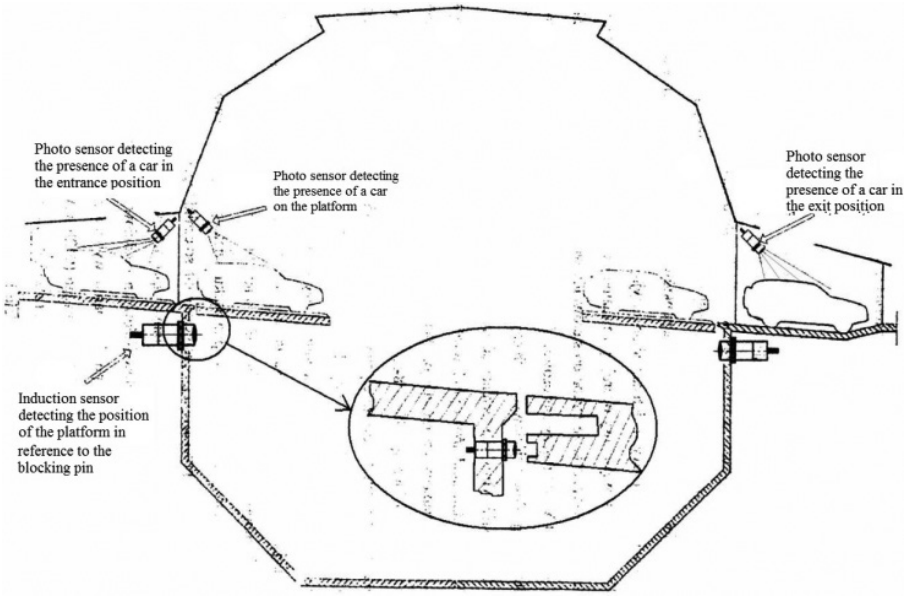


Fig. 6. Diagram of the arrangement of sensors for automatic control: entry – exit of a car [22]

Description of the optimisation programme

Pieces of the programme controlling operation of the car park responsible for searching for the optimal solution for the operation parameters: cost, road and efficiency, were presented in the development of the report on the research project [5].

The work covers an algorithm for determining "entry" and "exit" positions, subprogrammes checking the status of free spaces, subprogramme assigning "0" or "1" to each platform, "0" when there is no car on the platform and "1" when the platform is occupied, subprogramme counting "1". Knowing the number of cars and the distance of the platforms from the centre of the car park, the coordinates of the centre of gravity $x_{cc}y_{cc}$ are calculated.

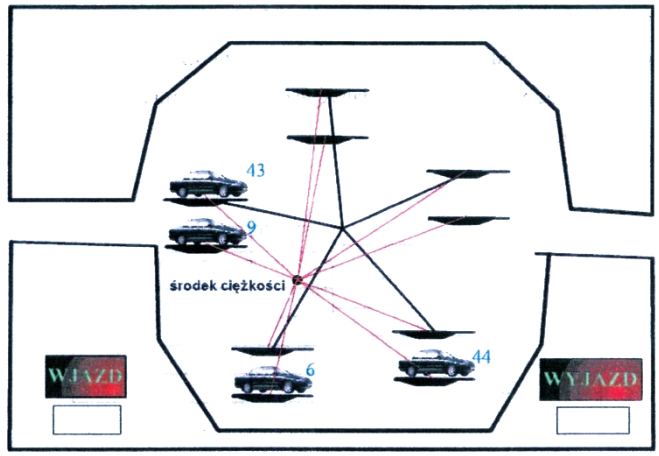


Fig. 7. Calculation of the centre of gravity

On this basis, in the next subprogramme, calculations are made to determine the length of sections from the centre of gravity to each of the individual platforms.

On the basis of these section lengths, the weight significance of each platform is determined later in the programme.

After calculating the length of the sections connecting the centre of gravity with each of the platforms, the programme sorts them in order from the largest to the smallest and assigns significance to each where 1 corresponds to the platform that is the nearest to the centre of gravity, while 10 to the platform the most distant to the centre of gravity. The next subprogramme assigns significance of the position, which informs about the distance (rotation angle) from the entry to all platforms, while the largest number represents the smallest angle and vice versa.

