MTK 18.

MŰSZAKI TUDOMÁNYOS KÖZLEMÉNYEK

18.

TUDOMÁNYOS BIZOTTSÁG/LEKTOROK
SCIENTIFIC ADVISORY BOARD/PEER REVIEWERS

ISSN 2393 – 1280

Erdélyi Múzeum-Egyesület Műszaki Tudományok Szakosztálya
Transylvanian Museum Society – Department of Technical Sciences

MŰSZAKI TUDOMÁNYOS KÖZLEMÉNYEK

18.

Szerkesztette / Edited by
BITAY ENIKŐ – MÁTÉ MÁRTON

ERDÉLYI MÚZEUM-EGYESÜLET
Kolozsvár
2023

A kötet megjelenését támogatta a Kolozsvár Polgármesteri Hivatala és Helyi Tanácsa,
a Kolozs Megyei Tanács, a Communitas Alapítvány,
a Magyar Tudományos Akadémia, a Bethlen Gábor Alapkezelő Zrt.,
az EME Műszaki Tudományok Szakosztálya

The publication of this volume was supported by the Cluj-Napoca Mayor's Office and
Local Council, the County Council of Cluj, by the Communitas Foundation,
the Hungarian Academy of Sciences, by the Bethlen Gábor Fund,
by the TMS – Department of Engineering Sciences

Kiadó/Publisher: Erdélyi Múzeum-Egyesület
Felelős kiadó/Responsible Publisher: Biró Annamária
Szerkesztette/Edited by: Bitay Enikő, Máté Márton
Olvasószerkesztő/Proofreader: András Zselyke (magyar), David Speight (English)
Műszaki szerkesztő/DTP: Szilágyi Júlia
Borítóterv/Cover: Könczey Elemér

Nyomdai munkálatok/Printing-work
F&F International Kft. Kiadó és Nyomda, Gyergyószentmiklós
Ügyvezető igazgató/Manager: Ambrus Enikő
Tel./Fax: +40-266-364171

online elérhető/online available at:
https://eme.ro/publication-hu/mtk/mtk-main.htm
DOI: 10.33895/mtk-2023.18

TARTALOM

CONTENT

Műszaki Tudományos Közlemények vol. 18. (2023) 01–04.
DOI: Hungarian: https://doi.org/10.33895/mtk-2023.18.01
English: https://doi.org/10.33894/mtk-2023.18.01

RESULTS OF RESEARCH RELATED TO REDUCING ENERGY CONSUMPTION IN THE ROCK WINNING PROCESS

József ANDRÁS,[1] József KOVÁCS[2]

[1] *University of Petroșani, Faculty of Mechanical and Electrical Engineering, Department of Mechanical, Industrial and Transportation Engineering. Petroșani, Romania, iosif.andras@gmail.com*

[2] *University of Petroșani, Faculty of Mechanical and Electrical Engineering, Department of Mechanical, Industrial and Transportation Engineering. Petroșani, Romania, kovacsi@mail.com*

Abstract

Energy consumption is the main parameter describing the efficiency of the rock winning process in mechanical excavation. Because the rock winning process is a very complex one, the strength properties of rocks prescribed by classical rock mechanics is not sufficient: Knowledge of cutting forces values, and of the shape and volume of removed chips – based on tool geometry and cutting head kinematics – has significant importance in the optimal design of winning machines, and in improving their operating parameters. The paper deals with some results in research performed at the University of Petroșani on these issues, both in theoretical findings and experimental assays in laboratory and on site.

Keywords*: rock cutting, energy, mechanical excavation, mining machinery.*

1. Foreword

Establishing the characteristics of the extraction of various rocks or useful minerals by excavators (**Fig. 1**), shearer-loader (**Fig. 2**), roadheader (**Fig. 3**) etc. is an important task that leads to the understanding of the phenomenon of rock cutting, the determination of the geometric and technological parameters of the excavation tool, the parameters of the removed rock chips and the relationships between the parameters of the process.

Knowing the laws of interaction between parameters, in practice it is possible to develop such excavation machines and technology that require minimal specific energy consumption, using the maximum efficiency of the machine, in different specific cases.

Mechanical excavation consists of the decomposition of the rock from its natural state in the rock mass.

The chips are separated from the rock by means of a sharp working tool, with suitable geometry and strength (**Fig. 4**) .

Fig. 1. *Bucket Wheel Excavator.*

Fig. 2. *Shearer loader.*

Fig. 3. *Roadheader.*

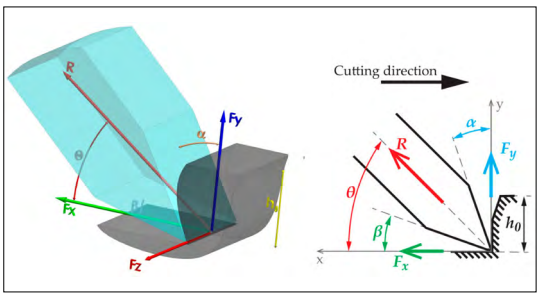


Fig. 4. *Geometric characteristics of the cutting tool (α – face angle, β – back angle, θ – cutting angle, h_0 – depth of cut) and forces acting on it (F_x – cutting force, F_y – pushing force).*

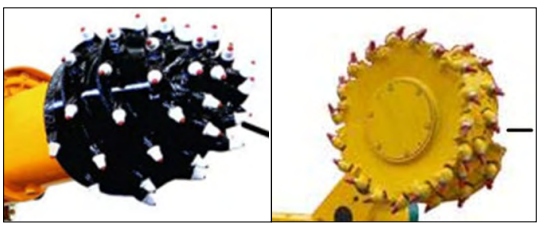


Fig. 5. *Different types of cutting head units.*

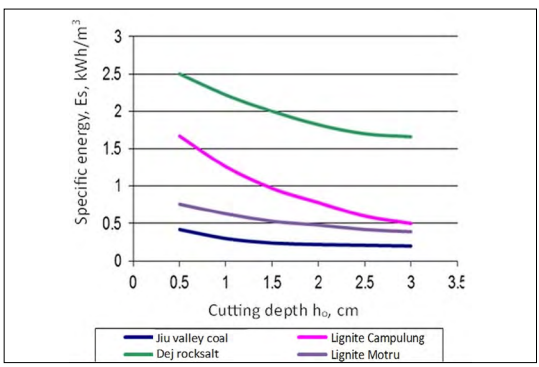


Fig. 6. *Specific energy consumption for the studied rocks as a function of the cutting depth.*

The working tools are placed according to an appropriate placement system on a rotary cutting head unit (**Fig. 5**), so that the machine can decompose the entire rock mass due to the spatial kinematics of the cutting head movement performed by a manipulator (arm).

In the case of coal, rock salt or other beneficial minerals and rocks, it is not enough to determine the strength of these by various conventional rock mechanics tests (compressive, tensile, shear strength, etc.), since excavation is a very complicated process that depends on the influence of many factors.

Its characteristics include: forces acting on the working equipment (drag bit, demolition tooth, cutting edge, etc.) (**Fig. 4**), specific cutting force, specific energy demand, required propulsion power, etc.

The cutting forces acting on the cutting device (tooth or drag bit), mainly the tangential cutting force and the normal force (**Fig. 4**), are the basic parameters used to calculate the torque of the cutting head, engine power and rock excavation efficiency.

On the other hand, the specific cutting energy is usually used to assess the cuttability of rock, which is one of the most significant parameters used both for performance evaluation and for evaluating the efficiency of excavation systems.

2. Analysis of the energy demand of various rocks

Fig. 6 shows the curves of specific energy consumption for the materials already mentioned, as a function of the cutting depth, which were plotted based on the results of laboratory measurements.

These curves are hyperbola-like, similar in shape, but significantly different in values.

It can be observed that in the case of rock salt the specific energy consumption is 1.5...2.5 kWh/m^3, in the case of the hardest domestic lignite it is 0.5...1.5 kWh/m^3, but in the case of hard coal it is only 0.25...0.5 kWh/m^3.

Since the difference between the analysed rocks is not so great as far as cutting factors are concerned, harder and softer materials were also taken into account as examples.

The previous frame was extended with clay, softer lignite and sandstone. Since the differences are already greater, we were forced to use the logarithmic scale, as can be seen in **Fig. 7.**

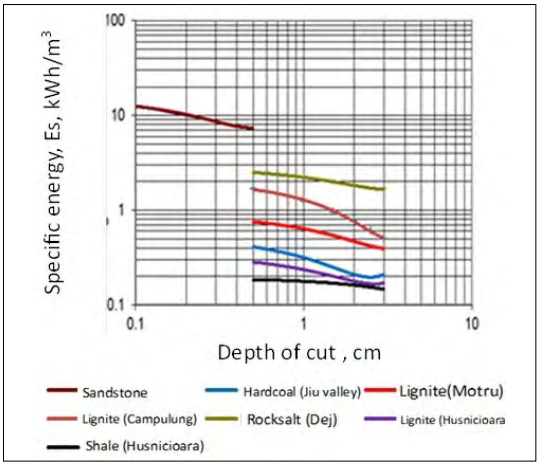


Fig. 7. *Specific energy consumption for the studied rocks as a function of the cutting depth.*

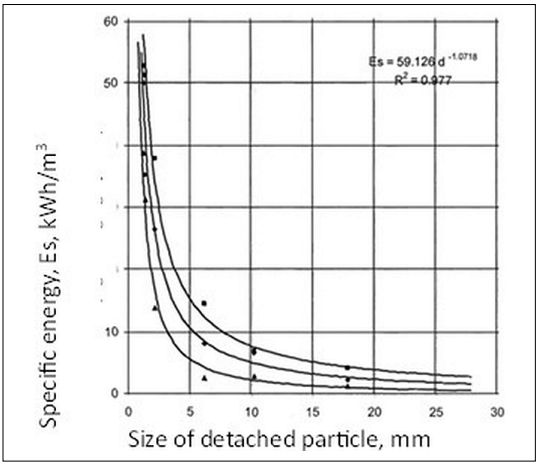


Fig. 8. *Dependence of specific energy on particle size.*

Based on the curves shown in **Fig. 7** the specific energy consumption was compared for the materials already listed.

Here, too, the order from clay to sandstone stands out, as the curves of rock salt and carbons are in between.

If we make the comparison, in terms of specific energy demand, it can be noticed that sandstone has a specific energy demand of 7...13 kWh/m^3, rock salt 1.8...2.6 kWh/m^3, hard lignite 0.5...1.8 kWh/m^3, coal 0.25...0.5 kWh/m^3, while clay has a specific energy demand of only 0.18...0.2 kWh/m^3.

It follows from the same source that the mechanized production of salt from Dej requires 4...5 times less energy than sandstone, which cannot be obtained on an industrial scale by cutting, but is 1.5...7 times more for coals and 10...13 times more for clay.

One of the main parameters describing the excavation process is the specific energy demand, which is a measure of the energy intensity of the ongoing process.

The author of paper **[3]** used more than 11,000 rock extraction experiments of extraction with blast and use of different machines (cutting machines, tunnel boring machines, shearer loaders, drilling rigs, etc.) establishing the dependence of specific energy consumption and removed particle size.

The measured or calculated, specific energy consumption E_s and the average dimension of the detached materia d the experimental point cloud is located within a band between two curves that approach the equilateral hyperbola. (**Fig. 8**), in which the function $E_s = f(d)$ is plotted.

Its descriptive regressive equation is:

$$E_s = 59{,}126\ d^{-1{,}0718}\ kWh/t \qquad (1)$$

As it can be seen, the exponent is -1.0718, which characterizes a quasi-equilateral hyperbola. On the other hand, it means that the product of specific energy consumption by the size of the particles (chips) produced is constant, regardless of the process used, i.e.:

$$E_s \cdot d \approx constant \qquad (2)$$

In previous papers **[1]** and **[2]** it has been proved that this invariant corresponds to the specific cutting force.

Thus, the relationship between the two essential parameters – specific energy consumption and specific cutting force – and based on this, the specific energy demand, can be calculated based on the specific cutting force, which can be determined by experimental means.

3. Conclusions

Energy demand is the main parameter that describes the efficiency of the rock winning process in mechanical extraction.

Knowledge of its parameters and of those influencing it plays an important role in the optimal design of extraction machines and in the improvement of their operating parameters.

We plotted the curves of the specific energy consumption of coal, two types of lignite and rock salt as a function of the cutting depth, the values of which were determined based on the results of laboratory measurements.

We did the same analysis for harder and softer rocks, because the differences are already greater, we were forced to use the logarithmic scale for comparison.

It has been proven, based on the processing of experimental results, that the product of specific energy consumption by the size of the particles (chips) produced is constant, regardless of the process used, and that this invariant corresponds to the specific cutting force.

Thus, starting from the relationship between the two essential parameters, it is possible to calculate the specific energy consumption based on specific cutting force, which can be established experimentally.

References

[1] I. Andras, M. Lazar, A. Andras: *The Correlation between Physical-Mechanical Properties and the Cutting Characteristics in the Case of Unminable Rock Formations From Husnicioara Lignite Open Pit*. 17th International Multidisciplinary Scientific GeoConference SGEM 2017, Sofia, 869–876. https://doi.org/10.5593/sgem2017/13/S03.110

[2] M. S. Nan, I. Kovacs, I. Andras, D. Jula: *Study of the Working Regime of the Bucket Wheel Excavators in the Conditions of Romanian Open Pit Lignite Mines*. 8th WSEAS International Conference on Simulation, Modelling and Optimization (SMO '08), Santander, Cantabria, Spain, September 23-25, 2008. 83–88. https://www.researchgate.net/publication/262329172_Study_of_the_working_regime_of_the_bucket_wheel_excavators_in_the_conditions_of_Romanian_open_pit_lignite_mines

[3] L. Ozdemir: *Comparison of Cutting Efficiencies of Single-Disc, Multi-Disc and Carbide Cutters for Microtunneling Applications.* No-Dig Engineering, 2/3. 2002.

Műszaki Tudományos Közlemények vol. 18. (2023) 5–8.
DOI: Hungarian: https://doi.org/10.33895/mtk-2023.18.02
English: https://doi.org/10.33894/mtk-2023.18.02

THE GREEN TRANSITION IMPACT ON MINERAL RAW MATERIALS SUPPLY

József ANDRÁS,[1] Yannick HARRELL,[2] Eugène TASHCHI[3]

[1] *University of Petroșani, Faculty of Mechanical and Electrical Engineering, Department of Mechanical, Industrial and Transportation Engineering. Petroșani, Romania, iosif.andras@gmail.com*

[2] *University of Petroșani, Faculty of Mechanical and Electrical Engineering, Department of Mechanical, Industrial and Transportation Engineering. Petroșani, Romania.*

[3] *University of Petroșani, Faculty of Mechanical and Electrical Engineering, Department of Mechanical, Industrial and Transportation Engineering. Petroșani, Romania.*

Abstract

Minerals are an essential component of many parts of today's fast-growing clean energy technologies – from wind turbines and power grids to electric vehicles. Demand for these minerals will grow rapidly as the clean energy transition accelerates. The paper examines some prospects for the transformation of the complex relationships between minerals and the energy sector in the green transition, focused mainly on electric vehicles expansion and related mineral resources demand increase.

Keywords*: green transition, mineral raw material, electric vehicles.*

1. Foreword

The global clean energy transitions will have far-reaching consequences for mineral demand over the next 20 years. By 2040, total mineral demand from clean energy technologies are projected to double in the STEPS (STATED-POLICIES-SCENARIO) and quadruple in the SDS (SUSTAINABLE DEVELOPMENT SCENARIO).[1]

In both scenarios, EVs and battery storage account for about half of the mineral demand growth from clean energy technologies over the next two decades, spurred by surging demand for battery materials.

Mineral demand from EVs and battery storage grows tenfold in the STEPS and over 30 times in the SDS over the period to 2040.

By weight, mineral demand in 2040 is dominated by graphite, copper and nickel. [2]

Lithium sees the fastest growth rate, with demand growing by over 40 times in the SDS. The shift towards lower cobalt chemistries for batteries helps to limit growth in cobalt, displaced by growth in nickel. [3]

The mineral requirement is actual for the following clean energy technologies:

- Solar PV (utility-scale and distributed)
- Wind (onshore and offshore)
- Concentrating solar power (parabolic troughs and central tower)
- Hydro-power
- Bioenergy for power
- Nuclear power
- Electricity networks (transmission, distribution, and transformer)
- Electric vehicles (battery electric and plug-in hybrid electric vehicles)
- Battery storage (utility-scale and residential)
- Hydrogen (electrolysers and fuel cells).

All of these energy technologies require metals and alloys, which are produced by processing mineral-containing ores.

Ores – the raw, economically viable rocks that are mined – are utilised to liberate and concentrate the minerals of interest.[4]

Those minerals are further processed to extract the metals or alloys of interest.

Processed metals and alloys are then used in end-use applications. While this analysis covers the entire mineral and metal value chain from mining to processing operations, we use “miner-

als" as a representative term for the sake of simplicity.

Minerals are not only used in the clean energy sector, but are also used widely across the entire energy system, in technologies that improve efficiency and reduce emissions.

For example, the most efficient coal-fired power plants require a lot more nickel than the least efficient ones in order to allow for higher combustion temperatures.

However, here we focus specifically on the use of minerals in clean energy technologies, given that they generally require considerably more minerals than fossil fuel counterparts.

2. Critical minerals and green transition

While minerals play a vital role in supporting clean energy transitions, energy is also crucial in the production of minerals.

Due in part to declining resource quality, the production and processing of energy transition minerals are energy-intensive, involving higher emissions to produce the same quantity of product. [5]

In recent years, mining and processing companies have faced growing pressure to address these and other issues related to their social and environmental performance. A growing number of consumers and investors are requesting companies to disclose targets and action plans on these issues.

An energy system powered by clean energy technologies differs profoundly from one fuelled by traditional hydrocarbon resources.

While solar PV plants and wind farms do not require fuels to operate, they generally require more materials than fossil fuel-based counterparts for construction minerals.

A typical electric car requires six times the mineral inputs of a conventional car and an onshore wind plant requires nine times more mineral resources than a gas-fired plant of the same capacity.

Since 2010, the average amount of minerals needed for a new unit of electricity generation capacity has increased by 50%, as the shares of renewable energy sources grow from the total capacity expansion..

The transition to clean energy means a shift from a fuel-intensive to a material-intensive system.

Which sectors do these increases come from? In climate-driven scenarios, mineral demand for use in EVs and battery storage is a major force, growing at least three times to 2040.

Lithium sees the fastest growth, with demand projected to grow by over 40 times in the SDS by 2040, followed by graphite, cobalt and nickel (around 20-25 times).

The expansion of electricity networks means that copper demand for grid lines more than doubles over the same period.

The rise of low-carbon power generation to meet climate goals also means a tripling of mineral demand from this sector by 2040. (**Fig. 1**).

Wind takes the lead, bolstered by material-intensive offshore wind. Solar PV follows closely, due to the sheer volume of capacity that is added. (**Fig. 2**)

Hydro-power, biomass and nuclear make only minor contributions given their comparatively low mineral requirements.

In other sectors, the rapid growth of hydrogen as an energy carrier underpins major growth in demand for nickel and zirconium for electrolysers, and for platinum-group metals for fuel cells. [6]

A typical electric car requires six times as many minerals as a conventional car, and an onshore wind farm requires nine times as many mineral resources as a gas-fired power plant of the same capacity. (**Fig. 3**)

Since 2010, the average amount of minerals needed for a new unit of electricity generation capacity has increased by 50%, as the shares of renewable energy sources grow from the total capacity expansion. [7]

As can be seen from the above, it is not only the supply of so-called critical mineral raw materials that causes problems, but also the need for traditional metals is seriously affected by the green transition. [8]

For example, aluminium and copper.

Aluminium is mainly needed to reduce the weight of electric cars to compensate for the weight of batteries.

The drastic increase in copper demand is caused by its need for renewable energy production and transportation technologies.

The production of both metals, on the other hand, is energy-intensive, at each stage of the production chain, so most of the surplus green energy is consumed by these processes, so the reduction of net emissions is questionable.

In the case of copper, the critical situation is based on the fact that existing and foreseeable reserves have a lower content, their processing is more energy-intensive, and according to Hubert, peak production will be achieved in a few years. [9]

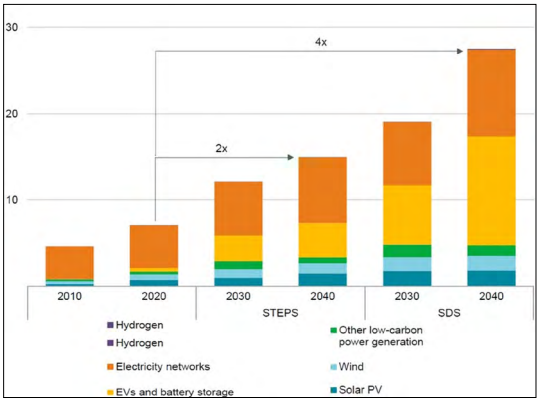


Fig. 1. *Dynamics of demand for minerals under different scenarios.*

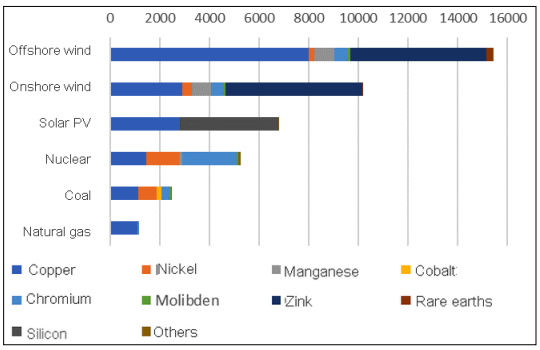


Fig. 2. *Comparison of metal requirements of different energy generation technologies.*

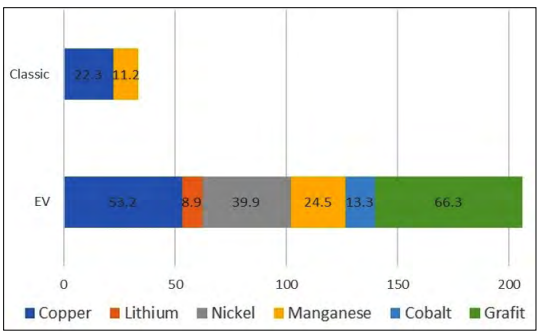


Fig. 3. *Comparison of metal requirements of conventional and electric cars.*

3. Europe and the supply of critical minerals

In September 2020 the European Commission published a set of policy documents to make Europe's raw materials supply more secure and sustainable.