Therefore, the entry of the next car on an empty platform is determined by the subprogramme that calculates the ratio of the significance of the entry and the significance of the length of the section from the centre of gravity (gravity significance). Therefore, every next entering car will be parked in the position with the highest significance.

In this way, the number of the platform to which the car should enter is determined so that the centre of gravity is close to the centre of the car park.

The algorithm of rotary car park operation

A detailed description of the car park operation during entry and exit is presented in [5]. Below, there are only its most important elements.

The algorithm of rotary car park operation during entry

After reading the bar code from the identifier, it is checked whether there are free spaces in the car park. If so, the entry gate (bram1_gor) is opened and, at the same time, the wheel rotation causes the provision of the appropriate platform to the "entry" position.

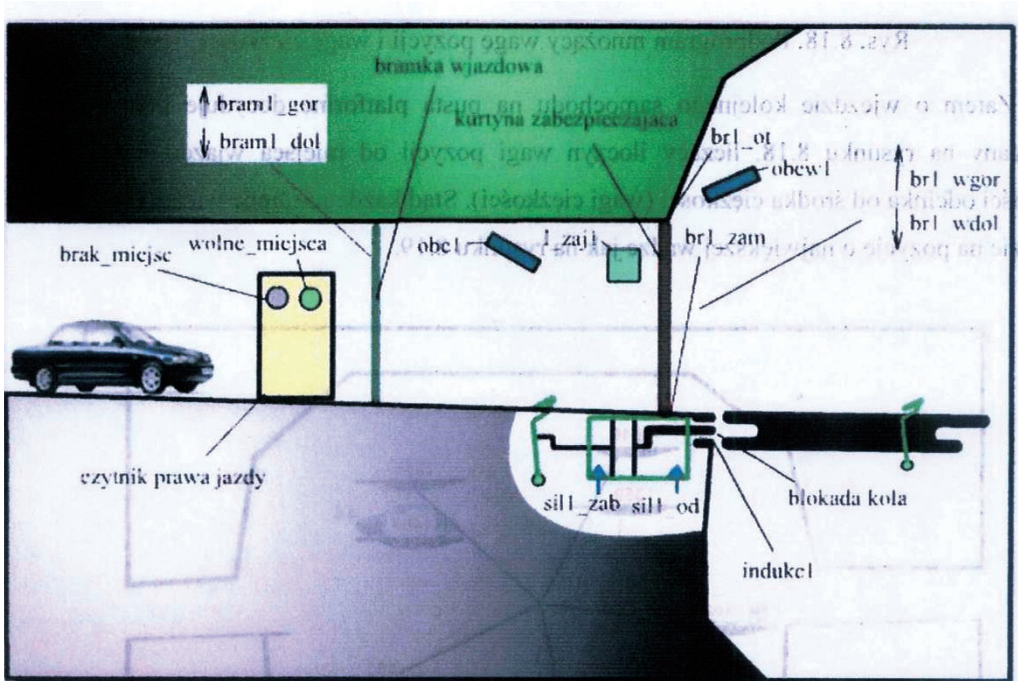


Fig. 8. Arrangement of entries and exits in the entry part [5]

After stopping the car behind the entry gate, the presence sensor (obe1) detects its presence. The driver re-inserts the identifier to the reader informing that he got out of the car. Then, the entry gate closes (bram1_dol). After, using an inductive sensor (indukc1), it is checked whether the wheel has stopped and the presence sensor (obew1) checks whether the place behind the curtain is free.

If both of these conditions are met, the security curtain (br1_wgor) is opened, the wheel is locked with a locking pin (sil1_zab) and the car is automatically shifted onto the provided platform. After the car has been detected by the presence sensor inside the car park (obew1) and waiting for two seconds, the security curtain (br1_wdol) closes and the wheel lock is released (sil1_od). The car is parked.

The algorithm of car park operation during exit

After reading the bar code from the identifier, it is checked whether the car is in the car park. If so, the car park is rotated, and an appropriate platform is set for the "entry" position. Then, using an inductive sensor (indukc3), it is checked whether the wheel has stopped and the presence sensor (obe3) checks whether the place in front of the curtain is free. If both of

these conditions are met, the security curtain (br3_wgor) is opened, the wheel is locked with the locking pin (sil3_zab) and the car is automatically shifted in front of the exit gate.