It updated its policy directions from previous studies to align with new 2030 and 2050 climate ambitions. [8]

The policy package extended the list of critical minerals to 30, compared to 14 in 2011.

Together with bauxite, titanium and strontium, lithium was added to the list for the first time in 2020, reflecting the region's ambition to nurture a battery and EV manufacturing industry.

Some EU member states have a strong metal refining and manufacturing base. Finland refines around 10% of global refined cobalt output. There are major manufacturers of solar PV components, wind turbines and EVs in the region.

However, the region is almost entirely reliant on external mining supplies for many energy transition minerals such as lithium, cobalt and REEs.

The European Union plans to seek opportunities for sourcing critical minerals domestically, for example by tapping opportunities for enhanced metal extraction in post-mining regions.

The EU Action Plan estimates that this could lead to 80% of Europe's lithium demand being supplied from European sources by 2025.

To implement the plan, the European Union established the European Raw Materials Alliance, which involves industrial actors along the value chain, member states and regions, trade unions, civil society, research organisations, investors and non-governmental organisations.

The alliance aims to diversify supply chains, attract investment into the raw material value chain, foster technology innovation and create an enabling framework for the circular economy.

4. Conclusions

Minerals are an essential component of many parts of today's fast-growing clean energy technologies, from wind turbines and power grids to electric vehicles.

Demand for these minerals will grow rapidly as the clean energy transition accelerates.

The global clean energy transition over the next 20 years will have far-reaching consequences for mineral demand.

Total mineral demand from clean energy technologies will double under the Sustainable Development Strategy (STEPS) and quadruple under the Sustainable Development Strategy (SDS) by 2040.

In both scenarios, electric vehicles and battery storage account for about half of the growth in mineral demand from clean energy technologies over the next two decades, driven mainly by growing demand for battery materials.

Minerals are used not only in the clean energy sector, but are widely used throughout the energy system, in technologies that improve efficiency and reduce greenhouse gas emissions.

The transition to clean energy means the transition from a fuel-intensive system to a material-intensive system.

References

[1] *Total Mineral Demand for Clean Energy Technologies by Scenario, 2010-2040*. IEA, Paris. https://www.iea.org/data-and-statistics/charts/total-mineral-demand-for-clean-energy-technologies-by-scenario-2010-2040-2

[2] Carrara S., Alves Dias P., Plazzotta B., Pavel C: *Raw Materials Demand for Wind and Solar PV Technologies in the Transition towards a Decarbonized Energy System.* EUR 30095 EN, Publications Office of the European Union, Luxembourg, 2020. https://doi.org/10.2760/160859

[3] Fizaine F., Court V.: *Renewable Electricity Producing Technologies and Metal Depletion: a Sensitivity Analysis Using the EROI*. Ecological Economics, 110. (2015) 106–118. https://doi.org/10.1016/j.ecolecon.2014.12.001

[4] Giurco D., Dominish E., Florin N., Takuma Watari, McLellan B.: *Requirements for Minerals and Metals for 100% Renewable Scenarios*. In: Teske S. (eds): *Achieving the Paris Climate Agreement Goals*. (2019) 137–457. https://doi.org/10.1007/978-3-030-05843-2_11

[5] Yannick Harrel: *Electromobility in the Automotive industry. A Technological and Geopolitical Shift*. Multidiszciplináris Tudományok, 11/5. (2021) 428–437. https://doi.org/10.35925/j.multi.2021.5.44

[6] Henckens T.: *Scarce Mineral Resources: Extraction, Consumption and Limits of Sustainability,* Resources, Conservation and Recycling, 169. (2021) 105511. https://doi.org/10.1016/j.resconrec.2021.105511

[7] Ballinger B. et al.: *The Vulnerability of Electric Vehicle Deployment to Critical Mineral Supply*. Applied Energy, 255. (2019) 13844. https://doi.org/10.1016/j.apenergy.2019.113844

[8] https://ec.europa.eu/commission/presscorner/detail/en/STATEMENT_22_5523

[9] https://www.goldmansachs.com/insights/pages/gs-research/copper-is-the-new-oil/report.pdf

Műszaki Tudományos Közlemények vol. 18. (2023) 9–17.
DOI: Hungarian: https://doi.org/10.33895/mtk-2023.18.03
English: https://doi.org/10.33894/mtk-2023.18.03

THE ROLE OF LAJOS MARTIN IN HUNGARIAN ACADEMIC EDUCATION IN CLUJ

Enikő BITAY[1, 2]

[1] *Sapientia Hungarian University of Transylvania, Faculty of Technical and Human Sciences, Department of Mechanical Engineering, Târgu Mureș, Romania.*

[2] *Transylvanian Museum Association (EME), Cluj-Napoca, Romania, bitay.eniko@eme.ro*

Abstract

Lajos Martin (1827–1897) was a mathematician, engineer, corresponding member of the Hungarian Academy of Sciences (1861), founding member of the Transylvanian Museum Society, first appointed professor at the Hungarian University of Cluj founded in 1872, one of the Hungarian pioneers of aviation, and also warden of the Evangelical church. In 2022 was commemorated the 125th anniversary of his death and also the 150th anniversary of the founding of the University of Cluj, of which he was professor, from the foundation until his death. He was also the rector of the institution in academic year 1895/96. His legacy (the floating wheel, his collection of papers containing his research and the Martin folder) was bequeathed by his children to the Transylvanian Museum Society, the institution where he gave many lectures and published the results of his research. The present study aims to present in more detail not only the part of his life revealed by his commemorative research, but also that part of his life which highlights his role in Hungarian education in Cluj, as educator and institution leader, marked by his faith in science.

Keywords*: education, research, history of science, Lajos Martin.*

1. Introduction

Martin Lajos Martin, mathematician, engineer, corresponding member of the Hungarian Academy of Sciences (1861), founding member of the EME, first appointed professor of the Hungarian University of Cluj, one of the Hungarian pioneers of aviation, and the caretaker of the Lutheran parish – in 2022, was commemorated on the occasion of the 125th anniversary of his death and the 150th anniversary of the founding of the University of Cluj, of which he was a professor from its foundation until his death, and rector in the academic year 1895/96. His legacy (the floating wheel, his collection of papers containing his research, the Martin folder) was bequeathed by his children to the Transylvanian Museum Association, the institution where he gave many lectures and published the results of his research. Research into the life, work and legacy of Lajos Martin has been ongoing for several years, and we intend to publish the partial results, interesting facts and testimonies of this research from time to time.

2. Collection of the legacy of scientists

Researching the career of a scientist-teacher is always an exciting task, especially when the collection of sources and data provides additional information that forces one to reinterpret and thoroughly examine many known and published data. Lajos Martin can be characterised as many things: a man with a thirst for knowledge; attention to technical creation; a conscious pursuit of innovative realisations, from dreams and ideas to discoveries and then to proof; a drive to patent and protect results, presenting them in lectures; continuous publication in several languages; education, nurturing the offspring of knowledge.... The only way to find out who he really was is to collect his legacy, to study his works, his publications and the literature about him. Collecting legacies is one of the EME's major tasks: to discover, collect, digitise, preserve, care for and make accessible the legacies of scholars who would otherwise be lost. In the case of Lajos Martin, this was an even more interesting task, since in 1913 his

Figure 1. *Portrait of Lajos Martin (foto: Elinger Ede, Budapest, 1881. Source: MTA–KIK, Ms. 10.206/91)*

children gave to the EME their father's invention, the hoverwheel, with several documents, drawings and papers, so that it would be preserved and stay at home. Over the years, they have suffered a miserable fate due to the two world wars and the dissolution of the EME, and the confiscation of its collections. The legacy is now in the care of two institutions. Lajos Martin's folder is kept by the Romanian State Archives, Cluj County Directorate, and the floating wheel is in the Transylvanian National History Museum.. A digitised version of all this, together with the material collected over the years, is available as the Lajos Martin Collection, together with the literature about him, in the EDA (Transylvanian Digital Repository) run by EME. [1]

3. Studies (training), (secondary school) teaching activities of Lajos Martin

Lajos Martin was born on 30 August 1827 in Buda, the seventh child of a family of twelve. He began his education as a home-schooled pupil and continued his studies at the Lutheran school in Buda. Even then, as a student, he had already taken a big interest in astronomy under the influence of one of his teachers. He then went on to study at the Catholic High School in Buda, where he acquired a high level of mathematical knowledge. During his university years, he showed great interest in various fields, studying two years at the Budapest University of Science in the humanities and two years at the University of Art and Engineering. He interrupted his studies at the outbreak of the War of Independence in 1848. Lajos Martin volunteered for the army, taking up arms for the ideal of freedom. After World War II, he hid for a time from the vengeance of the retaliating powers. He was captured in 1849 and after a few weeks of captivity was recruited as a soldier in the Austrian army.

After the suppression of the Revolution, between 1849 and 1851, he was a school servant in Naples at the military school. He was then transferred to the technical officers' school in Naples, where his superiors soon recognised his mathematical talent - he gave mathematics lectures to the students as a pastime - and transferred him to the military engineering academy. At the Vienna Officers' Academy, he was assigned to the last class for his outstanding knowledge. He became a first lieutenant and then was appointed teacher of geometry, engineering and architecture at the technical school in Krems at the age of 27. It was here that he made his first invention in 1856, the design of the rotary rocket. At the same time, he designed a rotary military propeller, which so impressed the Austrian high command that he was commissioned to build a navigable dirigible. Martin did not accept, because he saw the future of aviation in dynamic flight.

He left the army in 1859 and returned to Hungary. He tried his hand at private engineering in Buda. His excellent engineering design, submitted for a public tender, was noticed by the city council and he was elected chief city engineer. „At the end of 1860 he became the chief engineer of Buda, but at the beginning of the provizórium he resigned and became a teacher at the Körmöcbánya Real, where he served until the establishment of the University of Cluj.” [2]. Between 1859 and 1861 he worked as a private engineer in Budapest, then he also qualified as a teacher. In the meantime, Lajos Martin was elected a corresponding member of the Hungarian Academy of Sciences on 27 November 1861, and in 1862 he presented his thesis entitled The Forces of the Wing of a Bird.

Qualified as a teacher, he took his final examinations on 5-6 October 1862: written questions in quantitative science and power engineering, and then a written exam on 9 October, in addition to the above subjects, in German, geography, history and natural science.

In 1863, he became a teacher at the Real School in Körmöcbánya. On 9 October 1863, Lajos Martin passed the examination as a candidate teacher of geometry and geometric drawing at the elementary school in Körmőczbánya. This is evidenced by the test certificate: „Mr. Lajos Martin, teacher of Secondary School, is hereby declared fully qualified to teach Quantitative, Mechanical and Mechanical Engineering in German and Hungarian language highschools!" [3].

Subsequently, between 1864 and 1868, he became teacher at the Real School of Bratislava and at the Catholic High School. In the school year 1864-65, Lajos Martin taught quantitative science in classes IV, V and VI and mechanical engineering in class VI in the mixed-religion main school. He was also head of class V. [4]. He also taught quantitative science at the Main School in Bratislava in the academic year 1866-67 in classes IV, V and VI, and was also the head teacher of class IV. [5] Commissioned by the Minister of Education, he wrote a textbook in Hungarian on quantitative, geometric and descriptive geometry..

In 1868 he was a telegraph office manager in Pest, then in 1869 deputy director of the telegraph office in Debrecen.

In 1871 Lajos Martin moved to Cluj-Napoca, where he was appointed director of the telegraph office, and in 1872, at the age of 45, he was asked to become senior professor of mathematics at the University of Science in Cluj-Napoca. He served in education for 25 years until his death. In the 1890s, Lajos Martin lived in *Gyár* Street (today Eminescu street) in Cluj-Napoca (at the end of *Fürdő* Street), across the street from his friend Farkas Gyulai.

4. Lajos Martin, the engineer-researcher

He became more interested in aviation in the last decade of his life. Due to space limitations, we will present only a few data in chronological order, which give an idea of his research work in the field of aviation: In 1871-1875 he made his first model of a flapping-wing aeroplane; in early July 1893 he made a model of his floating wheel; in October 1893 at the EME meeting in Cluj-Napoca Martin presented the first model of the "floating wheel", in 1893 he patented the floating wheel; in 1894 he published his work The floating wheel compared with the Wellner wheel. On 31 March 1896, his patent No. 81,303 was registered in Germany; in 1895, he also made a speech as rector in which he stated that flight was the most important task of science; on 30 August 1896, he tested the floating wheel in the garden of the University of Cluj.

5. Lajos Martin and the Transylvanian Museum Association

Lajos Martin was a founding member of EME. He was a frequent lecturer and published the results of his research in EME's scientific publications. He embraced the EME creed and took every opportunity to actively participate in events, but he even contributed to the development of the collections, for example, in one record he donated 20 minerals to the mineral collection of the Unitarian school in Cluj [7].Lajos Martin's close relationship with the EME was not in vain, since in addition to his teaching and academic work, he found a suitable background for his aviation research, which he considered his heart's desire, both for the presentation of his results and for their publication and dissemination. His family was well aware of this when they wished to deposit his legacy at the EME: the floating wheel and its accompanying documentation in a folder. The Lajos Martins folder is a collection of 291 pages of documents. His biography, patents, correspondence, descriptions of inventions, drawings and related correspondence, as well as correspondence from him and his grandson of the same name. The documents are mostly in Hungarian or German, but also in English and French. The documents in the folder date from 1871 to 1913

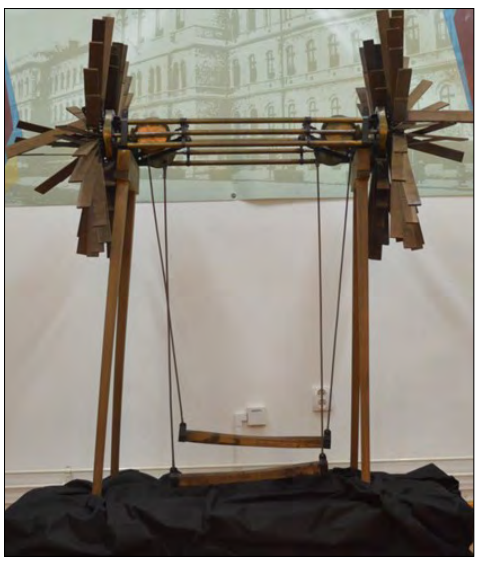


Figure 2. *Floating wheel of Lajos Martin (Foto: Bitay, 2013)* [6]

6. The role of Lajos Martin in Hungarian university education in Cluj

On 29 September 1872, he was one of the first to be appointed a full professor.: „On 29 September, in Ischl, His Majesty appointed as teachers at the University of Cluj... 3) for the Faculty of Human and Real Sciences, János Szamosi, Otto Hóman, Sándor Imre, Hugo Meltzl, Béla Szász, Lajos Felméri, Károly Szabó, Gedeon Ladányi, Henrik Finály, Lajos Martin, Antal Abt, Antal Fleischer and Antal Koch..” [8] about the proposal of the Hungarian Ministry, of the newly established University of Cluj and its appointed teachers, the *Budapest Közlöny* [9] and the *Magyar Polgár* [10] journals also report.

The observatory in Cluj Napoca operated between 1832 and 1872 under the direction of the piarist mathematical-astronomical teachers.. When the University of Cluj was established, the supervision and scientific management of the observatory was entrusted to mathematics professor Lajos Martin. Martin, appreciative of the old instruments, tried to add modern ones, two delta tubes, a theodolite, a comet finder and a pendulum clock. Between 1872 and 1885, the observatory was used by the university and the high school. Lajos Martin taught astronomy to university students in the observatory until 1893, when it was demolished along with the old university building. It was no longer located in the newly built university building, now the central building of the Babeş-Bolyai University. (A new observatory with modern equipment was established on *Tordai* road in 1924.) [11]

Lajos Martin was the first Vice Dean of the Faculty of Mathematics and Natural Sciences between 1872 and 1873, he taught elementary astronomy 3 hours per week, higher calculus 5 hours per week, integral calculus 5 hours per week, upper calculus 2 hours per week, astronomy 1 hour per week, application of external calculus (now differential calculus) to geometry 5 hours per week..

Lajos Martin is mentioned in the *Almanach* of the University as a "public ordinary teacher" between 1873-74, as a member of the Teachers' Examination Committee of the Secondary School, as a full member of the Hungarian Royal Society of Natural Sciences and the Hungarian Engineers' Society. His residence is listed as 8 Linczeg Street. At the university, descriptive geometry was taught continuously until 1918 [12]**.** Lajos Martin started teaching in the school year 1872/73 "descriptional geometry ", and in the second half of the school year 1873 he taught the subject of descriptive geometry for two hours a week. In 1875, he also taught projective theory, the theory of complex functions, calculus of variations and geometric drawing at the university, according to the university's *Almanach* [13].

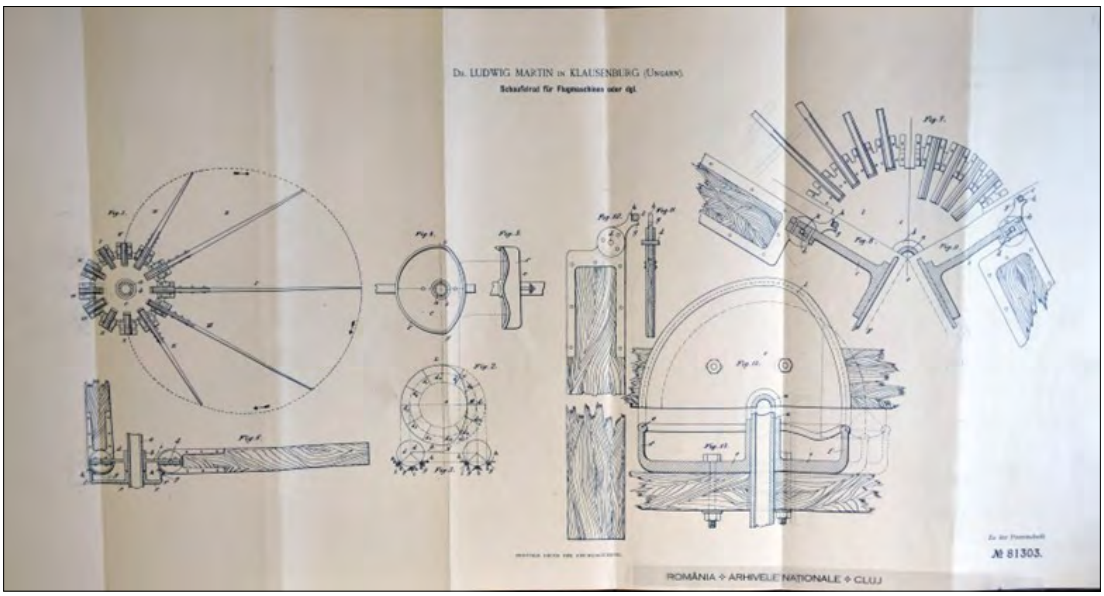


Figure 3. *Drawing of Lajos Martin's floating wheel [Source: Martin-mappa]*

7. Publications for teaching: books, notes

Lajos Martin's manuscripts for teaching purposes in his legacy testify to the time and attention he devoted to the preparation of his teaching materials. Thus we can find notebooks, manuscripts and printed manuals in the fields of mechanics, quantum mechanics, descriptive geometry and calculus of variations. Let us look at the five manuscripts below:

– Notebook. Notizen-Buch [14]
– Mechanics notes. Manuscript [15]
– Guide to teaching freehand geometrical drawing [16]
– The geometrical drawing according to patterns [17]
– Variatio calculation [18]

7.1. The notebook, NOTIZEN-BUCH, 1852

The notebook is a manuscript in German, consisting of 1 envelope, 49 written pages and 9 blank pages (30 sheets in total). The notebook, NOTIZEN-BUCH, dates from 1852, the writings contained therein are undated. We can only be certain that the entries were made in 1852 or later. No date or place of publication is given. During this period, Martin was a student at the technical officers' school in Naples.

The manuscript is believed to be a teaching note.