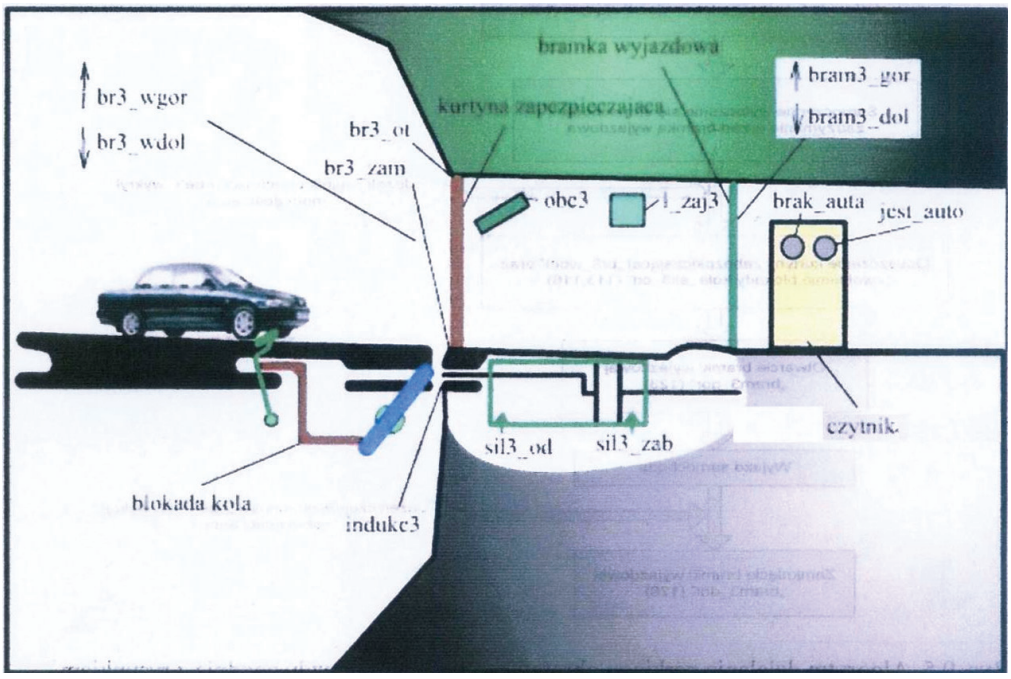


Fig. 9. Arrangement of entries and exits in the exit part [5]

After the car has been detected by the presence sensor (obe3) and waiting for two seconds, the security curtain (br3_wdol) closes and the wheel lock is released (sil3_od). Five seconds after the car is detected by the presence sensor (obe3), the exit gate opens (bram3_gor) and the driver drives away. After driving away, the exit gate closes (bram3_dol).

The process visualisation system has been made with the use of the InTouch 8.0 programme. InTouch consists of three main programmes: Application Manager, WindowMaker and WindowViever programmes[15].

The creation of a visualisation (simulating) window for the designed process was conducted using both ready-made templates (wizards) and special graphic objects drawn for this purpose, due to their lack in the programme libraries.

The screen window presents the "circle" of the car park with platforms and cars rotating with them. After entering a car on a platform, its identification number will appear along with it, which, like a car, is moving in a rotational movement and is inseparably attached to it until the car has left.

The analysis of the conducted simulation tests showed the correctness of the adopted optimisation methods. Simulation of the programme written in the InTouch environment for the adopted parameters of the rotary car park gave positive results. For several dozens of tests, the simulation of the "entry and exit" of cars into the car park did not show any errors. The

tests concerned only the rotation of the drum (car park circle) and torque optimisation. The correctness of the selection and operation of the entire programme for individual mechanisms was checked in real conditions on the car park prototype [5].

In Fig. 10–12, examples of screen views (graphic windows) designed to visualise the operation of a rotary car park are presented.