The first part of the note is a description of the theory of finite differentials, followed by a section on vector summation and then the theory of permutations. The following chapter, "Fundamenten des Analysis", discusses the properties of functions. This is followed by static torque, and then various statical problems, among which I think I can discover the calculation of the friction of a rope in a suspended pulley, the calculation of the centre of gravity of a circular slice, and the study of a cylindrical body moving down a ramp. The next chapter of the note introduces the basics of fluid mechanics. The note concludes with a section on finite differentials of multi-variable functions. It is interesting to note that most of the note is devoted to finite differentials, while the flow theory section is based on differential calculus

7.2. Mechanics notes

Manuscript, in a booklet of 42 pages (33 written and 9 blank) (21 sheets in all). No date of publication or location where he was at the time and wrote this note. Manuscript, which may have been intended as a teaching note.

The mechanics-notes starts with an introduction to the basic concepts of dynamics. He formulates the equations of motion on the basis of differential theory, and then moves on to a discussion of motion in a central force field, which he calls "central motion". This is followed by a description of the laws of motion of the Solar System. These are followed by the motion of the planets, presumably in elliptical orbits, the oblique deflection in unresisted or resistive media, and the periodic motion of bodies, presumed to be spherical, on various plane curves under the influence of gravity. The chapter is called "On pendulum motion".

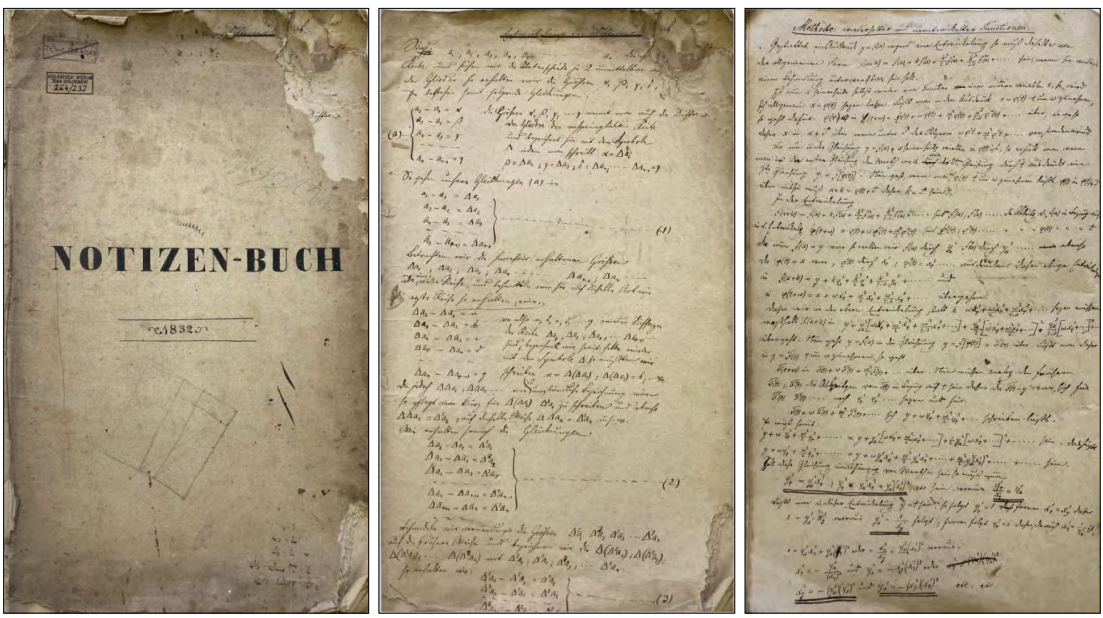


Figure 4. *The cover of Lajos Martin's notebook cover and the first and 42nd pages. (1852)*

The note concludes with a chapter on the motion of the material point system.

7.3. Guide to teaching freehand geometrical drawing

The book is a guide to teaching freehand geometric drawing, a handbook for teachers who are already familiar with the subject but need to supplement their theoretical knowledge with practical exercises. The book is therefore a rich collection of forms from which the teacher can choose the most suitable for teaching and putting into practice the subject matter. The 28-page manuscript contains 295 (geometrical) graphs (drawings): I. The straight line, II. The closed perimeter graphs, III. The regular polygons, IV. The simplest applications of regular polygons, V. Star-shaped polygons, VI. Applications of regular and star-shaped polygons, VII. Lines of overall dimensions, VIII. Application of the contour, IX. Including the chapters on straight line roses.

7.4. The geometrical drawing according to patterns

This note was originally intended to be entitled: Guide to teaching geometricál drawing according to patterns. It has been crossed out and corrected to read: Geometrical drawing by patterns. It is intended for use in gymnasiums and industrial schools.. At that time Lajos Martin was a full professor at the Main School of Economics in Bratislava, a full member of the Society of Engineers and a corresponding member of the Hungarian Academy of Sciences. („Martin Lajos a mennyiségtan rend. tanára a pozs. vár. Főreáltanódának, a m. tud. Akadémia lev. –; a m.k.term.tud társulat s a magy. Mérnök-egylet rend. Tagja”).

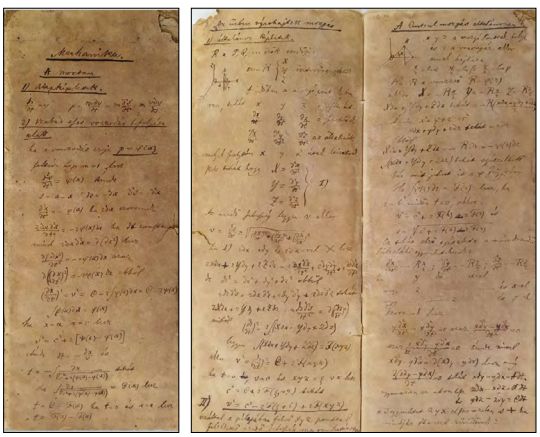
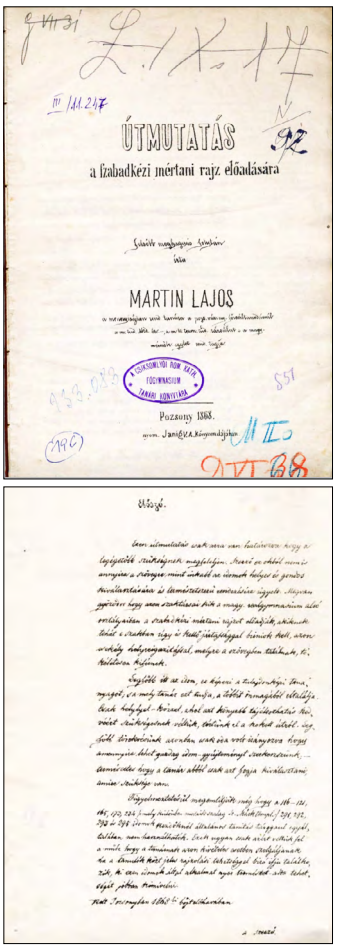


Figure 5. *Pages 1 and 6-7 of Lajos Martin's manuscript Mechanika.*

The book, which is a continuation of the previous one (Guide to teaching freehand geometrical drawing), contains 13 sheets of text and 200 geometric forms. It is a guide to teaching geometric drawing from patterns and how to use wire and wood patterns for teaching freehand geometry in the teaching of illustrative geometry.

7.5. Variatio calculation

His notes in Hungarian attest to the importance he attached to education in his mother tongue. Lajos Martin wrote in the foreword to his Variatio Calculus handbook: "The present book is intended to serve only as a handbook for my lectures, ignoring the higher needs and owes its existence only to the fact that there is not one of our textbooks written in Hungarian which deals with variation calculus adequately". The Variatio Calculation is therefore



6. ábra. *Cover and first page of Lajos Martin's manual " Guide to teaching freehand geometrical drawing ", with the foreword.*

a handbook to accompany the lectures, which is also a niche, because there was no other textbook written in Hungarian that discusses variation calculus in sufficient detail. The handbook is based on Strauch's two-volume work [19] but does not follow Strauch's theory of variational calculus in all respects, and supplements it where necessary. He intends his manual to be a resource, an aid to the development of a better variational calculus.

8. The role of Lajos Martin in the management of the university

Lajos Martin also took a role in the management of the university, and "was dean of the faculty of mathematics and natural sciences in 1877/8, 1882/3, 1883/4 and 1891/2. In 1872/3, 1878/9, 1884/5, 1886/7, 1887/8 and 1892/3 he was the vice dean of the university, and in 1895/6 he was the rector of the university" [20]. [13, 21, 22]

As already mentioned, Lajos Martin was appointed the first Vice-Dean of the Faculty of Mathematics and Natural Sciences between 1872 and 1873. [23] Since the foundation of the university, he has therefore held intermittent positions of leadership, as vice-dean, dean, and finally rector in 1895/96. At that time, the leadership positions were one-year terms, alternating, for example, from Vice Dean to Dean, then from Dean to Vice Dean, thus maintaining continuity in leadership, with long-term plans and guidelines.

On his way to the top of the university's leadership, in his rectorial inaugural speech in 1895, he discussed his own field of research, identifying aviation as the most important task of sci-

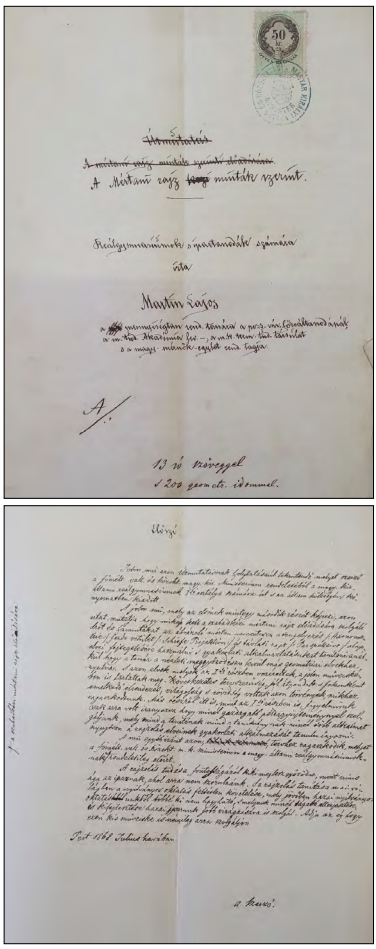


Figure 7. *Cover and first page of the manuscript of Lajos Martin's The Geometrical Drawing by Patterns with the preface (Pesta, 1868 july)*

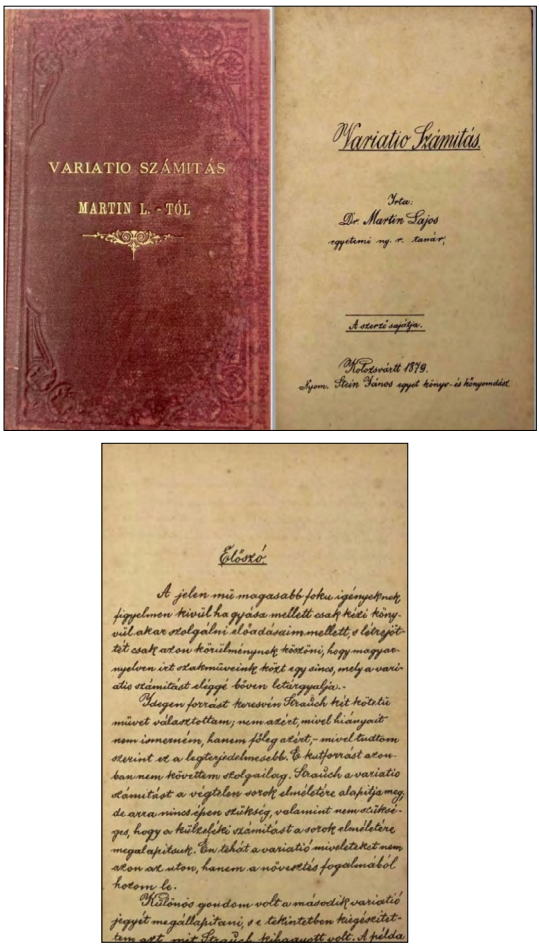


Figure 8. *Inside and outside cover of Lajos Martin's handbook Variational Calculation, fragment of the preface (Cluj, 1879. may)*

ence. We might think that this was a day of joy for him, when he was at the height of his career, when he could talk about his favourite science as rector, but it was not a day to experience his complete success, a peaceful moment. All this was overshadowed by a sad family tragedy. In Sándor Márki's diary of 15 September 1895 we read the following note: "... we attended the opening of the university, this time in the completed eastern wing of the university's own central building. The new rector, Lajos Martin, had caught up with the fate of recent rectors, and had to give his inaugural speech in the knowledge that one of his daughters was lying in a coffin...". [2] For Lajos Martin's eldest daughter, his second-born child Zsuzsanna (b. 12/12/1866, Bratislava) died in her 29th year on 14 September 1895, the day before the opening, after a long period of suffering. Her funeral was held on the afternoon of 16 September 1895, "with great condolences" [24] .

It is important to stress that Lajos Martin was not only an inventor, mathematician, mechanic and pioneer of the idea of flight, but also an excellent teacher. For half a century (25 years of which he spent at the University of Science in Cluj-Napoca) he taught and educated generations to think scientifically. Teaching ha beend his life-long vocation, whether as a school servant, military officer, teacher in a real school or as a university teacher (professor).

9. Lajos Martin in the Lutheran parcel of the Házsongárd cemetery in Cluj

Lajos Martin died at 5 o'clock in the morning of 4 March 1897. His funeral was held on 6 March in the Lutheran cemetery of Házsongárd in Cluj. Dr. Gyula Farkas, Dean of the Faculty of Mathematics, gave a eulogy at the funeral. In the diary of Sándor Márki we read this note: "In today's issue of Kolozsvár I wrote anonymously about his death: 'His human masterpieces wanted to ascend into the air / And while he himself remained below, his soul flew into the sky. / He now knows the way and whispers to his descendants from there / The great secret that lifts the terrestrial to the sky". [2]

10. Conclusions

To successfully bring a technical creation to life, the genius, dedication and persistence of the creator are not always enough; it is equally necessary to have the right conjuncture, the right people and the power of shared recognition to breathe life into the invention and launch it into the world of application..

Unfortunately, there was no national interest in the "invention(s)", and the researcher was not willing to sell it abroad, so he struggled to develop it as long as he could, with limited financial resources. A review of the results of his very rich scientific work, which was linked to aviation, reveals a scientist who was well ahead of his time, although some of his calculations were incorrect. He tried to be open on controversial issues, but this does not detract from the value he had, and it is our task, the task of posterity, to appreciate, value and preserve his memory at all times. It is important to identify, organise and preserve the technical heritage that we have. It is primarily an inspiring example for the younger generation, but it is also a rich source of ideas and further research.

Acknowledgements

I would like to thank the EME Technical Sciences Division for their support for the publication of this paper. Thanks also to Adrienn Székely and János Talpas for their help in digitising the documents discovered in the Hungarian Museum of Technology and Transport.

Figure 9. *Restored tombstone of Lajos Martin in the Házsongárd lutheran cemetery.*

References

[1] Bitay E.: *Szemelvények Martin Lajos (1827–1897) műszaki hagyatékából*. Műszaki tudományos közlemények, 6. (2017) 27–38. https://doi.org/10.33895/mtk-2017.06.03

[2] Erdész Á. (szerk.): *Márki Sándor naplói II. 1893–1903*. Gyula, 2018. 99., 115., 171.

[3] *Martin Lajos vizsgalati bizonyítvány*. Magyar Műszaki és Közlekedési Múzeum Műszaki Tanulmánytár, Témagyűjtemény 264, 6. boríték, 182.

[4] Érdemsorozat az 1864/65 tanév második felében. Pozsony, 1865. 1.

[5] Érdemsorozat az 1867/68 tanév második felében. Pozsony, 1867. 1.

[6] Bitay E., Máté M.: *Martin Lajos a feltaláló mérnök és lebegőkereke*. Műszaki Tudományos Füzetek - FMTÜ, 18. (2013) 74–82. https://doi.org/10.36243/fmtu-2013.11

[7] Kovács J.: *Értesítés az unitárius iskolákról. 1. A kolozsvári főtanodáról.* Keresztény Magvető, 14/4. (1879) https://eda.eme.ro/handle/10598/19319

[8] Márki S.: *Az egyetem eszméjének s a kolozsvári M. Kir. Ferencz-József-tudomány-egyetemnek története*. In: *A kolozsvári Magyar Királyi Ferencz-József-tudomány-egyetem története és statisztikája. Kolozsvár Acta Reg. Scient. Universitatis Claudiopolitanae Francisco-Josephinae Anni MDCCCXCV—VI. Fasciculus II*. Ajtai K. Albert könyvnyomdája. 1896. 33.

[9] Budapesti Közlöny, 6. évf. 226. szám, 1872. október 3. Thursday, 1805. 1.

[10] Magyar Polgár, VI. évf. 227. szám, Kolozsvár, 1872. oktober 4. Friday. 2.

[11] Józsa J.: *Tanulmányok. Erdély régi csillagvizsgáló tornyai.* Kriterion, Cluj-Napoca, 2016. 162–163.

[12] Oláh-Gál R.: *Az ábrázoló geometria erdélyi oktatásának történetéhez*. Természet Világa, 2012. 143/10. 458. https://eda.eme.ro/handle/10598/27193

[13] *A kolozsvári Ferencz-József Tudományegyetem Almanachja* MDCCCLXXIII–MDCCCXCVI-ról (1873–1896).

[14] Martin L.: *Jegyzetfüzet/Notizen-Buch*. 1852 (?). https://eda.eme.ro/handle/10598/33388

[15] Martin Lajos: *Mechanika*. Manuscript. https://eda.eme.ro/handle/10598/33386

[16] Martin L.: *Útmutatás a szabadkézi mértani rajz előadására*. Janig V. A. kőnyomdája, Bratislava, 1868. february. https://eda.eme.ro/handle/10598/27004

[17] Martin L.: A *mértani rajz minták szerint. Reálgymnásiumok s ipartanodák számára*. Pest, 1868. https://eda.eme.ro/handle/10598/33387

[18] Martin L.: *Variatio számitás*. Stein János egyet. könyv- és kőnyomdász, Kolozsvár, 1879. https://eda.eme.ro/handle/10598/27003

[19] Strauch G. W.: *Theorie und Anwendung des sogenannten Variationscalcul's,* 1–2. kötet. Verlag u. Meyer & Zeller, Zürich, 1854.

[20] *Gyászrovat: Martin Lajos*. Kolozsvár, Magyar Polgár, XI. évf. 1897. márc. 5., 52. sz. 3. p.

[21] *Acta Reg. Scient. Universitatis Claudiopolitanae Francisco-Iosephinae* Anni. MDCCCLXXII–MDCCCXCVI. (1872–1896)

[22] *A kolozsvári Ferencz-József Tudományegyetem Tanrendje* az MDCCCLXXII–MDCCCXCVI. tanév I-ső felére. (1872–1896) .

[23] *A kolozsvári M. Kir. Ferencz-József-Tudomány-Egyetem története és statisztikája*. Kolozsvár, 1896, 36.,100.

[24] *Gyászrovat. Kolozsvár*, 9/211. 1895. september 16.

Műszaki Tudományos Közlemények vol. 18. (2023) 18–24.
DOI: Hungarian: https://doi.org/10.33895/mtk-2023.18.04
English: https://doi.org/10.33894/mtk-2023.18.04

EXERGY ANALYSIS OF A VAPOUR COMPRESSION HEAT PUMP USING ENVIRONMENTALLY FRIENDLY REFRIGERANTS

János DÓSA,[1] Ovidiu-Bogdan TOMUȘ[2]

[1] *University of Petrosani, Faculty of Mechanical and Electrical Engineering, Department of Mechanical, Industrial and Transportation Engineering, Petrosani, Romania, iondosa@upet.ro*

[2] *University of Petrosani, Faculty of Mechanical and Electrical Engineering, Department of Mechanical, Industrial and Transportation Engineering, Petrosani, Romania, bogdantomus@upet.ro*

Abstract

The paper presents the comparative exergetic analysis of heat pumps using refrigerants with different ozone depletion potential and global warming potentials. Globally, the air conditioning of houses is seen as a major energy challenge as the economic development of countries, the majority situated in hot climates, will lead to a growing demand for chilling. Added to this is the increasingly widespread use of heat pumps in areas with a temperate climate to increase energy efficiency. As a result, the environmental impact of refrigerants from ozone depletion potential and global warming potential point of view cannot be overlooked. On the other hand, an environmental benefit of high energy efficiency is the reduction of greenhouse gas emissions from fossil fuel combustion for various purposes. The purpose of this approach is to highlight how the use of different types of refrigerants will impact the exergy efficiency of heat pumps.

Keywords*: heat pump, exergy efficiency, refrigerants, environmental impact.*

1. Introduction

Energy savings are a major concern in the modern world, as since the first energy crisis in 1973 mankind has basically suffered a chronic energy shortage, one that worsens once in a while, usually triggered by major political developments.

The energy crisis is aggravated by increasingly pressing environmental problems, which is why new, environmentally friendly energy production technologies are gaining more and more prevalence.

One such technology is the heat pump, which is a heat engine used to extract heat from a lower-temperature environment and deliver it to a higher-temperature environment, using a refrigeration cycle.

The purpose of its use is the management of thermal energy, during which ambient heat and waste heat can be used for heating and domestic hot water production.

Heat pumps use electricity to transfer heat from a cool space to a warm space but they can be used for cooling too. During the heating season, they move heat from the environment into your home and during the cooling season, heat pumps move heat from your house into the environment. Their efficiency results from the fact that they transfer heat instead of generating it.

Globally, the cooling of houses is considered a significant energy challenge, as the economic development of countries with mostly warm climates leads to an increase in the demand for air conditioners. Added to this is the increasingly widespread use of heat pumps in areas with a moderate climate in order to increase energy efficiency.

As a result, the environmental effects of refrigerants cannot be ignored.

On the other hand, the environmental advantage of high energy efficiency is the reduction of greenhouse gas emissions generated during the burning of fossil fuels for various purposes.

The purpose of this approach is to highlight how the use of different types of refrigerants will impact the exergy efficiency of heat pumps.

2. The vapour compression heat pump

Vapour compression heat pumps are perhaps the most widely used type of heat pumps. The first of these types worked with R-717 (ammonia) refrigerant, and this refrigerant is still common in large refrigeration systems today. Ammonia is a dangerous compound for both the environment and people as it is toxic and explosive. Later, propane (C_3H_8), methyl chloride (CH_3Cl), sulfur dioxide (SO_2), and several other compounds were used. The Freon family is the trade name for halogenated hydrocarbons, which are widely used in the refrigeration industry due to their excellent properties. Freons, or CFCs as they were also known, have been widely used due to their excellent stability and safe use: they are non-flammable and less toxic than other refrigerants. Unfortunately, one of their properties has become very dangerous: if freon reaches the upper atmosphere, its chlorine content destroys the ozone layer, which protects the earth's surface from the ultraviolet radiation of the Sun.

CFC refrigerants were replaced by newer and less polluting refrigerants, hydro-chlorofluorocarbons (HCFCs), such as R-22, which is used in most household refrigerators today, and HFCs (such as R-134a), which became popular in automobiles. Currently, researchers are developing more environmentally friendly refrigerants in order to replace these refrigerants as well, so more and more research is addressing this topic.

One such refrigerant, which is completely neutral from an environmental point of view and non-toxic, is water vapour (R-718). So, in the following, ammonia (R-717) and the more modern, environment-friendly refrigerant (R-134a) and heat pumps operating with water vapour (R-718) will be analysed in terms of energy efficiency and environmental point of view.

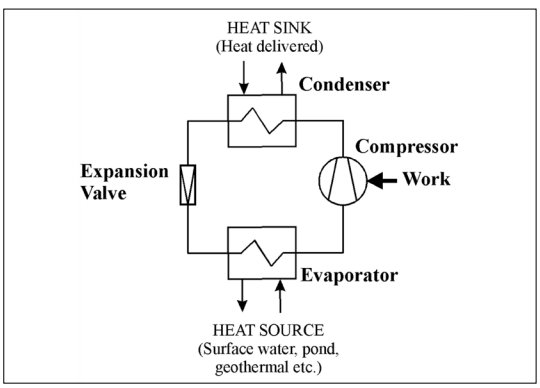


Fig. 1. *Schematics of vapour compression heat pump.*

2.1. The vapour compression heat pump

The water-to-water vapour compression heat pumps are suitable for transferring heat from a lower temperature source to a higher temperature sink [1] [2]. This makes them the first choice for low-temperature waste heat recovery systems.

This type of heat pump (Figure 1) is suitable for producing domestic hot water using waste heat from water below 40 °C and heating it up to 50-80 °C.

The mathematical model of such a heat pump was developed within the framework of the paper [3] so in the following the thermodynamic characteristics of the heat pump will be calculated based on that model.

Figure 2 shows the cycle of a vapour compression heat pump [4]. The development of the mathematical model of the vapour compression heat pump is based on this cycle.

2.2. Vapour compression heat pump using R-718 refrigerant

In Figure 3 ea vapour compression heat pump using R-718 (water vapour) refrigerant is presented [4] followed by its cycle in T-s diagram (Figure 4).

Common elements can be recognized in both figures, such as the valves (throttle valves are marked FS1 and FS2 in Figure 3 and the compressor, which is a turbo-compressor labelled with Tk in the case of the heat pump using R-718 refrigerant.

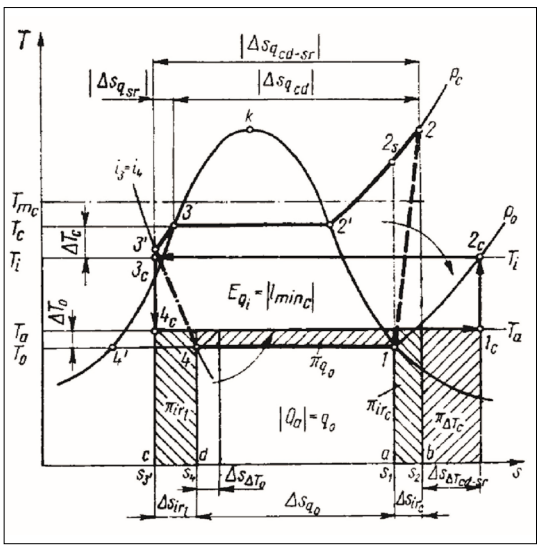


Fig. 2. *Vapour compression heat pump cycle T-s diagram.*

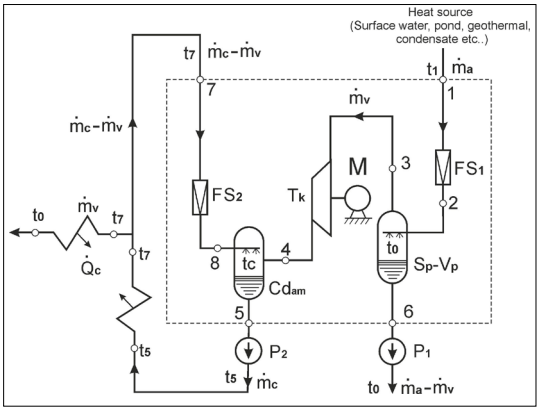


Fig. 3. *Schematics of vapour compression heat pump using water vapour (R-718) as refrigerant.*

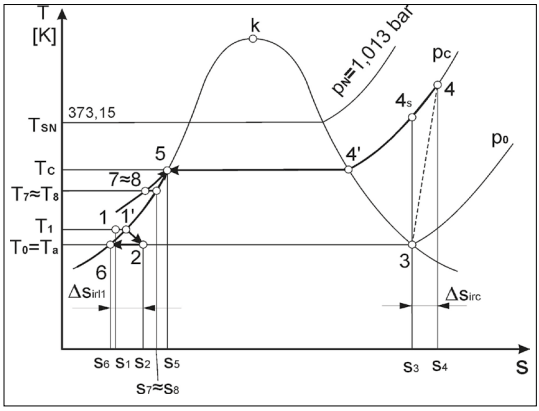


Fig. 4. *Vapour compression heat pump cycle T-s diagram for R-718 refrigerant.*

Due to the refrigerant used, the construction of the second heat pump is much more complicated, and as a result, a few common characteristics were selected for the purpose of comparison, namely: the power required for operation *Pe* [kW], the theoretical μ and actual μ_e coefficient of performance (COP), and the exergy efficiency ηE.

3. Efficiency of heat pumps

The mathematical model of the vapour compression heat pump was presented within the framework of paper [3] and based on this, the thermal characteristics of the heat pump assuming that it works with ammonium refrigerant (R-717) and R-134a refrigerant were calculated.

The mathematical model of the heat pump operating with water vapour refrigerant (R-718) can be found in literature [4–6] so in the following, the thermal characteristics of the heat pump based on those models were calculated.

The following assumptions were made in order to perform the calculus:

– Q – required heat delivery must be at least the amount of heat absorbed from the heat source Q = 30 kW (which is enough to meet the heating needs of a medium-sized household);
– T_i = 55 °C – temperature of delivered hot water;
– T_a – ambient temperature (heat source), which is set to 30, 35 and 40 °C respectively.

In addition, depending on the type of heat pump, we perform the calculations with the following data:

– for vapour compression heat pump:
  • ΔT_c = 10 °C – temperature excess required for heat transfer in condenser (heat delivery);
  • ΔT_0 = 5 °C – temperature drop required for the heat transfer of the evaporator;
  • T_{sr} = 5 °C – temperature drop needed for sub-cooling;
  • refrigerant used R-717 (ammonium) and R-134a;
  • η_{em} = 0,9 – compressor efficiency.
– for the water vapour compression heat pump calculations:
  • temperature of vapour-liquid separator $t_0 = t_a$ = 20, 25, and 30 °C, depending on heat source temperature
  • turbo-compressor internal efficiency η_i = 0,72,
  • mechanical efficiency η_{em} = 0,9.

In Table 1 the thermal characteristics of the heat pump operating with ammonia (R-717) refrigerant, in Table 2 the thermal characteristics of heat pump operating with R-134a refrigerant can be found, while Table 3 presents the thermal characteristics of the heat pump operating with water vapour (R-718).

The data in the tables are presented in Figures 5–10 in order to highlight the dependencies and to more easily compare the different characteristics of heat pumps operating with three types of refrigerants.

Figure 5 highlights that as the temperature rises, the amount of power required to operate the circuit decreases.

At the same time, it is clear that the heat pump operating with water vapour refrigerant requires less energy to operate the cycle.

The power required to operate the cycle is the largest in the case of the heat pump operating with R-134a refrigerant (Figure 5).

Table 1. *Thermal characteristics of a heat pump operating with R-717 refrigerant*

Temperature of heat source, T_a [°C]	30	35	40
Power supplied to the cycle, P_e [kW]	5,03	4,30	3,59
Exergy loss due to irreversibility of compression π_{irc} [%]	9,11	8,57	7,46
Exergy loss due to irreversibility of expansion π_{irl} [%]	5,55	4,90	4,22
Exergy loss due to irreversibility of heat transfer in evaporator π_{q0} [%]	9,43	11,12	13,45
Exergy loss due to irreversibility of heat transfer in condenser $\pi_{\Delta Tc}$ [%]	26,09	29,20	33,82
Theoretical coefficient of performance , *COP* μ	6,62	7,74	9,29
Actual coefficient of performance, *COP* μ_e	5,96	6,97	8,36
Exergy efficiency η_E	50,48	47,17	42,47

Table 2. *Thermal characteristics of a heat pump operating with R-134a refrigerant.*

Temperature of heat source, T_a [°C]	30	35	40
Power supplied to the cycle, P_e [kW]	5,54	4,52	3,75
Exergy loss due to irreversibility of compression π_{irc} [%]	12,15	9,21	7,61
Exergy loss due to irreversibility of expansion π_{irl} [%]	15,56	14,58	12,37
Exergy loss due to irreversibility of heat transfer in evaporator π_{q0} [%]	8,40	10,51	12,81
Exergy loss due to irreversibility of heat transfer in condenser $\pi_{\Delta Tc}$ [%]	16,96	21,87	25,35
Theoretical coefficient of performance , *COP* μ	6,01	7,37	8,89
Actual coefficient of performance, *COP* μ_e	5,41	6,64	8,01
Exergy efficiency η_E	46,91	43,82	41,86

Table 3. *Thermal characteristics of the heat pump operating with R-718 refrigerant.*

Temperature of heat source, T_a [°C]	30	35	40
Minimum power supplied to the cycle, P_{min} [kW]	2,1	1,64	1,44
Exergy loss due to irreversibility of compression π_{irc} [%]	16,47	17,86	19,34
Exergy loss due to irreversibility of expansion π_{irl} [%]	8,90	10,71	13,11
Exergy loss due to irreversibility of heat transfer in evaporator $\pi_{\Delta Tc}$ [%]	29,91	30,30	31,67
Theoretical coefficient of performance, *COP* μ	6,37	7,51	9,1
Actual coefficient of performance, *COP* μ_e	5,62	6,60	7,95
Exergy efficiency, ηE	44,63	41,13	35,87

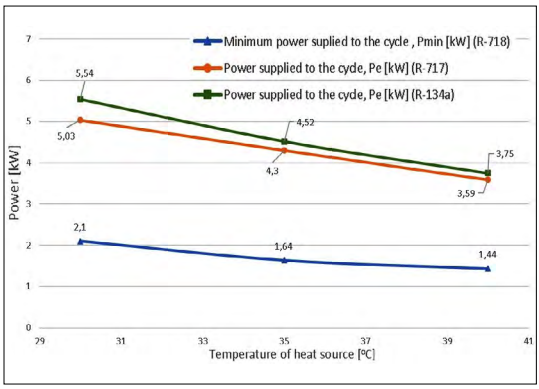


Fig. 5. *Power supplied to the cycle.*

Figure 6 shows that there is the same dependence between the temperature of the heat source and the exergy efficiency as in the case of the supplied power, namely as the temperature rises, the exergy efficiency decreases.

From this point of view as well, the heat pump operating with ammonia (R-717) refrigerant performs better, the exergy efficiency is higher in all cases, while in the case of R-134a refrigerant it is located between the two, but at high heat source temperatures it is almost equal to the exergy efficiency of ammonia refrigerant.

The most important thermal characteristics of a heat pump is the theoretical coefficient of performance (COP) μ and the actual coefficient of performance, μ_e.

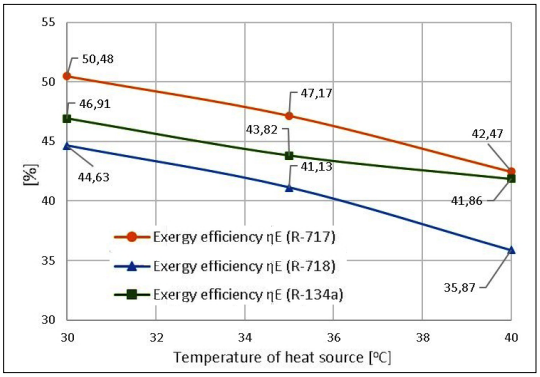


Fig. 6. *Exergy efficiency.*

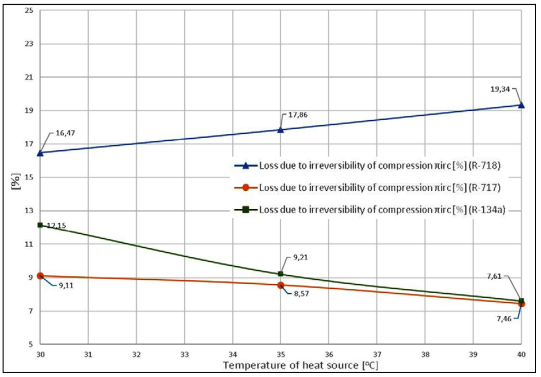


Fig. 8. *Exergy loss due to irreversibility of compression* π_{irc}.

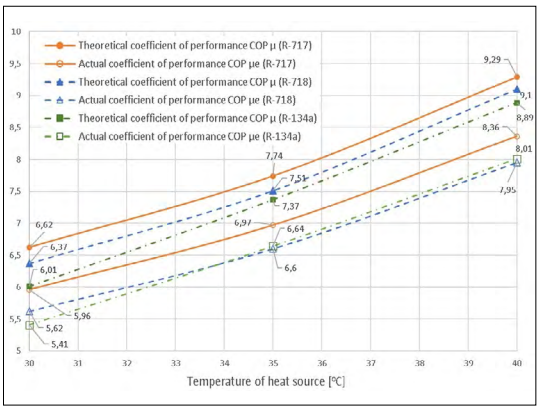


Fig. 7. *Theoretical and actual coefficient of performance (COP).*

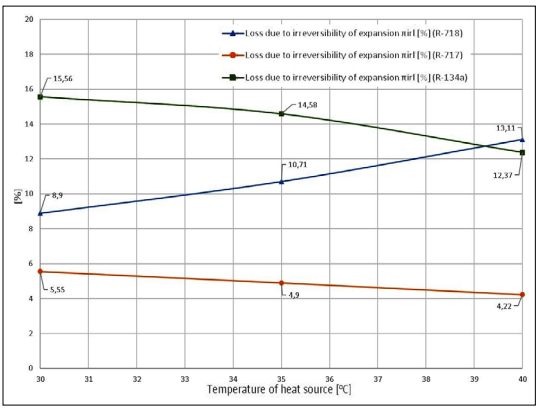


Fig. 9. *Exergy loss due to irreversibility of expansion* π_{irl}.

Figure 7 presents the variation of the theoretical COP μ and the actual COP μ_e as a function of the temperature of the heat source.

In case of the use of the three types of refrigerants, as the temperature of the heat source increases, the theoretical performance factor μ and the actual performance factor με also increases.

As presented in **Figure 7** the theoretical performance factor is the highest for ammonia (R-717), followed by water vapour (R-718) and R-134a. In terms of actual performance, at higher heat source temperatures, R-134a has a better performance than R-718 refrigerant (water vapour).

Figures 8., **9.** and **10** highlight the following exergy losses: the exergy loss due to the irreversibility of the compression process π_{irc} [%], the exergy loss in the throttle and expansion valve π_{irl} [%], the exergy loss in the condenser due to the irreversibility of the heat transfer process $\pi\Delta T_c$ [%].

The exergy efficiency diagram presented in Figure 6 allows us to conclude that the ratio of losses remains the same, so the largest exergy losses will occur in the case of the heat pump operating with water vapor (R-718) refrigerant, followed by the heat pump operating with R-134a refrigerant, while in the case of the heat pump operating with ammonia (R-717) refrigerant the smallest exergy losses are expected.

This is also the case in **Figure 8** where the exergy loss due to the irreversible nature of the compression process π_{irc} is presented.

Value of exergy loss due to the irreversibility of the compression process π_{irc}, is the highest in the case of water vapor (R-718) refrigerant and increases with the temperature of the heat source, while it decreases in the case of R-717 and R-134a.

It should be noted that the construction of a heat pump using water vapor is much more complicated due to the refrigerant used, so only the losses that occur in each case are shown.

In **Figure 9** the order of magnitude of the losses due to irreversibility of expansion π_{irl} rises with

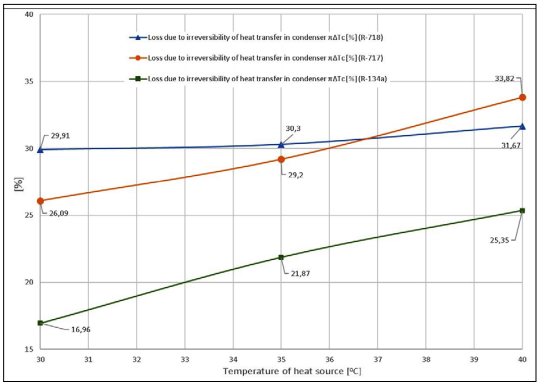


Fig. 10. *Exergy loss due to irreversibility of heat transfer in condenser, $\pi\Delta T_c$.*

the temperature of the heat source, for R-718 refrigerant, while for R-717 and R-134a decreases.

The schematics of the vapour compression heat pump using water vapour (R-718) as refrigerant shows there are two throttle valves.

This can lead to the conclusion that losses due to irreversibility of expansion have to be greater that one in the case of the heat pump using R-717 or R-134a refrigerant. But in **Figure 9**, it can be seen that only at high heat source temperatures, exergy loss due to irreversibility of expansion π_{irl} is higher. Correlating these with data in **Figure 5** where the power required to operate the cycle can be found, the conclusion is that the lower value of the loss due to irreversibility of expansion occurs due to the lower flow rate of refrigerant passing through the valves, when using R-718 refrigerant. The nature of variation shows that this loss is directly proportional to the temperature of the heat source.

In **Figure 10** exergy loss due to irreversibility of heat transfer in condenser, $\pi_{\Delta Tc}$ is presented.

Exergy loss due to irreversibility of heat transfer in condenser $\pi_{\Delta Tc}$ mrises for all types of refrigerants used.

For lower temperatures of heat source exergy loss due to irreversibility of heat transfer in condenser $\pi_{\Delta Tc}$ is higher for R-718 refrigerant, while at higher temperatures of heat source this type of loss is higher for R-717 refrigerant, while for R-134a is the lowest in all cases.

4. Conclusions

The thermodynamic characteristics of the heat pump operating with ammonia (R-717) refrigerant are better than the characteristics of the heat pump operating with water vapor (R-717) refrigerant, while the values of characteristics of the heat pump operating with R-134a refrigerant are usually in between.

An exception is the power required for operation of the cycle, which is much smaller in the case of a heat pump working with water vapour (R-718) refrigerant, approximately 2.5 times smaller. Values are: 3.59 kW for the heat pump operating with ammonia (R-717) refrigerant compared to the heat pump operating with water vapour refrigerant, which requires only 1.44 kW.

As for the heat pump working with R-134a refrigerant, the power required for operating the cycle is 3.75 kW.

The theoretical coefficient of performance (COP) μ in the case of the heat pump operating with ammonia refrigerant, according to the heat source temperature of 40 °C is 9.29, while in the case of the heat pump operating with water vapor refrigerant it is 9.1, so it is only 2% higher and is also very close in the case of R-134a, 8.89.

The actual coefficient of performance (COP) μ_e in the case of the heat pump operating with ammonia refrigerant according to the heat source temperature of 40 °C is 8.36, while in the case of the heat pump operating with water vapour refrigerant it is 7.95, which is 5.2% higher, and the value of the actual coefficient of performance for the heat pump working with R-134a refrigerant is 8.01, located between the two values.

These values of the coefficient of performance (COP) show that although the thermodynamic characteristics of the heat pump operating with ammonia refrigerant (R-717) are better, they are not significantly better than in the case of R-134a refrigerant, and in fact, they do not display large differences compared to the results of the water vapor refrigerant (R-718).