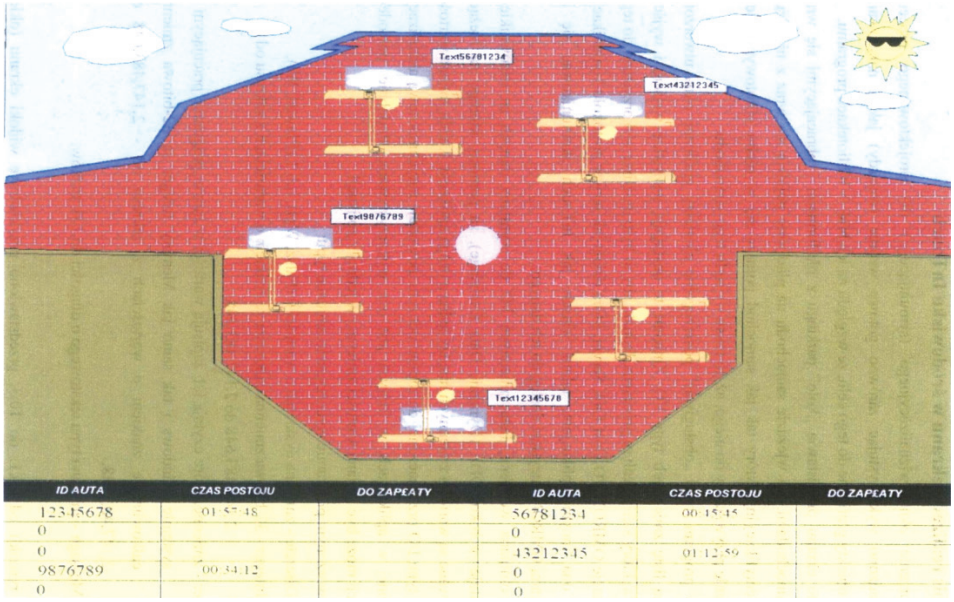


Fig. 10. Screen presenting the circle of the car park with platforms and cars placed on them [5]

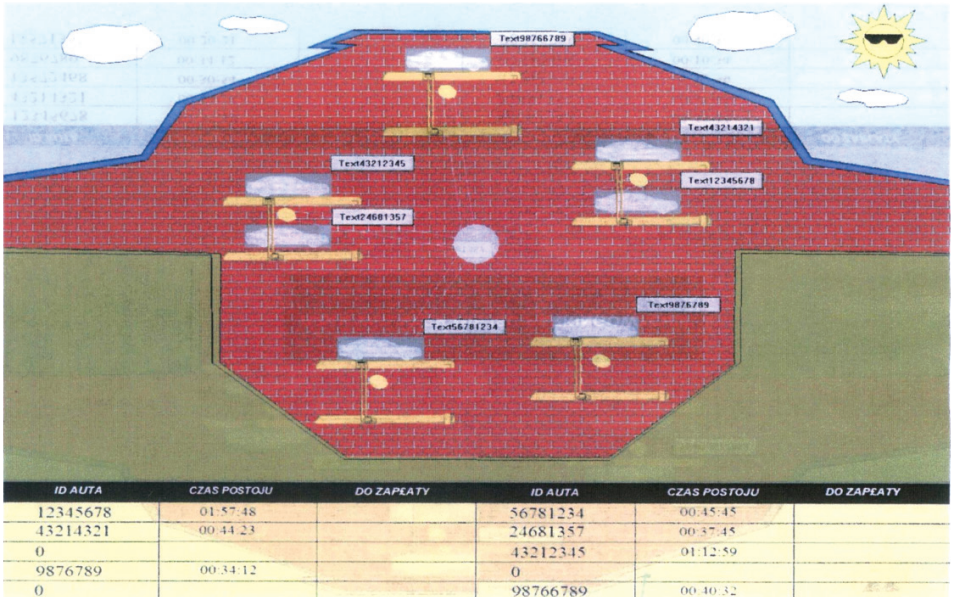
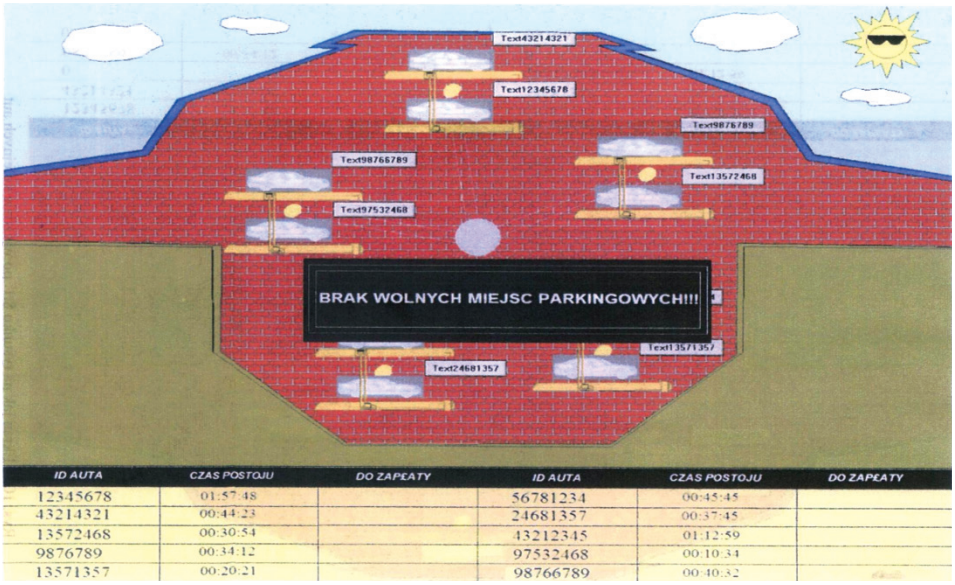


Fig. 11. Screen presenting the moment after the entry of other cars [5]



ID AUTA	CZAS POSTOJU	DO ZAPŁATY	ID AUTA	CZAS POSTOJU	DO ZAPŁATY
12345678	01:57:48		56781234	00:45:45	
43214321	00:44:23		24681357	00:37:45	
13572468	00:30:54		43212345	01:12:59	
9876789	00:34:12		97532468	00:10:34	
13571357	00:20:21		98766789	00:40:32	

Fig. 12. Screen presenting the "alarm" message informing about the lack of free parking spaces [5]

5. Structural development of the car park model and detailed development of the selected units

The PO – 10 rotary car park model was designed as a reduction model at scale of 1÷12.5 (due to the availability of car models in this size scale). The model's construction was preceded by computer tests of the nominal model as well as structural and operational analyses.

As a result of these works, the design of the model was created.

The model consists of three basic units: a mobile part with car parking platforms, an electronic part: a drum drive with rotational speed control, an electronic part allowing automatic entry to a parking space and exit after parking. The original, characteristic way of entry and exit from a parking space is the gravitational movement forced by the inclination of entry and exit channels, which constitute the parking space.

For utility reasons, individual parts of the model were made of various materials. Mechanical parts were made of metals; other parts were made of plastics with high resistance to impact and abrasion. The model is controlled by means of an electronic, ergonomically designed panel.

The functional and kinematic structure consists of entrance blockades, entry doors, a rotary drum with a unit of five "gondolas" with two platforms in each, that is, ten in total, exit blockades on platforms and exit doors. The drive structure of the model consists of four systems: drive of the entry blockades, drive of entry and exit doors, drive of a rotary drum and drive of platform exit blockades.

The main drive of the model was executed with an OMRON servo motor with a torque of 2.7 Nm and a maximum rotational speed of 2000 rpm. The motor has a built-in encoder with

a resolution of 10,000 pulses/rotation. It is possible to set the position with a resolution of up to 1/50,000 of rotation with the "lowering" belt transmission.

After switching on the power supply, the drum with "gondolas" can be positioned in relation to the limit switch built near the entry base position of the zero "gondola". In relation to entry and exit, the remaining "gondolas" are based in relation to this base position.

The locks of the right and left wheels in the entry and also in the exit are set by means of DC motors with transmissions. The lock operation is controlled by testing contacts of the limit switches indicating the status of locking and releasing.

The PLC exercises supervision over the model operation. The model is operated by a human via a control panel with a plastic keyboard.

The external appearance of the model was created as the last component of the whole. It is the result of the proper display of the construction and mechanical as well as an electric and electronic unit.

It consists of a base, a full back wall, a partially glazed front wall, a rounded glazed top cover adapted to the drum shape, entry ramp, exit ramp, visible inner walls and invisible reinforcements.

The model walls are made of 8mm thick metaplex. The visible parts are covered with HPL plastic, epoxy resin 1 mm thick in RAL 7037 colour. This material is highly resistant to mechanical shock and abrasion as well as chemically resistant. The "glazed" front wall and "shade" are made of a 4 mm colourless metaplex. If necessary, both parts are easily removable. Lids are similarly removable above the entry and exit ramps, and above the chamber at the back of the model holding drives and electronics. This gives convenient access to mechanisms intentionally hidden to achieve an aesthetic effect of the entire model.