If the exergy efficiency is analyzed, the heat pump operating with ammonia refrigerant (R-171) displays better values, followed by R-134a and the R-718 refrigerant being in last place.

According to the heat source temperature of 40 °C the exergy efficiency of the heat pump operating with ammonia refrigerant is η_E=42.47 %, while in the case of the heat pump operating with water vapor refrigerant η_E=35.87 %, 18.4 % higher. In the case of using the modern R-134a refrigerant, this value is η_E=41.86%, which is barely perceptibly lower than the R-717 value.

The differences between the values (the actual coefficient of performance μ_e and exergy efficiency η_E) especially between R-718 and the other refrigerants, are higher than the tolerable margin

of error for the preparation of the energy balance, which is 3% [7], so they cannot be ignored.

There is one last comparison criterion left, the effect of the refrigerant on the environment.

From this point of view, the heat pump working with water vapour refrigerant (R-718) is by far the best choice, because, in the case of an accident, water vapor is completely neutral from an environmental point of view, while ammonia is particularly harmful being toxic, and moreover, can be explosive.

As for R134a, it is a modern refrigerant with a global warming potential (*GWP*) of 1430 (relative to CO_2) and an ozone depletion potential (*ODP*) of zero [8]. It should be mentioned here that although ammonia is toxic and explosive, its global warming potential (*GWP*) is zero and its ozone depletion potential (*ODP*) is also zero [9].

So, in the event that it is necessary to use heat pumps in places where the risk of refrigerant leakage is unacceptable from a work safety regulations point of view, such as underground mines, tunnels, etc. the use of heat pumps with water vapor refrigerant is a perfect solution and does not generate much higher operating costs from an energy usage point of view.

References

[1] Radcenco, V. et. al.: *Processes in Refrigeration Equipment*. Didactică si Pedagogică Pulishing House, Bucharest, 1983, 372–390.

[2] Energy Department of the USA https://www.energy.gov/eere/buildings/downloads/water-source-heat-pumps (accessed on 2022 september 15)

[3] Dosa, I.: *Power Plant Waste Heat Recovery for Household Heating Using Heat Pumps*. Multiscience-XXVIII. microCAD International Multidisciplinary Scientific Conference, Miskolc, 10-11 april 2014, 1–8. https://doi.org/10.13140/rg.2.1.2616.2406

[4] Radcenco, V. et. al.: *Heat Pumps*. Tehnica Publishing House, Bucharest, 1985, 65–211.

[5] Radcenco, V. et. al.: *Refrigeration Installations and Cryogenics: Problems and Applications.* Tehnica Publishing House, Bucuresti, 1987,75–100.

[6] Laza I.: *Refrigeration Plants*. Universitatea Tehnica, Timisoara, 1995,60-125

[7] ***, *Guide to Development and Analysis of Energy Balance*. M.O. of Romania, part. I, nr.792/11.11.2003.

[8] Gasservei SA (accessed on: 2022. september 30.) https://gas-servei.com/shop/docs/technical-data-sheet-r-134a-gas-servei.pdf

[9] Danfoss (accessed on: 2022. september 30.) https://www.danfoss.com/en/about-danfoss/our-businesses/cooling/refrigerants-and-energy-efficiency/refrigerants-for-lowering-the-gwp/ammonia-nh3/

Műszaki Tudományos Közlemények vol. 18. (2023) 25–28.
DOI: Hungarian: https://doi.org/10.33895/mtk-2023.18.05
Ebglish: https://doi.org/10.33894/mtk-2023.18.05

EXAMINATION OF SYSTEMATIC ERRORS OF A ROUNDNESS MEASUREMENT EQUIPMENT

Erzsébet EGYED-FALUVÉGI,[1] Tamás KARÁCSONY-FEJÉR,[2] László JAKAB-FARKAS[3]

[1] *Sapientia Hungarian University of Transylvania, Faculty of Technical and Human Sciences, Department of Mechanical Engineering, Târgu Mureș, Romania, faluvegi.erzsebet@ms.sapientia.ro*

[2] *Sapientia Hungarian University of Transylvania, Faculty of Technical and Human Sciences, Department of Mechanical Engineering, Târgu Mureș, Romania, karacsony.tamas@student.ms.sapientia.ro*

[3] *Sapientia Hungarian University of Transylvania, Faculty of Technical and Human Sciences, Department of Mechanical Engineering, Târgu Mureș, Romania, jflaci@ms.sapientia.ro*

Abstract

The profile in the frontal section of a solid of revolution forms a perfect circle if all of its points are at an equal distance from the center line. The value of the deviation from circularity is very important and it is measured with a dedicated measuring device. This paper examines the systematic errors in a measuring device in which the measured workpiece is supported and rotated between two live centers, while a dial indicator detects the deviations. Regarding the structure, this paper proposes a model consisting of five systematic error measurements which considers the eccentricity of the live centers and the tilt of the dial indicator axis.

Keywords*: roundness, systematic error, coaxiality.*

1. Introduction

Circular features are the most common in machined parts which are three dimensional solids of revolution. The profile in the frontal section of a solid of revolution forms a perfect circle if all of its points are at an equal distance from the center line

According to the standard [1, 2] in every section perpendicular to the main axis, the effective profile must fall within two concentric circles. The tolerance is the radial distance between the two concentric circles.

The deviation can manifest in ovality or polygonal features [3, 4]. In practice, this deviation can even be seen with the naked eye: the shape of the part is not circular; however, this is impossible to measure with conventional measuring devices. Therefore, the construction of a measuring device was required, one which allows us to determine the amount of circular deviation.

2. The construction of the circularity measuring device

Measuring circularity requires a dedicated measuring device, which allows the user to measure the circular profile in every cross-section of the part. There are two different methods for such devices [4]. The first method consists of positioning the workpiece in a clamping device horizontally held up by either one or multiple points of contact. During the measuring process, it revolves around its axis and is swept by an indicator. This is the method recommended by the authors of study [5]. In the second method, the workpiece is positioned vertically with a precision bearing on one of its ends, coaxially aligned with its main axis. After examining and comparing both methods and considering their costs, we have chosen to use the first one, in which the workpiece is positioned horizontally between two centers and revolves around its axis. An indicator is used to sweep the outer profile of the part (Fig. 1). The jig holding the indicator slides horizontally and can be brought into contact with the profile of the workpiece at any point on its surface.

The device allows us to measure parts of various lengths and diameters with limitations on their length and diameter due to the construction of the device. The measurements of the largest part that fits the device are as follows: length 195 mm, diameter 80 mm. One of its drawbacks is that it is unsuited for measuring bores.

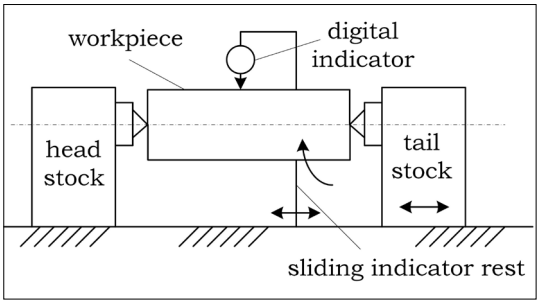


Fig. 1. *The theoretical model of the device.*

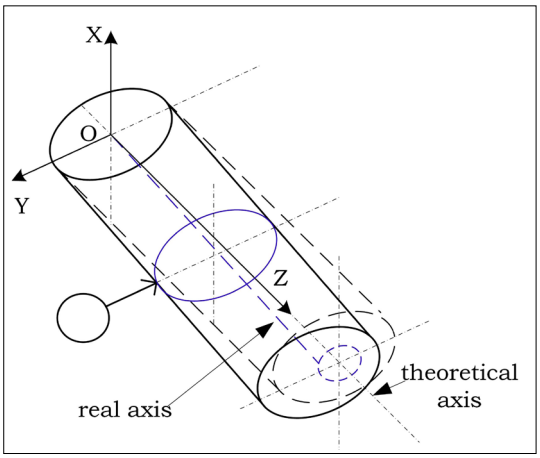


Fig. 2. *Examining the coaxiality of the live centers, three-dimensional model.*

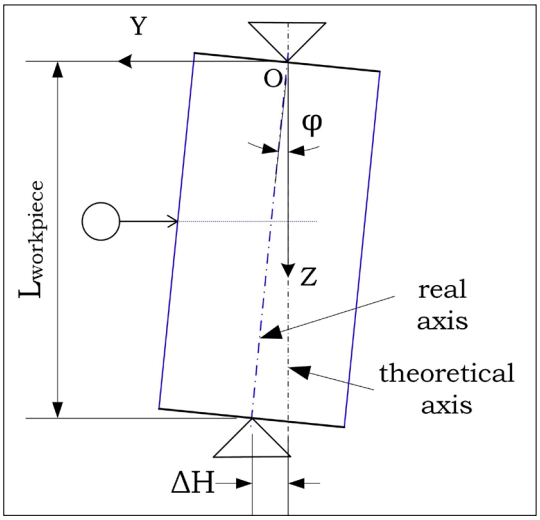


Fig. 3. *Examining the coaxiality of the live centers in the horizontal plane.*

3. Examination of systematic errors

Scientific research considers errors to be a deviation between real and theoretic (nominal) values. There are two major categories of errors: random and periodic errors. This paper examines the latter; therefore, it does not address random errors.

There are numerous methods for determining periodic errors [6–11]. Study [6] formulates and elaborates on the concept of periodic errors and uses a V-block for moving the measuring device. Studies [7, 8] measure the circularity as a function of the points of measurement, the radius of the indicator ball and eccentricity. Studies [9, 10] measure the circularity as a function of the points of measurement, the radius of the indicator ball and eccentricity. Studies [11] examines the coaxiality of a turbine shaft, however, this study proposes a model with seven measurement errors.

Regarding the structure, this paper proposes a model made of two systematic error measurements which considers the eccentricity of the live centers and the tilt of the dial indicator axis.

3.1. Examining the coaxiality of the live centers

During assembly or manufacturing, it is possible for the live centers to shift radially in relation to each other. Figure 2 eshows a three-dimensional sketch which includes the conical surface. This surface is the result of the difference between the ideal and real axes of the centers. The measured profile is an ellipse in every measuring plane.

The coaxiality problem of the centers is broken down into two directions.

3.1.1. Examining the coaxiality of the live centers in the horizontal plane

The centers are offset from one another along the OY axis by a distance of ΔH (Figure 3). Since this distance has a constant value, which results from the construction of the device, it has a φ inclination angle:

$$\varphi = \operatorname{arctg} \frac{\Delta H}{L_{darab}}. \qquad (1)$$

According to (1), if ΔH is constant and the length of the workpiece increases, the value of φ angle decreases continuously.

3.1.2. Examining the coaxiality of the live centers in the vertical plane

The centers are offset from one another along the OX axis by a distance of x_p (Figure 4).

Since this distance has a constant value, which results from the construction of the device, it has a θ inclination angle:

$$\theta = \arcsin \frac{x_P}{R}, \tag{2}$$

where R is the radius of the workpiece.

This vertical shift is shown as ΔV on the indicator:

$$\Delta V = R(1 - \cos\theta). \tag{3}$$

According to (3), if x_p is constant and the value of R continuously increases, the value of ΔV decreases

3.2. Tilt in the axis of the digital indicator

Figure 5 shows the tilt in the axis of the digital indicator. In this case, there is an angle of δ between the axis of the indicator and the horizontal plane.

Due to the tilt, the indicator does not show the rea Δy difference but:

$$m = \frac{\Delta y}{\cos\delta}. \tag{4}$$

If the value of δ continuously increases for a given Δy difference, the height m shown on the indicator also increases, however, this increase is inconsequential.

3.3. Compensation for systematic errors

Currently the measuring device is in the planning and construction phase. The construction allows the indicator to move along the OX and OY axes. The actuators for these movements are electric motors. Both the operation of the motors and the processing of the data from the indicator will be performed via software.

For measuring coaxiality, the centers have a two-step cylindrical surface with known dimensions and tight tolerances. Indicating either cylindrical surface in any given cross section results in a circle. The center points of these circles can then be determined. These two center points on the same live center form an axis and if the coaxial deviation has a value different than zero, then the two centers have two separate axes. This distance can then be determined via software.

The tilt in the axis of the indicator is created using a part with a known Δy radial deviation. During production or assembly, a $\delta \neq 0°$ angle may appear, however, if this is kept below 5° its effect is insignificant since the Filetta 611-343 indicator has a resolution of 0.01 mm and an accuracy of ±0.02 mm.

4. Conclusions

Measuring circular deviation is a measuring method used in practice, one for which the industry uses various devices. This paper focuses on the periodic errors in a circularity measuring device intended for laboratory use. Familiarity with the device is crucial for its successful operation. The paper proposes an error measuring model comprised of five parts.

The paper proposes a model comprised of two parts measuring periodic errors and also deals with compensating methods for said error. Since the measuring device is currently under planning and construction, the paper contains no measured results.

References

[1] EN ISO 1101: *Geometrical Product Specification (GPS) – Geometrical Tolerancing – Tolerances of Form, Orientation, Location and Run-out*, 2017.

[2] EN ISO 12181: Geometrical Product Specification (GPS) – Roundness, 2011.

[3] Ionescu N., Vișan A., Manolache D., Nistor C.: *Tolerances Design*. Editura Printech, București, 2016, 54–57.

[4] Curtis M. A., Farago F. T.: *Handbook of Dimensional Measurement*. 5. ed. Industrial Press Inc., South Norwalk, 2014. 377–406.

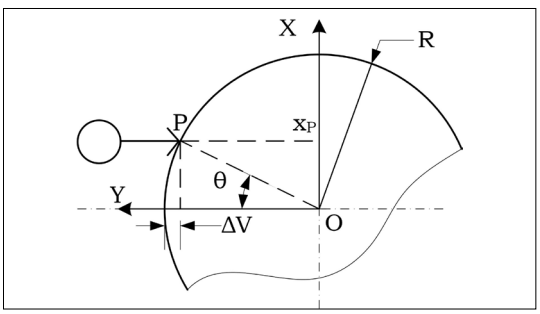


Fig. 4. *Examining the coaxiality of the live centers in the vertical plane.*

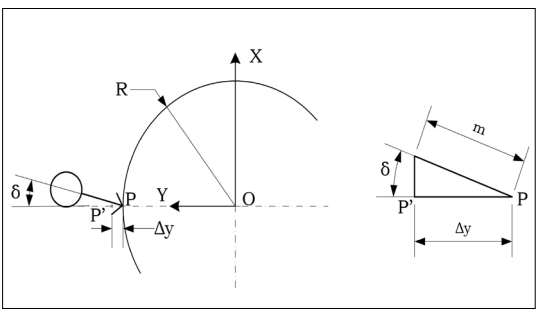


Fig. 5. *Tilt in the axis of the digital indicator.*

[5] Tero M., Papp I.: *Dispozitive de control*. Editura Universității Petru Maior, Târgu-Mureș, 2002, 71–108.
[6] Adamczak S., Janecki D., Stepien K.: *Qualitative and Quantitative Evaluation of the Accuracy of the V-Block Method of Cylindricity Measurement*. Precision Engineering, 34. (2010) 619–626. https://doi.org/10.1016/j.precisioneng.2010.03.004
[7] Ciobotă D., Palade D. D., Stanciu D., Abălaru A.: *Consideration Regarding Roundness Measurement of Closed Profiles and Open Profiles*. UPB Sci. Bull., Series D, 75/2. (2013) 129–140.
[8] Zhao Z., Li B., Zhang G., Yu H., Shang M.: *Influence of Eccentricity and Tilt of Cylindrical Part's Axis on the Measurement Results of Its Diameters*. Measurement, 138, (2019), 232–239. https://doi.org/10.1016/j.measurement.2019.01.085
[9] Sun C., Wang L., Tan J., Zhao B., Tang Y.: *Design of Roundness Measurement Model with Multi-Systematic Error for Cylindrical Components with Large Radius*. Review Of Scientific Instruments, 87. (2016) 955–961. https://doi.org/10.1063/1.4941679
[10] Sun C., Wang H., Liu Y., Wang X., Wang B., Li C., Tan J.: *A Cylindrical Profile Measurement Method for Cylindricity and Coaxiality of Stepped Shaft*. The International Journal of Advanced Manufacturing Technology, 111. (2020) 284501502856. https://doi.org/10.1007/s00170-020-06296-5
[11] Liu Y., Li R., Sun C., Mei Y., Wang X., Tan J., Wang H.: *A Coaxiality Measurement Model Based on Multi-Systematic Errors Separation for Turbine Shaft*. Measurement, 186. (2021) https://doi.org/10.1016/j.measurement.2021.109975

Műszaki Tudományos Közlemények vol. 18. (2023) 29–36.
DOI: Hungarian: https://doi.org/10.33895/mtk-2023.18.06
English: https://doi.org/10.33894/mtk-2023.18.06

THE STUDY OF THE PROFILE ERRORS OF STRAIGHT TEETHED SHAPER CUTTERS IN THE CAD-ENVIRONMENT

Károly-István GÁL,[1] Márton MÁTÉ[2]

[1] *Sapientia Hungarian University of Transylvania, Faculty of Technical and Human Sciences, Department of Mechanical Engineering, Târgu Mureș, Romania. gal.karoly@ms.sapientia.ro*

[2] *Sapientia Hungarian University of Transylvania, Faculty of Technical and Human Sciences, Department of Mechanical Engineering, Târgu Mureș, Romania. mmate@ms.sapientia.ro*

Abstract

It is well known that straight-teethed shaper cutters present theoretical profile error, which leads to a deviation of the cut gear tooth profile from the involute profile. Taking advantage of the CAD environment, we have visualized and studied shaper cutters at different stages of sharpening. This is advantageous because the effect of any shape or dimension modification of the rake face can be immediately and easily controlled in the design process, thereby saving precious design time.

The aim of our research was to create a 3D model in a CAD environment to visualize different sharpening methods and the profile of shaper cutters at different wear stages, and to estimate the inherent profile errors. The aim of using different sharpening methods was to improve the lateral cutting geometry (especially the lateral relief angle) and also to reduce or – at least – keep the profile error within the accepted limits. During the measurements, both the rake angle and the cutter rack profile angle were changed in order to improve the geometry. The edge was generated using numerically computed points and Autodesk Inventor's special commands. Comparative analyses were performed.

Keywords*: shaper cutter, profile error, CAD, rake face, rake angle, relief engle, cutter rack profile angle.*

1. Basic terminology

1.1. Structure of the classic cutting whee

The Fellow's cutter is essentially a gear with teeth having a rake angle (γ_v) and a relief angle (α_v) [1, 2, 3] as shown in Figure 1.

The rake face is straight circular cone, coaxial with the shaper cutter. Based on the literature, we choose the rake angle $\gamma_v = 5°$ and accept the meshing of it with a straight profile rack type tool which causes profile error [3].

The computing of the profile angle α_s of the Fellow's cutter's meshing rack is based on the approximation presented in the literature [1, 3], where the shaper cutters are considered to have infinite teeth, thus the involute profile is distorted into a straight profile, which causes the profile error.

The tip-relief surface is designed with the use of the characteristic property of involute gears, i.e. profile shifting, which results in a linear decrease of the addendum diameter (*Da*). The profile shifting reaches its maximum in the base plane section (I), is zero in the reference plane II, while plane III reaches its minimum value (ξ_{min}) Plane III also delimits the useful range (*Hh*) of the cutting edge wheel. The total height (*Ht*) is larger than this value to provide the strength and rigidity of the tool [3].

The computing of characteristic circle radii of the cutter is similar to the computing of a cylindrical gear [1–3], thus this calculus is omitted.

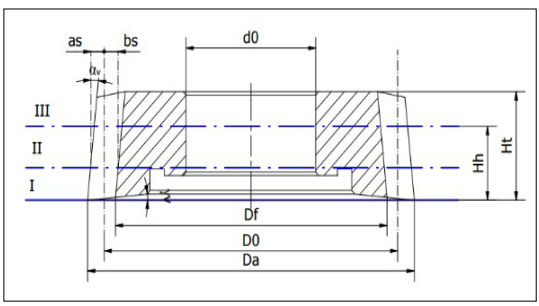


Fig. 1. *Constructive dimensions of the Fellow's cutter.*

1.2. The shape of the side edge

The side edge of the shaper cutter is the intersection of the rake and the lateral relief faces. The equations of the edge [3–6] are written from the parametric equations of the helical involute tooth surface combined in the implicit equation of the conical rake face. In the present communication, we also consider the change in the position of the frontal surface due to re-sharpening. The geometrical relations between the considered surfaces are given in Figure 2.

The position vector of an arbitrary point Q of the side relief surface is given by

$$\overrightarrow{OQ} = \overrightarrow{OG} + \overrightarrow{HG} + \overrightarrow{HQ} \tag{1}$$

whose matrix form related to the frame $Oxyz$ is:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R_b \begin{bmatrix} cos(\varphi - \eta) \\ sin(\varphi - \eta) \\ 0 \end{bmatrix} + \frac{PE}{2\pi}\varphi \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + \tag{2}$$

$$+u \begin{bmatrix} sin\,\beta_b\, cos(\varphi - \eta) \\ -sin\,\beta_b\, cos(\varphi - \eta) \\ -cos\,\beta_b \end{bmatrix}$$

The two chosen parameters for the surface representation are the rotation angle φ of the generating line holding plane while this rolls without slipping on the basic cylinder and the distance u from the point H corresponding to the angle of rotation φ, where the generating line is tangent to the basic helix.