Dimensions of the model with the base are; length – 2040 mm, width – 600 mm, height with base – 1600 mm + feet, the model itself – 1200 mm.

Tests of gravitational traffic of cars at the entry and exit of the car from the parking space have been conducted.

The "model" test showed sufficient inclination of the channels at the entry and exit 2°. It is an inclination at the same time ensuring stopping the car before and after entering the platform in the parking space.

Tests conducted under real conditions (with 5 cars of different weights) in the range of inclination of entry of 2–4° also showed a sufficient inclination of 2°.

Drum rotation of the model is driven by a gear with a toothed belt and an electronic option to change the rotation speed.

The drum rotation is forced by the drive unit, the motor of which, through the flexibly mounted "Nord" reducer, drives the chain gear, whose pin chain is multi-point-mounted to the external frame of the drum. The total drive ratio is $i = 2550$, which gives the drum speed $n = 0.57$ rpm.

In addition, the inclusion of a frequency inverter in cooperation with the electric motor makes it possible to slow down or speed up the drum rotation.

The analysis of the reduction model and, in particular, the reliability of the car's movement within the car park has created the need to solve the technical problems of blockades. The

problem of blockades of the car entry and exit from the parking platform has also been solved structurally.

The reduction model made can be used in further stages of work, to simulate tests defining the performance parameters of the future prototype.

6. Correlation of model parameter tests with computer tests – determining the durability and reliability of the model

The PO10 rotary car park model reflecting the adopted function of parking cars according to the developed parking technologies allowed for the correlation of results obtained by computer with the results of model tests. The structure and construction of the car park model were fully possible. Computer-developed construction elements as well as static and kinematic systems allowed for the creation of elements, units and assembly of a model fulfilling operating conditions in accordance with the assumed functions.

Also, a full correlation was obtained between computer model tests in the scope of production, assembly and operation of the electronic part controlling the reduction model.

Model tests in combination with the reduction model made, extended the scope of these tests to examples of constructional and electronic solutions that can be used in the design and implementation of the prototype. The results of these works also fully correlate with model tests.

Computer model tests have created the possibility of design and implementation of a fully functional reduction model, the implementation of the idea of a new method for parking cars.

The durability and reliability of the reduction model were tested in a 200-hour operation of the model in 90% operation cycles with "full load". The number of "entries" and "exits" was variable, on average 9 per hour, 1800 during the test period.

In the mechanisms of blockades (thresholds), no damage was found during the testing period - proper opening and closing of thresholds.

Car models also moved freely on entry and parking channels in accordance with the direction of their inclination. However, during the tests, it was found that a more favourable shape of the "channels" would be a section of the circle instead of a rectangle (more favourable tyre position, lower rolling resistance).

Durability, wear of individual elements of the car park model was only possible within the scope of the structure made under the test conditions. Elements and units of the car park did not show measurable wear [5].

7. Summary

The problem of "parking cars" is a significant problem of road traffic. The direction of improvement in the problem of parking cars is to create places, ways to leave the car for a certain time, parking it safely. Correction of this problem is to be conducted in the tests on new, functionally efficient methods of parking cars. The offered "rotary car park" is such a solution.

The patent, literature, Internet researches and own experience confirmed the novelty, the originality of the parking method and associated "rotary car park" with parking technology. Computer modelling and tests have been conducted in the field of construction and automation of the car park operation. Modelling of the mechanical part confirmed the possibilities and pointed to specific solutions of the construction units. Similar results were obtained in the tests of automatic, electronic control of the "car park" operation. The adopted methodology and research programmes were correctly and effectively selected, which also allowed to obtain the right correlation between computer tests and designs of structural and electronic solutions. The results of these works significantly created the construction possibilities for the implementation of the "car park" prototype.

The analysis of the model and its tests pointed to the positive features of this solution for parking cars.

The created reduction model made designed, in further stages of work, to simulate tests defining the performance parameters of the future prototype.

An important task of the model is also to be its illustrative, promotional and advertising role. Such a developed model can be the basis for the development of a prototype.

8. Conclusions

The basic conclusion from the presented work is the advisability of its continuation in the scope of further tests on the presented parking method and its technology, as well as the development of the scope of work for the performance of full technical documentation, creating and testing a prototype, implementation into production.