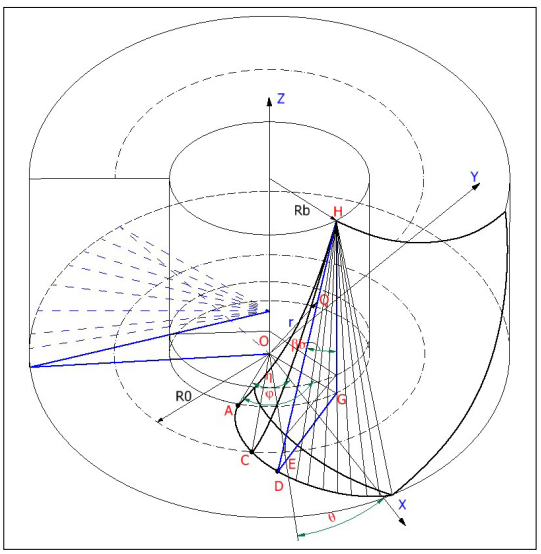


Fig. 2. *Geometric relationships of surfaces.*

The helix parameter can be written as

$$P_E = \frac{2\pi R_b}{tg\ \beta_b}, P = \frac{R_b}{tg\ \beta_b}, tg\ \beta_b = tg\ \alpha_v\ sin\ \alpha_s \tag{3}$$

With the linear transformation $v = u/R_b sin\beta_b$ the equations of the side relief surfaces becomes:

$$\begin{cases} x(\varphi, v) = R_b(cos(\varphi - \eta) + v\ sin(\varphi - \eta)) \\ y(\varphi, v) = jR_b(sin(\varphi - \eta) - v\ cos(\varphi - \eta) \\ z(\varphi, v) = P(\varphi - v), j \in \{-1; 1\} \end{cases} \tag{4}$$

Note that j = 1 results in the surface shown in Figure 2, while for $j = -1$ it is symmetrical with respect to the Oxz plane.

The angle η s the angle between the polar radius of the base circle point of the involute and the symmetry axis of the complete tooth profile in the Oxy plane, calculated from the second basic equation of evolutionary trigonometry [2]:

$$\eta = \frac{\pi}{2z_s} + 2\frac{\xi_s}{z_s} tg\ \alpha_s + inv\ \alpha_s \tag{5}$$

The geometric interpretation of the angles η and φ is illustrated in Figure 3.

The implicit equation of the rake face, based on Figure 2, is:

$$x^2 + y^2 = (R_a - z\ ctg\ \gamma_v)^2 \tag{6}$$

Assume that due to re-sharpening the rake face moves a distance δ_h talong the z-axis, in its positive direction. Then equation (6) takes the following form:

$$x^2 + y^2 = (R_a - (z - \delta_h)\ ctg\ \gamma_v)^2 \tag{7}$$

From equations (4) and (7), the following relationship between the parameters φ and v results:

$$\varphi(v; \delta h) = \frac{(R_a(\xi s) - R_b \cdot \sqrt{1 + v^2})tg(\gamma_v) + \delta_h}{P} + v \tag{8}$$

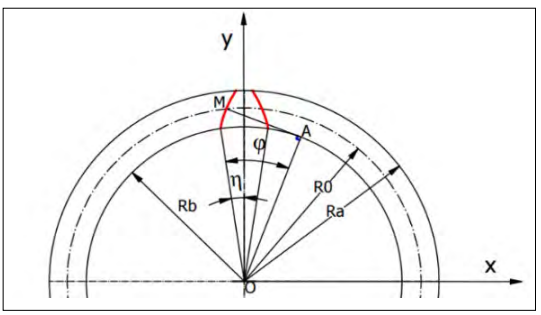


Fig. 3. *Involute's position parameter η and developing parameter ϕ.*

The coordinates of the points on the side edges are given by the coordinate functions (4), with the interparametric dependence (8).

2. Computation of the theoretical profile error

Due to the form of the rake face the tool edge points result one by one from different involutes comprised in different planes perpendicular to the cutter's axis. With increased re-sharpening the profile shifting decreases, thus the points of the implied profiles become closer and closer to the base circle. As a result, the edge form of the shaper cutter changes after each sharpening.

In the main movement, the cutter's edges mesh a gear: this is the equivalent gear of the cutter. Its tooth profile, according to the previous statements and equations (4-8), differs from the involute. The profile error is defined as the difference between the involute tooth profile of a gear with the same number of teeth as the shaper cutter and the tooth profile of the equivalent gear, measured perpendicular to the involute. The latter is referred to in the following as the control profile. The profile error of the cutting wheel can be calculated from the projection of the control profile and the tool edge perpendicular to the *Oxy* plane.

The reciprocate positions of the control profile and the edge projection are shown in **Figure 4**. As can be seen in the figure, the pitch circle points of the control profile and the edge projection coincide. The profile error is positive if the edge projection is outside the control profile and negative if it is inside. As known from **[1–5]**, a positive profile error causes a lightly dedendum undercut of the machined wheel, which ensures an unobstructed connection. Negative profile defects are undesirable as they cause tooth addendum width increasing, which may result in tooth crowding during coupling.

The value of the profile error, for an arbitrary point D of the edge projection equals the length of the segment ED. Its calculation is carried out as shown in **Figure 4.** The selected points of the edge are assumed to be known. Henceforth, we denote by R_b the base radius of the tool involute and by r_b the base radius of the control involute, so that

$$R_0 = \frac{R_b}{\cos\alpha_s} = \frac{r_b}{\cos\alpha_0}.$$

The coordinates of an arbitrary point D of the tool edge projection are (x^D, y^D). Thus:

$$DO = \rho = \sqrt{(x^D)^2 + (y^D)^2} \tag{9.a}$$

$$\angle DOT = \theta = \arccos\frac{r_b}{\rho} \tag{9.b}$$

$$\psi = \operatorname{arctg}\frac{-x^D}{y^D} \tag{9.c}$$

$$\angle COy = \operatorname{arctg}\frac{-x^C}{y^C} \tag{9.d}$$

$$\angle BOT = \theta - \psi + \angle COy + \operatorname{inv}\alpha_0 \tag{9.e}$$

$$ED = DT - ET = \sqrt{\rho^2 - r_b^2} - r_b(\angle BOT) \tag{9.f}$$

To apply the formulas (9.a–f) we need to compute the coordinates of the pitch circle point of the edge projection C. From the equations of the involute curve **[4]**, **Figure 5**.

$$\begin{cases} x(\varphi;\eta) = R_b(sin(\varphi-\eta) - \varphi\, cos\,(\varphi-\eta)) \\ y(\varphi;\eta) = R_b(cos(\varphi-\eta) + \varphi\, sin\,(\varphi-\eta)) \end{cases} \tag{10}$$

we compute the parameter value φ_0 corresponding to the pitch circle:

$$\varphi_0 = \sqrt{\frac{R_0^2}{R_b^2} - 1} = tg\,\alpha_s \tag{11}$$

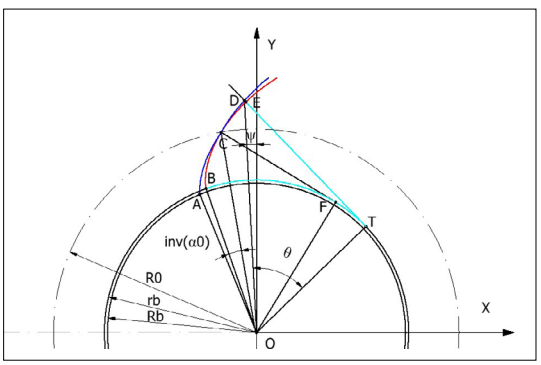


Fig. 4. *Illustration of the edge and control profile.*

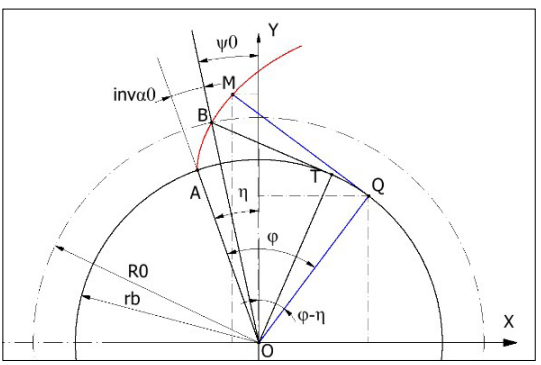


Fig. 5. *Computation of the arbitrary M point of the control involute-*

Thus, the angle between the position vector of point *C* and axis *Oy* is

$$\theta_0 = arctg\frac{x(\varphi_0;\eta)}{y(\varphi_0;\eta)} = inv\,\alpha_s - \eta \qquad (12)$$

A cylindrical helix is started from the point *C* of the tool string, whose parameter *P* is known from (3):

$$\begin{cases} x(w) = R_0 cos\,(\theta_0 + w) \\ y(w) = R_0\,sin(\theta_0 + w) \\ z(w) = Pw \end{cases} \qquad (13)$$

The pitch circle point of the edge is obtained by intersecting the helix given by eq. (13) and the conical rake face. From equations (13) and (7) it follows:

$$R_0^2 = (R_a - (Pw - \delta_h)\,ctg\,\gamma_v)^2 \qquad (14)$$

And from there

$$w_0 = \frac{1}{P}\left((R_a - R_0)\,tg\,\gamma_v + \delta_h\right) \qquad (15)$$

This will give the coordinates of the pitch circle point of the edge projection

$$\begin{cases} x^C = R_0 cos\,(\theta_0 + w_0) \\ y^C = R_0\,sin(\theta_0 + w_0) \\ z^C = (R_a - R_0)\,tg\,\gamma_v + \delta_h \end{cases} \qquad (16)$$

The points of the control profile (16) are calculated using the following equations:

$$\begin{cases} x^M(\varphi) = r_b(sin(\varphi - \eta_0) - \varphi\,cos\,(\varphi - \eta_0)) \\ y^M(\varphi) = r_b(cos(\varphi - \eta_0) + \varphi\,sin\,(\varphi - \eta_0)) \\ \eta_0 = arctg\dfrac{x^C}{y^C} + inv\,\alpha_0 \end{cases} \qquad (17)$$

3. Numerically generated curves

By fitting the equations obtained above into the Mathcad environment, the shaper's involute, the edge curve and the control involute can be plotted with high accuracy.

Equal distanced points are used to represent the curves in order to make the numerical interpolation as accurate as possible. It is a well-known fact [2, 3, 6, 7] that if the running parameter φ of the involute follows an arithmetic progression, then the lengths of the arc-sequences between the points result in a geometric progression. Thus, a linear scaling of φ is not recommended. For equal division, the following computational procedure is used:

– we specify the number of nodes on the considered curve segment;

– calculate the total length of the arc to be divided;

– calculate the arc length of the subdivision;

– iteratively determine the parameter sequence $\varphi_i,\ i \in \overline{0, N-1}$ defining the equidistant nodes.

The elementary involute arc length, according to equations (10), is

$$ds(\varphi) = \sqrt{\dot{x}^2 + \dot{y}^2}\,d\varphi = R_b\,\varphi d\varphi \qquad (18)$$

The interval of the parameter φ fits:

$$\sqrt{\frac{R_t^2(\xi_s)}{R_b^2} - 1} \le \varphi \le \sqrt{\frac{R_a^2(\xi_s)}{R_b^2} - 1} \qquad (19),$$

where $R_t^2(\xi_s)$ and $R_a^2(\xi_s)$ are the involute root radius and the addendum radius corresponding to the maximum tool profile shifting value, respectively.

Denote the corresponding parameter values by φ_t and φ_a respectively. The total length of the involute arc, by integrating the arc length (18) is

$$L = \int_{\varphi_t}^{\varphi_a} d\varphi = \frac{1}{2}R_b(\varphi_a^2 - \varphi_t^2) \qquad (20)$$

The *N* arc points determine *N*−1 involute-arc segments of equal length ΔL = L/(N−1). This is used to construct the corresponding parameter sequence, with the following recursion:

$$\varphi_0 = \varphi_t,\ \Delta L = \int_{\varphi_j}^{\varphi_{j+1}} d\varphi,\ j \in \{0,1,\dots,N-2\} \qquad (21)$$

After computing we obtain the following series:

$$\varphi_{j+1} = \sqrt{\frac{2\Delta L}{R_b} + \varphi_j^2}, j \in \{0,1,\dots,N-2\} \qquad (22)$$

Figure 6 shows N=13 equidistant points of the involute of a shaper cutter with z_s=27 teeth and m=3.75 mm module, while the tooth surface – a helical involute – is shown in **Figure 7.** Here also

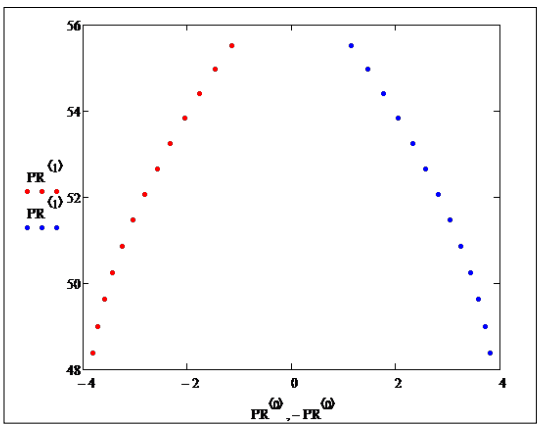


Fig. 6. *Equally distanced shaper involute points.*

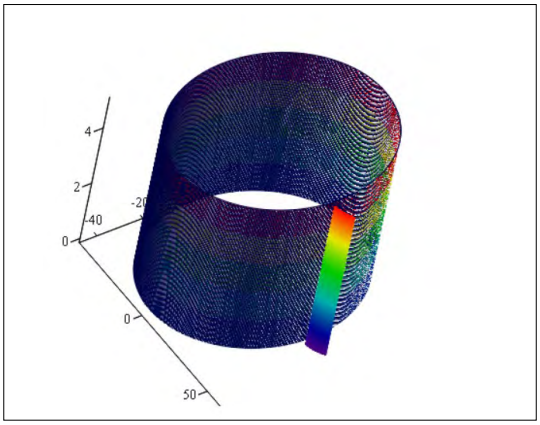


Fig. 7. *Tool tooth surface built on the involute.*

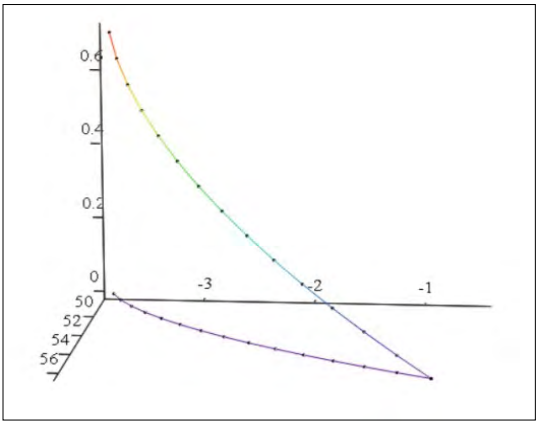


Fig. 8. *Equidistant divided tool edge curve and its projection.*

was considered the same equidistant division of the involute, with the control of the v parameter limits in such manner that the surface points fall between the two limiting planes of the cutter. Using the 3rd of equations (4), for each value of φ, the limits of v can be computed.

The profile shifting factor ξ, as a function of the tool wear value δ_h, influences the shape and the limits of the edge curve.

Figure 1 shows that the translation of the rake face cone by δ_h causes a displacement of the same amount of the tooth basic plane, so that we have to expect a $\Delta R_a = \delta_h \operatorname{tg} \alpha_v$ addendum radius decrease. Due to the linear relationship between the addendum radius and the profile shift, this determines a profile shift decrease of $\Delta\xi = \Delta R_a / m$ Thus, the tool profile shifting value corresponding to δ_h becomes

$$\xi(\delta_h) = \chi = \xi s - \frac{\delta h \cdot tan\,(\gamma v)}{m} \qquad (23)$$

Accordingly, the integration limits in formula (19) will change.

For the tool edge curve, the expression of the elementary arc length in (18) is much more complicated than in the case of the involute. The complicated computational formulae would introduce significant rounding error anyway, so we start the side-edge subdivision from the equidistant subdivision of the base involute, assuming that it will not differ significantly from the equidistant subdivision due to the specific geometry of the rake face and the low rake angle value. Therefore, we assign to each of the nodes of the basic plane involute a helix of parameter P that fits the side relief face. The edge points result as intersection points of the twist line and the rake face.

The polar coordinates of the j^{th} member of the point series generated by formula (22), in correspondence with the involute geometric properties, are polar radius

$\rho_j = R_b\sqrt{1+\varphi_j^2}$, and polar angle

$\theta_j = -\eta + \varphi_j - \operatorname{arctg}\varphi_j$, The equations of the twist line from here can be written, in formal identity with equations (13). On the basis of (15), the value of the v parameter, for each j, results as

$$w_j = \frac{1}{P}\left(\left(R_a - \rho_j\right) tg\,\gamma_v + \delta_h\right) \qquad (24)$$

The edge curve calculated for the previous numerical example is illustrated in **Figure 8**.

4. CAD model in Autodesk Inventor environment

4.1. Construction of the body model

The body model is built on four involute curves created in parallel planes arranged in descending order of profile shifting. These are defined for $\xi \in \{0{,}3;\ 0;\ -0{,}3;\ -0{,}605\}$. Since the tooth surface of the tool is a helical involute surface, we computed their initial angle η. The height of the planes above the basic plane xOy was calculated using the formula given in **[3]**

$$h_i = (\xi_s - \xi_i)\, m\, ctg\, \alpha_v \qquad (25)$$

The projections of the four involutes in the xOy plane are illustrated in **Figure 9.** The generated involutes were connected using the Autodesk Inventor command „Loft". After realizing one tooth this was extended to obtain the 3D model of the shaper cutter **[8]** (**Figure 10**).

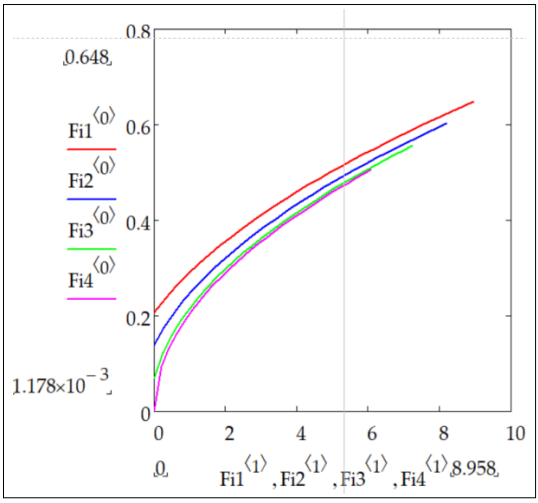


Fig. 9. *Illustration of involutes with different profile shifts.*

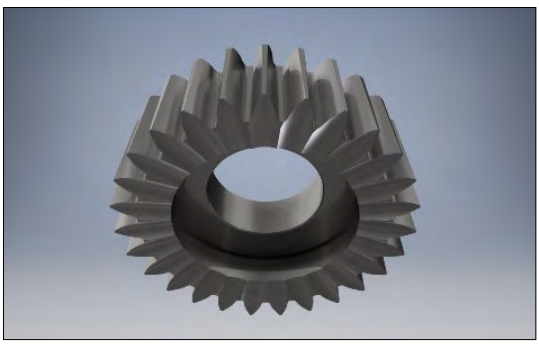


Fig. 10. *Shaper cutter 3D model.*

4.2. Function of the model

With the model, any attempt to modify both the shape and the dimensions of the rake face can be immediately and easily checked in the design process, saving precious time. Secondly, different sharpening methods and sharpening stages can be predicted, which facilitates both the design process and the prediction of the tool parameters and, based on this, the accuracy of the gears produced with a given tool. The model allows the inspection of the distribution and the maximum value of the theoretical profile error in a graphical environment.

5. Case study

Two profile error analysis methods have been developed in a CAD environment:

Method I.: in this case, the numerically generated equivalent gear involute was implemented in Autodesk Inventor and the tool edge was developed in Inventor using the specific commands of the software.

The measurement method comprised the following steps:

Intersection of the model body with the rake face. Classical conical surface, common for all teeth, or individual rake face for each tooth, can be here adopted.

Graphical output of the edge curves.

Projection of the edge curves in the plane xOy perpendicular to the axis.

Adding of the involute profile of the equivalent gear (number of teeth and module of the shaper cutter but the rack profile angle α_0) to the projection curve, ensuring the overlap of pitch circle points of both.

Now the two curves plotted before will be intersected by a tangent line to the basic circle of the equivalent gear.

The profile error equals the length of the segment cut by the two curves.

Measurements were taken on the involute inner circle (R_t) and head circle (R_a) of the Fellow's cutter.

In our study, several sharpening stages were simulated and two sharpening methods were simultaneously analyzed: classical conical sharpening and cylindrical surface sharpening [8]. The variation of the errors was studied for the variation of the rack profile angle, the sharpening stage and the rake angle. The results obtained are presented in Tables 1–4.

Table 1. *Measured profile error values, for conical rake face*

Cases	R_t [mm]	R_a [mm]
α_s, $\xi = 0.3$	0.005	0.021
α_s+10', $\xi = 0.3$	0.001	0.004
α_s−10', $\xi = 0.3$	0.015	0.033
α_s, $\xi = 0$	0.008	0.015

Table 2. *Measured profile error values, for cylindrical rake face*

Cases	R_t [mm]	R_a [mm]
α_s , $\xi = 0.3$	0.001	0.008
α_s+10', $\xi = 0.3$	0.007	0.012
α_s−10', $\xi = 0.3$	0.011	0.016
α_s , $\xi = 0$	0.003	0.003

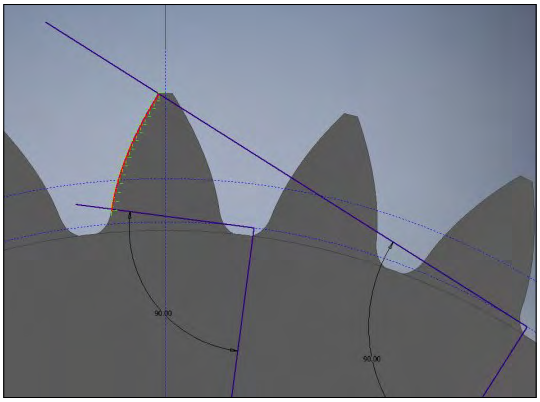


Fig. 11. *Profile error measurement Case I.*

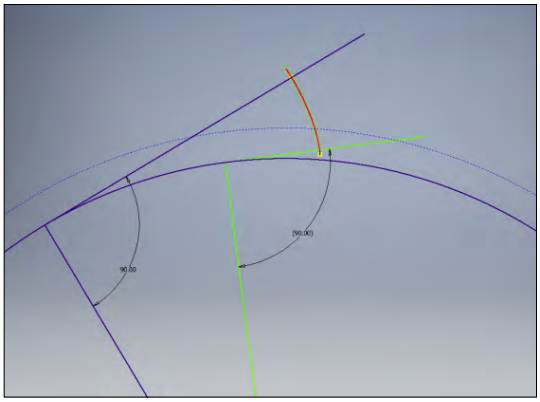


Fig. 12. *Profile error measurement Case II.*

In **Table 1** the theoretical profile error of a new shaper cutter with maximum profile shifting is shown, considering a classical circular cone rake face with $\gamma_v = 5°$. The tool's rack profile angle is set to $\{\alpha_s - 10'; \alpha_s; \alpha_s + 10'\}$ In the last row of the table, the errors were investigated for the nominal value of α_s, but for a tool profile shifting of $\xi = 0$.

In the second table, the error values were obtained in the case of applying a cylindrical rake face for each tooth. The parameters of the cylinder are radius $\rho_h = 12$ mm, and the axis inclination angle 5°.

Method II.: in this case, both the equivalent gear involute and the tool edge projection curve were numerically obtained and after that implemented in Autodesk Inventor, where and measurements were performed. In this way, we eliminated the need to build a spatial model.

Table 3. *Method II., errors on the involute inner circle*

α_s	$\gamma_v = 5°$	$\gamma_v = 10°$	$\gamma_v = 17°$
$\delta_h = 0$ mm	0.006	0.014	0.023
$\delta_h = 5$ mm	0.007	0.013	0.023
$\delta_h = 10$ mm	0.007	0.014	0.023
$\delta_h = 20$ mm	0.007	0.014	No data

Table 4. *Method II., errors on the addendum circle*

α_s	$\gamma_v = 5°$	$\gamma_v = 10°$	$\gamma_v = 17°$
$\delta_h = 0$ mm	0.019	0.038	0.066
$\delta_h = 5$ mm	0.017	0.032	0.049
$\delta_h = 10$ mm	0.016	0.026	0.033
$\delta_h = 20$ mm	0.012	0.015	No data

The measurement procedure is similar to the method outlined in Case I., except that both the control profile and the edge curve were generated numerically (**Figure 12**).

The second measuring method was used only for the case of conical rake face. We investigated the variation of the profile errors for several stages of tool wear and for a constant value of the tool's generating rack profile angle α_s.

6. Conclusions

The distribution of the errors differs from the results obtained using the pure mathematical investigation model. As illustrated in **Figure 3,** the projection of the tool edge is tangent to the control profile on the pitch circle and is located outside it. In contrast, in the Autodesk Inventor environment, in 95% of the cases, we found that only the edge projection segment between the pitch circle and the addendum circle is located outside of the control profile. The section between the involute inner circle and the pitch circle fits within the control profile in the majority of cases. Results of the present study show that the relative position of the two curves fits the mathematical model only for the case of α_s+10°, ξ(0.3) and conical rake face; however, in this case the profile error was found to be the smallest.

Looking at the results of cases I and II, it can be said that changing the tool generating rack profile angle is not advantageous in all cases for profile error evolution.

Changing the classical literature-based tool rack profile angle was useful in only one case, as it is supported by the obtained results; otherwise, increasing or decreasing it leads to an increase in the profile errors in most cases.

The investigation of the cylindrical rake face proved to be effective. The profile errors on the addendum circle were significantly smaller than those arising by the use of the conventional rake face.

With increases in rake angle γ_v the profile error gradually increases, but the chip removal capability is greatly improved. Our measurements show an increasing trend from the inner involute circle to the addendum circle.

Too high a rack angle value is detrimental to the tool-life because the possible number of re-sharpening decreases drastically.

As δ_h increases, the profile error on the involute inner circle remains constant, while it gradually decreases on the addendum circle.

References

[1] Hollanda D.: *Aşchiere şi scule.* Reprografia I. I. S. Tg. Mureş, 1994. 234–240.

[2] Szeniczei L.: *Az általános fogazás.* Nehézipari Műszaki Könyvkiadó, Budapest, 1958. 49–50.

[3] Máté M.: *Hengeres fogaskerekek gyártószerszámai.* Erdélyi Múzeum-Egyesület, Kolozsvár, 2016. https://doi.org/10.36242/mtf-12

[4] Máté M.: Az egyenesfogú metszőkerék szerszámkapcsolószögének optimálása. Műszaki Tudományos Füzetek – FMTÜ I. sz. (1996) 12–15. https://doi.org/10.36243/fmtu-1996.03

[5] Máté M., Kántor A., Laczkó-Benedek B.: *Metszőkerékkel lefejtett fogaskerekek profilpontosságának vizsgálata.* Műszaki Tudományos Közlemények, 7. (2017) 279–282. https://doi.org/10.33895/mtk-2017.07.62.

[6] Máté M., Hollanda D.: *Az egyenesfogú metszőkerék geometriai modell számítógépes kiértékelésének hibái.* Műszaki Tudományos Közlemények, 10. (2019) 53–58. https://doi.org/10.33894/mtk-2019.10.06

[7] Máté M., Hollanda D., Tolvaly-Rosca F., Forgó Z., Egyed-Faluvégi E.: *Synthesis of a Profile Errorless Involute Shaper Cutter with Cylindrical Rake Face.* In: 2019 IEEE 19th International Symposium on Computational Intelligence and Informatics and 7th IEEE International Conference on Recent Achievements in Mechatronics, Automation, Computer Sciences and Robotics (CINTI-MACRo), Szeged, 14–16 November 2019. 71–76. https://doi.org/10.1109/CINTI-MACRo49179.2019.9105302

[8] Tolvaly-Roşca F.: *A számítógépes tervezés alapjai: AutoLisp és Autodesk Inventor alapismeretek.* Erdélyi Múzeum-Egyesület, Kolozsvár, 2009. https://doi.org/10.36242/mtf-07

Műszaki Tudományos Közlemények vol. 18. (2023) 37–44.
DOI: Hungarian: https://doi.org/10.33895/mtk-2023.18.07
English: https://doi.org/10.33894/mtk-2023.18.07

RANDOM NUMBER GENERATOR

Katalin HARANGUS,[1] András KAKUCS[2]

[1] *Sapientia Hungarian University of Transylvania, Faculty of Technical and Human Sciences, Târgu-Mureș, Romania, katalin@ms.sapientia.ro*

[2] *Sapientia Hungarian University of Transylvania, Faculty of Technical and Human Sciences, Târgu-Mureș, Romania, kakucs2@ms.sapientia.ro*

Abstract

Illustration plays an important role during education: The Galton board is a suitable tool for illustrating random processes and explaining probability distributions. We have created this tool in a virtual version, which facilitates data collection for statistical processing of experimental data and also enables the study of non-symmetrical distributions. The random processes on the device are simulated, which requires a random number generator. Since there were some doubts about the software-generated pseudo-random numbers, we created a true random number generator based on the input noise of microcontrollers.

Keywords*: random number generator, Galton board.*

1. Introduction

During the education of university students, we would like to provide them with fundamental knowledge from the field of probability theory and statistics. At least as much as clarifies the nature of random phenomena and introduces practical, numerical relationships that describe them. Since this field belongs to mathematics, its thorough understanding and in-depth study require a mathematical way of thinking. Hence comes the biggest challenge in accomplishing this task: the engineering way of thinking is not the same as that of a mathematician. While the latter usually does not tie concepts to the real world, the engineer practically lives and breathes it. Experience shows that even at a basic level, the theoretical math knowledge of engineering students is incomplete, and they do not really see the need to fill this gap. The situation is even worse for students in non-engineering majors, where there is no math education at all. Therefore, in teaching subject areas that require abstract thinking, the use of illustrative tools that connect abstract concepts to reality plays a crucial role **[1, 2]**. For example, demonstrating the operation of various algorithms using software and hardware tools. Based on this, the idea arose that in clarifying the nature of random processes, which is followed by establishing quantitative relationships, we could "rely" on the Galton board.

So a virtual Galton board was developed. The original version of the Galton board is a tool used to illustrate normal distribution. The virtual version simulates the occurring natural phenomenon, and the user can intervene in it, allowing for different distributions to be visualized. The simulation is done using computer tools, which not only makes the phenomenon observable but also allows for data to be collected "on the go", which can later be processed and analyzed by computer program.

The virtual Galton board simulates random phenomena, so a random number generator is needed. Previous experience has shown that working with pseudo-random numbers generated by algorithms doesn't always lead to the expected results. For example, even when a few hundred consecutive random numbers are generated, they don't always distribute uniformly. To overcome this problem, the aim was to produce true random numbers using hardware, which didn't incur additional costs in building the virtual Galton board, as it was achieved through a microcontroller that controlled it. This article describes the implementation of this idea.

The didactic tool itself was presented and published in our presentation "Didactic Tool for Intuitive Random Process Visualization" at the "25th International Conference on Interactive Collaborative Learning" held in Vienna in late September 2022. The second part of this article provides a brief summary of what was discussed at the conference.

2. Generating random numbers

Generating random numbers is usually simple in most programming environments because there is typically a function [3], specifically designed for this purpose, such as:

- in C (and therefore when programming with Arduino) and C++, the *rand()* function returns an integer between 0 and *RAND_MAX*, where the upper limit depends on the programming environment but is guaranteed to be at least 32767;
- in Visual Basic the *Rnd()*, function represents a random number from the [0, 1) interval;
- in Excel the *RAND()* function returns a random number from the [0, 1) interval, or the *RANDBETWEEN(bottom, top)* function returns a random integer in the [*bottom*, *top*] interval.

These functions use some algorithm to determine the next value of the nominally uniformly distributed random number, which is why they are computed rather than truly random. They are called pseudorandom numbers. In practice, consecutive numbers appear random, but they always follow the same sequence. To reduce the likelihood of repetition, there is usually an option to set the starting point of the sequence. In C, there is the *randomSeed(x)* function, while in Visual Basic, there is a *Randomize(x)* statement, both of which set the beginning of the sequence to the number given as the parameter *x*. Of course, if the same parameter is used every time, the sequence will always start in the same place, so starting the sequence from a location determined by the computer's clock, for example, can make it more random (the parameter could be the number of seconds elapsed since midnight, for instance).

If we need true random numbers, we can generate them using some physical device [3]. This can be, for example, dice or a hardware random number generator. Their operation is not based on an algorithm, but on some physically random process. The RPG100 integrated circuit available on the market generates 16-bit random numbers based on semiconductor noise, but this circuit needs to be connected to a computer.

Looking for a simpler and cheaper solution, we built our own version based on the noise of the analog input of microcontrollers.

Although it is slower than the circuit mentioned as an example, theoretically we can generate random numbers with any resolution using it. The implementation is based on the easily programmable and computer-connectable microcontroller Arduino and ESP development boards available to anyone: we tested the Arduino Mega [4] and the ESP32-es [5] versions. The main difference between the two is the speed, with the ESP32 being significantly faster.

The operating principle is based on sampling the randomly fluctuating voltage of the microcontroller's "floating", unconnected inputs. To clarify the input state, it is usually connected to the ground with a larger resistor, or in the case of digital inputs, to the positive supply voltage. These resistors ("pull down", "pull up") are sometimes built-in elements of the microcontroller, which can be enabled through software, otherwise we have to connect them to the circuit ourselves. To make the input state random, we don't need to connect anything to it, and we have to bypass the internal pull down/pull up resistor (if there is one).

If we sample a floating digital input, its state will randomly alternate between LOW and HIGH. If we sample an analog input, the sampled voltage value will randomly fluctuate within the range of the smallest and largest possible values.

In both cases, the obtained signal depends on the noise voltage of the corresponding input. There are several explanations for what the source of this noise could be, probably due to the combined effects of these sources. One source is the thermal noise of the semiconductors (and conductors and dielectrics), which is also the basis of the RPG100 integrated circuit mentioned above. Another important component is "atmospheric noise", which can be heard as radio static from an untuned station. The latter has several components, stemming from natural phenomena and human activity. Some of the components from human activity can show regular repetition, such as the 50 Hz noise of the electrical network, so it does not develop completely randomly.

The third source of input noise in microcontrollers is, according to our experience, the microcontroller itself and the surrounding circuit. Thus, the 32nd input of the Wemos Lolin32 clone we tested (which is a development board built on an ESP32) is particularly noisy, much noisier than

the other inputs, which is probably due to some sort of design flaw.

The Arduino samples the voltage of the analog input with 10-bit resolution, while the ESP32 with 12-bit resolution, which can be read as an integer using the *analogRead()* function.

If nothing is connected to an analog input, the length of the wires that act as an antenna for environmental, atmospheric noise will be minimal. In this case, the sampled voltage changes chaotically, depending on the noise of the development board components, so there is little regularity in it. According to experience, the noise voltage changes too slowly, so there is no significant difference between the magnitude of two consecutive signals. If we used this sampled voltage to generate a random number, there would not be a significant difference between consecutive numbers.

To generate true random numbers with our random number generator, we utilize a flaw in the sampling circuit (the analog-to-digital converter) where the last one or two bits of the number representing the signal magnitude obtained by sampling even in the case of a stable voltage measurement, are fluctuating. This helps to eliminate any regularly changing components of the noise.

The operating principle can be algorithmically described as follows:

– we sample the floating analog input;
– we determine the last one or two bits of the obtained digital signal;
– we add these bits to a sequence of bits;
– when this sequence is long enough, we interpret it as a random number (e.g. with 16 bits, we get a random number between 0 and 65535).

In our program written in C and applied to Arduino and ESP32, with keeping only the last bit, it looks like this **[6]**:

```
B = byte(analogRead(Pin0) &
0b00000001) |
(byte(analogRead(Pin1) &
0b00000001) << 1) |
...
(byte(analogRead(Pin7) &
0b00000001) << 7);
```

where *analogRead(PinX)* is the value read from the analog input, *0b00000001* is a mask (with only the last bit being 1). The & operation is used to apply this mask (bitwise "and"), which results in only keeping the last bit of the read value (the rest is zeroed out). This bit is shifted left by << *n* (where n determines where the kept bit should go), and then the result is converted to a one-byte number (*byte(...)* – this is necessary because the value returned by the *analogRead() function is not a one-byte data)* –, then with the | operation (bitwise "or"), it is inserted into the already created sequence. The result will be one byte. If a two-byte number is needed, it can be obtained from two bytes with the $B1 + 256 \times B2$ operation.

To speed up the generation of random numbers, we sample eight different analog inputs simultaneously to create the eight random bits.

If we only need a single random "yes/no" value (such as in building a virtual Galton board), then the process is even simpler: we just need to observe whether the integer value read from the analog input is even or odd, and the generated random number is the remainder of dividing this value by two, either 0 or 1.

3. Testing the hardware random number generator

During testing, the program generates two one-byte data in each cycle, from which a random integer between 0 and 65,535 is calculated. Experience shows that after sampling, a little time needs to elapse for the circuits to return to their initial state, otherwise the obtained quantities do not behave randomly. If we were to solve the random number generation with a single input sampling, the process would be very slow due to the waiting time between determining two consecutive bits. Therefore, we sample eight different inputs in sequence without waiting, hoping that they do not affect each other. We briefly interrupt the program execution between the determination of the bytes (8-8 bits). From these numbers, we generated sequences, patterns of 16,384 elements. By sorting these patterns in increasing order, we obtained the graphs shown in **Figure 1.**

We found that the length of the waiting time between consecutive samplings affects the quality of the acquired dataset. Without waiting, the datasets contain a large number of zero values. A waiting time of 2 ms already improves the situation, but both the ESP32 and the Arduino Mega show strongly nonlinear behavior.

A 10 ms wait already significantly improves the situation and seems to be sufficient for the Arduino. The clock frequency of the Arduino is an order of magnitude lower than that of the ESP, so the delay due to the longer cycle time is also added to the 10 ms.

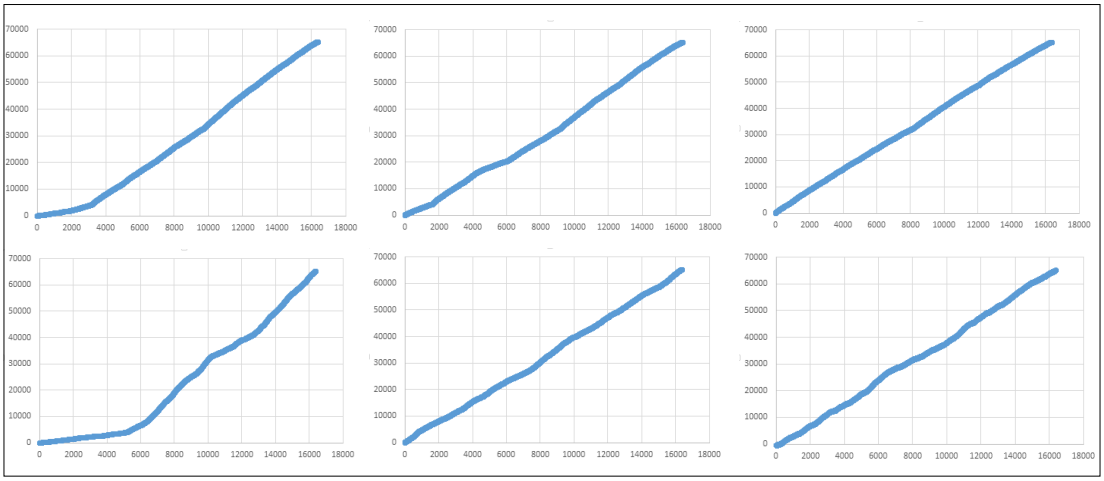


Fig. 1. *N Increasingly ordered samples.*
Top: ESP32, Bottom: Arduino Mega. From left to right: 2ms, 10ms, and 25ms delay between two consecutive samples. The horizontal axis shows the index of the generated random number, while the vertical axis shows its value.

A 25 ms wait leads to a linear approximation for the ESP32 as well.

In the case of the Arduino Mega, we obtained a better quality data set than the one sampled with a 25 ms wait presented in **Figure 1** (even better than the example shown with a 10 ms wait), which we analysed further for two reasons:

– because it belongs to the weaker quality samples obtained;
– because it was easier to build our demonstration circuit with this circuit.

During the analysis, we calculated the empirical average and standard deviation of the sample: $m = 32518{,}35$, $\sigma = 19234{,}06$. Since the obtained numbers should have a uniform distribution between 0 and 65,535, the theoretical average would be 32,767.50, and the standard deviation would be 18,918.32. We observe that these are relatively close to each other.

We calculated the empirical distribution function of the sample, as well as the value of the theoretical distribution function calculated for the sample members. Based on these two sets of numbers, we performed a goodness-of-fit test using the χ^2 test (*CHISQ.TEST*) in Excel, and the returned value was 1, indicating perfect fit (even though we can see that it is not exactly perfect), so we can say with a high degree of confidence that the generated random numbers are uniformly distributed.

Due to the suspiciously good fit, we also performed another non-parametric test, the Kolmogorov–Smirnov test. This test decides whether the empirical and theoretical distribution functions can be considered identical or not based on the largest difference between them. The change in the absolute value of D_i is shown in **Figure 2.** The maximum value of D_i must be compared with a critical value, which depends on the level of confidence and the number of samples: if it exceeds the critical value, then it cannot be accepted at the given level of confidence that the sample has the specified theoretical distribution

The highest value of D_i should be compared with a critical D_{cr}, value dependent on the level of reliability and the number of samples: if it exceeds the critical value, then it cannot be accepted at the given level of reliability that the sample follows the theoretical distribution.

In the case of our sample, which consisted of 16,384 numbers of not very good quality, the K-S test led to a result that contradicted the χ^2 test at the usual levels of confidence. Therefore, we could not prove that the distribution is uniform. However, if we perform the K-S test not with the distribution function determined by the elements of the sample, but with the histogram obtained by grouping them into intervals, then a positive result can be obtained as well.