For the planned further work on the "rotary car park", it is necessary to perform full structural, mechanical and automation calculations of the car park as well as to develop and prepare technical documentation. A separate task is to document the "underground" part of the car park, which is within the scope of construction tests. Making cost estimates: technical documentation with the "underground" part of the car park, making a prototype of the PO-10 or PO-20 car park.

The implementation of the aforementioned objectives would allow for the introduction of a new, original scientific achievement for use, resulting in an improvement of the increasing inconvenience of parking cars.

References

[1] EN ISO 9001:2008 Quality management systems – Requirements.
[2] ISO 14000 Environmental management.
[3] ISO 14050:2009 Environmental management – Vocabulary.
[4] ISO 26800:2011, Ergonomics – General approach, principles and concepts.
[5] Stolarski B. et al., *Modelowanie i optymalizacja obrotowego parkingu samochodowego, projekt badawczy [Modelling and optimisation of rotary parking system, research project]*, Politechnika Krakowska, Kraków 2011.
[6] Borek J., Mruk A., Stolarski B., *Studium obrotowego parkingu samochodów [Study of a rotary car park]*, Mechanics and Mechanical Engineering, Vol. 9, No. 1/2005.
[7] https://www.youtube.com/watch?v=q3Kju88On7E (access: 7.11.2017).
[8] Borek J., Stolarski B., *Wielopoziomowy, wolno stojący, obrotowy parking samochodowy – studium [Multi-level, free-standing rotary car park – study]*, Unpublished work, Kraków 2005.
[9] Korzeniowski W., *Parkingi i garaże dla samochodów osobowych [Car parks and garages for passenger cars]*, WC0113, Warszawa 2000.
[10] Kwaśniewski J., *Programowalne sterowniki przemysłowe w systemie sterowania [Programmable industrial controllers in the control system]*, Author's edition, Kraków 1999.
[11] Urbanowicz H., *Napęd elektryczny maszyn roboczych [Electric drive of working machines]*, WNT, Warszawa 1979.
[12] Miedziński B., *Elektrotechnika. Podstawy i instalacje elektryczne. [Electrotechnology. Basics and electrical installations]*, PWN, Warszawa 2000.
[13] Przeździecki F., *Elektrotechnika i elektronika [Electrical and Electronic Engineering]*, PWN, Warszawa 1982.
[14] Niderliński A., *Systemy komputerowe automatyki przemysłowej [Computer systems for industrial automation]*, Vol. 2, WNT, Warszawa 1985.
[15] Wonderware, *InTouch – Podręcznik użytkownika [User Guide]*, Kraków 1997.
[16] Stolarski B., *Metody taksonomiczne w technologii samochodów [Taxonomic methods in car technology]*, Monograph No. 97, Politechnika Krakowska, Kraków 1990.
[17] Taguchi G., *Introduction to Quality Engineering, Designing Quality into Products and Process*, translation, Sekkei Sha, Tokio 1986.
[18] http://www.woehr.de/-internetowy (access: 7.11.2017).
[19] http://www.morski.com.pl/ – web portal of the Polish integrator of the Wöhr company (access: 7.11.2017).
[20] http://www.sels.com.pl/index2php (access: 7.11.2017).
[21] Kacznia J., Polański Ł., Diploma thesis.
[22] Skrzyniowski A., *Modelling and optimization of rotary parking system, Scientific Conference on Automotive Vehicles and Combustion Engines*, KONMOT, IOP Publishing, IOP Conf. Series: Materials Science and Engineering 148, 2016, 012051 doi:10.1088/1757-899X/148/1/012051.
[23] Stańko K., *Przegląd i charakterystyka systemów parkingowych [Overview and summary of parking systems, Mechanics]*, Czasopismo Techniczne 9-M/2012.

[24] http://blog.immobile.com.pl/historia-rozwoju-automatycznych-systemow-parkowania-na-swiecie/ (access: 9.03.2017)
[25] http://dokk1-parkering.dk/en/plant.aspx (access: 20.03.2018).
[26] Borek J., *Wielopoziomowy, wolnostojący parking samochodowy (parking obrotowy) [Multi-Level Car Park (Rotary Car Park)]*, Zeszyty Naukowo-Techniczne Stowarzyszenia Inżynierów i Techników Komunikacji w Krakowie [Research and Technical Papers of Polish Association for Transportation Engineers in Cracow], Conference Materials, 2007.