As a check, we performed the same calculations with pseudorandom number sequences generated in Excel, and the results were similarly good or even better.

However, there was a criticism regarding the software-generated sequences, that if they are not long enough, they may show a distribution far from uniform. Therefore, we divided our hardware-generated number sequence into smaller subsamples, shorter number sequences, in their order of creation, and examined how their empirical mean and variance behave. In **Figure 3** we can see the case of our number sequence divided into 256 subsamples of 128 elements, where the empirical mean varies between 28,307.41 and 36,125.93.

The greatest deviation from the mean of the entire sample is shown by the value of 28,307.41 (in the 92nd subset). The graph of the ordered numbers in this subset can be seen in **Figure 4.** According to the chi-squared test and the K-S test (which was performed with 128 numbers and not using a histogram approximation) conducted on this subset, the hypothesis of a uniform distribution can still be accepted with very high probability.

Additional tools are often used to examine the uniformity of possible repetitions and the distribution of consecutive (i.e., non-ordered) numbers.

A commonly used solution is the visual representation of number sequences bit by bit, where the 0 bits are black and the 1 bits encode a white pixel in a black and white bitmap. If the distribution is truly uniform, then this image should be uniformly grey and should not show any repetition or pattern. We applied this principle a little differently by assigning not one bit but one byte to each pixel, so the pixels in the resulting image can have 256 different colours. This image should also be free of repetitions and patterns and should be uniformly grey, so there should be no red or blue shaded spots on it, for example. For the described sample, we obtained a 256×128

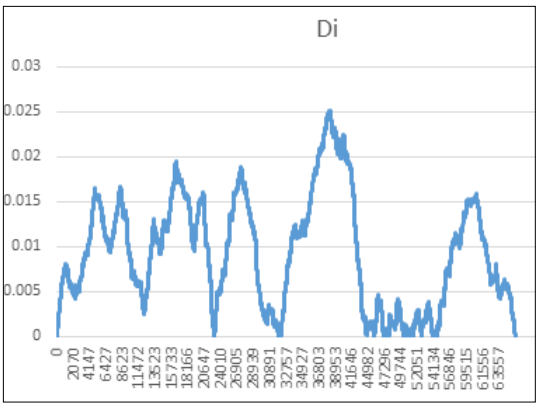


Fig. 2. *The absolute magnitude of the difference between the empirical and theoretical density functions.*

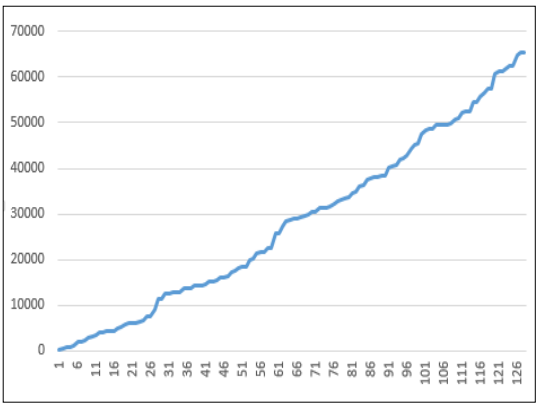


Fig. 4. *Graph of the elements of the sub-sample showing the largest deviation from the average.*

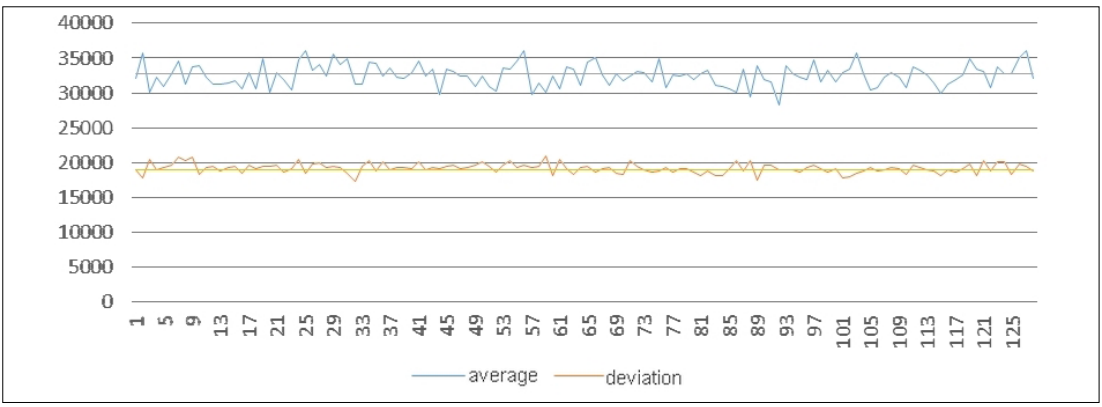


Fig. 3. *The variation of empirical mean and standard deviation around the theoretical values for the sample divided into 128-element strings.*

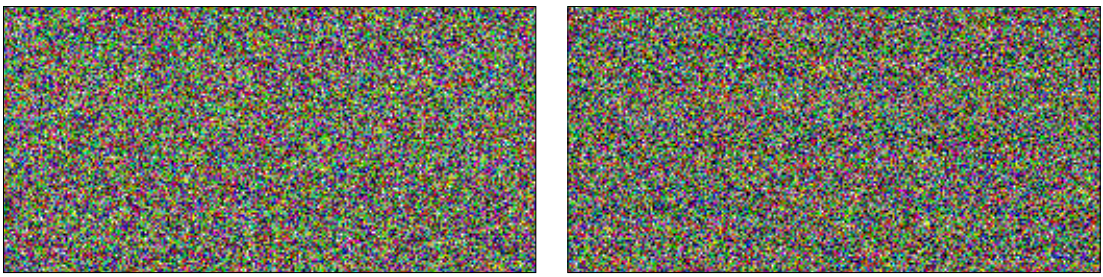


Fig. 5. *Bitmap representation of random numbers (left: Arduino Mega, right: ESP32).*

pixel sized, bitmap-format image on the left side of **Figure 5.** On the same figure, the rectangle on the right shows the result obtained with the number sequence generated by ESP32 with a waiting time of 25 ms. By examining the images, we can be sure that the obtained number sequences are indeed uniformly distributed.

4. The Galton Board as a didactic tool

In 1889, Francis Galton, an English polymath, built a device known as the Galton board to illustrate random processes intuitively. The Galton board consists of a sloping board with nails arranged in a chessboard-like pattern at equal distances from each other. At the bottom of the board, compartments are formed where the disks or balls launched from the same point above collect.

As the disk slides downward, it collides with the first nail, which prevents it from freely sliding further. Here, the disk bounces once and randomly, with probability p or q, it will slide to the right or left side of the nail, respectively, and continue downward until it reaches the next row of nails. In the "classic" version $p = q = 0.5$. Therefore, if we slide enough disks down, approximately half of them will end up on the first nail of the second row, and half will fall on the second nail.

In the second row, the left or right deviation is repeated. The probability of deviation to the right or left is the same, so the probability of further progress along any possible path is halved. However, the middle nail in the second row can be reached by the sliding disk in two different ways, so the probability of reaching it is the sum of the two appropriate probabilities (**Figure 6**).

The situation becomes more complicated in the third and subsequent rows. The calculations necessary for explanation can be made more manageable by constructing the Pascal triangle in Excel (**Figure 7**).

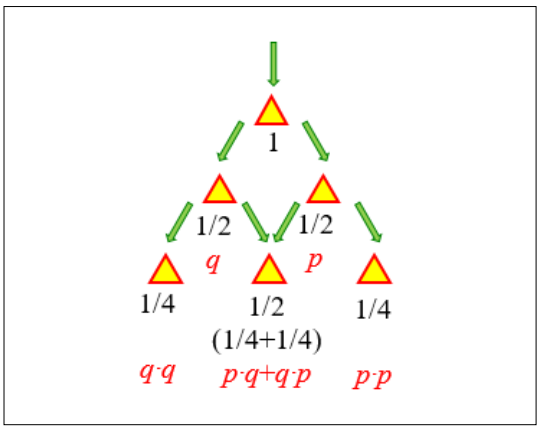


Fig. 6. *Collision with the second nail.*

5. The virtual Galton board

Our visualization tool is a virtual Galton board (**Figure 8**). In reality, it is a tangible device that simulates the phenomena that occur on the actual board. The device's "heart" is an Arduino Mega 2560 development board, to which we connected a 32×32 RGB LED matrix. The "disk" is a pixel (a coloured LED on the board), which falls downward at a constant (user-modifiable) speed. If it hits a "nail" (which is also a pixel) during its descent, we use a random number generator to determine which direction it should deviate (0 – left or 1 – right).

The probability of the disks deviating left or right on the original Galton board can be 0.5, but in our upgraded version, we can also adjust it to match the tilt of the stand, allowing us to study the general binomial distribution, not just the symmetric case corresponding to $p = 0{,}5$. For example, if we raise the left side of the device, the disks are more likely to fall to the right, and their distribution becomes skew.

The angle of deviation α of the stand from the vertical position is measured by an analog MMA7261QT accelerometer sensor.

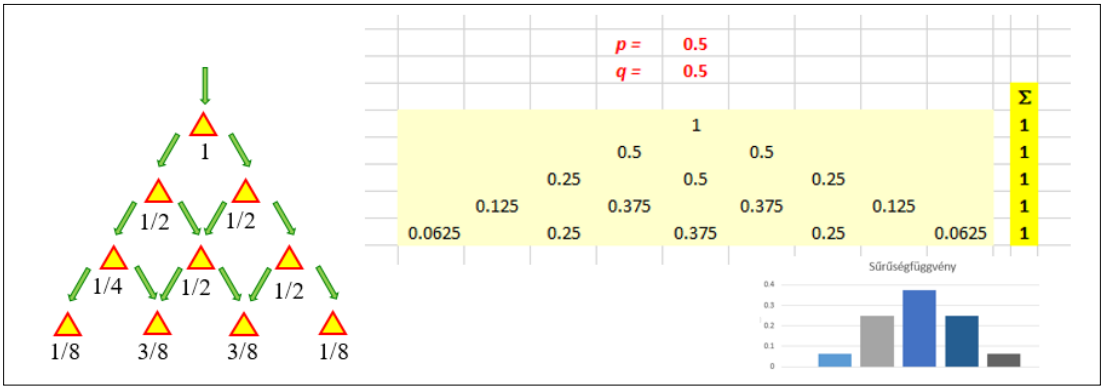


Fig. 7. *Collision with further nails.*

If the Galton board is not tilted, the disks bounce to the right or left with equal probability. If we tilt the board, this probability changes to:

$p = 0{,}5 + \sin \alpha$.

This formula only makes sense if α is in the [–30°, +30°] interval; if $\alpha < -30°$, then $p = 0$, and if $\alpha > +30°$, then $p = 1$.

Only a limited number of disks can fit in the holder, so if it gets full, we need to scale down the column heights. Therefore, after a while, the number of disks no longer corresponds to the number of illuminated pixels.

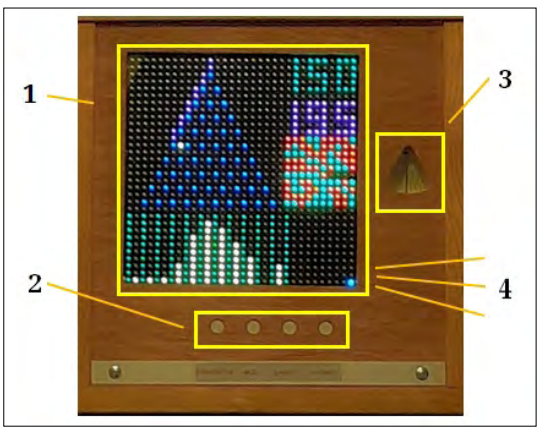


Fig. 8. *The virtual Galton board*
1- 32×32 LED panel
2- Control buttons, in order: start/stop, mode selection, animation slowdown, animation speedup;
3- Inclinometer showing the probability p calculated from the sine of the tilt angle;
4- Mode indicator:
Top LED: true random numbers,
Middle LED: pseudorandom numbers,
Bottom LED: pseudorandom numbers influenced by the tilt angle.

After every hundredth rolling ball, the Arduino sends the number of collected balls in the holder through the serial port, which can be copied from the "Serial monitor" window of the Arduino IDE program for further analysis. We can draw histograms of the frequencies, calculate the empirical mean and variance of the distribution, and perform statistical tests.

6. Conclusions

Discussions on easily constructible hardware random number generators often refer to the noise of floating (unconnected) inputs of microcontrollers, from which randomly varying data can be obtained through sampling. If we need random numbers in programming a microcontroller system, we don't really need separate devices, just the code that generates the random numbers. However, based on our experience, the state of an input, whether it's analogue or digital, changes randomly but too slowly, so the distribution of consecutive random data will not be satisfactory. Therefore, we further developed the basic idea by working with the least significant bit of randomly sampled analogue input values, and if we need random numbers from a certain range, rather than simple "yes/no" outputs, we combine the resulting sequence of bits.

Experience has shown that if the analogue inputs are sampled too frequently, the result obtained will not be as desired: most likely, the sampling circuit will not have enough time to return to its initial state. If we want to generate a long sequence of random numbers, reading the random bits from a single analogue input would be a time-consuming process due to the insertion of gaps between samples. Therefore, to determine a byte of a random number, we read the state of

not one, but eight inputs at the same time (actually one after the other, without inserted interruptions).

The numbers obtained in this way have shown a uniform distribution. Based on these, normally distributed numbers can be generated using the Box-Müller method.

References

[1] Mcloughlin C., Krakowski K.: T*echnological Tools for Visual Thinking: What Does the Research Tell Us?* Paper presented at the Apple. University Consortium Academic and Developers Conference, James Cook, 13.1–13.12 (2001)

[2] Khasanovna U. N.: *The Importance of Didactic Tools in the Teaching of Educational Science*. Academicia Globe: Inderscience Research, 2. (2021) 76–79.

[3] Johnston D.: *Random Number Generators—Principles and Practices: A Guide for Engineers and Programmers.* Walter de Gruyter GmbH & Co KG, 2018

[4] *** Arduino Mega Rev. 3. Documentation, https://docs.arduino.cc/hardware/mega-2560

[5] *** ESP32 Reference, https://docs.espressif.com/projects/esp-idf/en/latest/esp32/hw-reference/index.html

[6] *** Arduino Language Reference, https://www.arduino.cc/reference/en/

Műszaki Tudományos Közlemények vol. 18. (2023) 45–48.
DOI: Hungarian: https://doi.org/10.33895/mtk-2023.18.08
English: https://doi.org/10.33894/mtk-2023.18.08

A MACRO- AND MICROSCOPIC INSPECTION OF THE SAFETY CRITICAL COMPONENTS

Huszák Csenge,[1] Schramkó Márton,[2] Kovács Tünde Anna[3]

[1] *Óbuda University, Doctoral School on Safety and Security Sciences. Budapest, Hungary, huszak.csenge@bgk.uni-obuda.hu*

[2] *Óbuda University, Doctoral School on Materials and Technologies Sciences. Budapest, Hungary, schramko.marton@bgk.uni-obuda.hu*

[3] *Óbuda University, Bánki Donát Faculty of Mechanical and Safety Engineering, Institute of Materials and Manufacturing Science, Department of Materials Technology. Budapest, Hungary, kovacs.tunde@bgk.uni-obuda.hu*

Abstract

In many cases related to critical infrastructures, we find equipment that is extremely important from a safety point of view and key safety-critical components in them. Such parts are turbine blades, engine elements of jet aircraft, as well as high-pressure steam lines of power plants and pressure equipment for liquid gases. In the case of safety critical components, non-destructive and the destructive tests are indispensable. The damage to these components can cause catastrophic failures that need to be avoided. During the inspection of the components, microscopy tests give information about macrostructure and microstructure. Carrying out the tests consists of several steps, during which it is necessary to follow the relevant standards. During the examination of macrostructure, conditions that may lead to premature failure and breakage of the component can be detected. Getting to know the microstructure can provide information about manufacturing, welding, or microstructural defects, which can be a key issue from the point of view of safe operation.

Keywords: *microscope, destructive testing, welding, microstructure.*

1. Introduction

Numerous safety-critical components are known, and their failure can lead to disasters. Currently, there are no specific regulations for the inspection of these components. In many cases, components containing welded joints, pipelines, or turbine blades are considered safety critical. There are regulations for inspecting welded joints, such as examinations targeting macrostructure and microstructure analysis. In this paper, we aim to analyse the sample preparation for inspections used in welded structures and emphasize the importance of training the personnel responsible for performing this task.

2. Sample Preparation

During the preparation the size and shape of the sample are crucial. The object to be examined may be too small to grind the surface (e.g., the cross-section of a needle or a wire) or too large (e.g., a welded piece). If the object to be examined is small in size, an embedded sample of the appropriate size can be prepared, allowing us to further prepare the surface. However, if it is too large, it is necessary to create a representative sample specific to the material under investigation [1, 2, 3]. During the cutting process, it is important to ensure that the material structure of the sample is not altered. Therefore, cutting should not cause the specimen to heat up or undergo plastic deformation, as both can lead to changes in the material structure and yield false examination results. When examining welded joints, it is advisable to follow the recommendations of the standard regarding the location of sample extraction [4, 5]. Cutting can be performed using water jet cutting or machining. In practice, specialized laboratory cutting equipment can also be used, typically employing a diamond abrasive cutoff wheel, adjust-

able speed, and water cooling to perform the sample cutting. For thin sheets and foils, cutting with scissors is also an option. The cut surface will inevitably undergo deformation, so it is necessary to further work on the cut surface by grinding whenever possible [1, 2, 3].

3. Specimen mounting

Mounting refers to encapsulating the machined sample or small-sized specimen in a polymer mounting material. Depending on the type of resin used and the nature of the sample, cold or hot mounting methods can be applied. The primary objective here is to ensure that neither the structure nor the chemical composition of the sample is altered during the mounting process [1, 2, 3]. In cold mounting, a two-component resin is typically used. During the curing of the resin, minimal reaction heat is generated, but this does not cause structural changes in the case of metals. To prevent any accidents, it is necessary to follow the recommended proportions and mixing sequence provided by the manufacturer of the mounting material in all cases. Figure 1 shows the samples prepared using the cold mounting method.

Hot mounting is performed by heating and polymerizing the powder mounting material (such as Bakelite, epoxy, acrylic, etc.). The temperature for hot mounting can reach up to 200°C, depending on the specific mounting material. However, it is crucial to consider that for certain metals with low melting temperatures, this temperature range can cause changes in the microstructure, such as recrystallization.

4. Grinding

The surface of the sample is ground using progressively finer abrasive wheels. The abrasive material used can be silicon carbide-based or diamond grinding discs. Abrasive materials are commercially available in various grit sizes ranging from P60 to P4000. Abrasive materials with grit sizes of P2000 to P4000 are considered polishing agents as well. Whether manual or automated grinding is employed, during each grinding step, the sample is rotated 90 degrees to remove the previous grinding scratches. Thus, grinding eliminates the previous grinding scratches with perpendicular and finer scratches (see Figure 2). Grinding is performed with a continuous water supply, where the water helps remove loose particles, debris, and heat generated by friction.

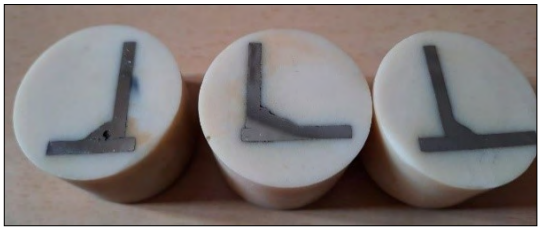


Figure 1. *Samples using cold mounting method.*

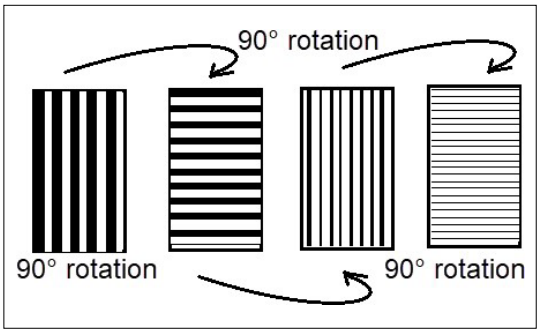


Figure 2. *Steps of grindings.*

5. Surface polishing

The next step is to polish the surface prepared to the desired level of smoothness. Polishing can be performed through mechanical, chemical, or electrolytic methods. Mechanical polishing is carried out using a polishing disc or a polishing cloth. The material of the polishing cloth can be natural or synthetic. For polishing, a polishing paste or a distilled water suspension containing aluminium oxide, magnesium oxide, or diamond particles ranging from 0.05 to 15 μm in size, can be used, which should be compatible with the material and the hardness of the polishing cloth. During manual polishing, the sample should be moved in a circular motion opposite to the direction of rotation on the polishing disc. After polishing, the surface is washed with distilled water followed by alcohol. The applicability of alcohol should be checked for non-metallic samples. The surface is then dried with warm, blown air. Once the surface is polished, macroscopic examinations can be conducted. At 50 times magnification, certain cracks, welding defects, and inadequate fusion become visible. Figure 3 shows a polished and etched sample obtained from a welded joint. In the case of cast iron, the graphite or certain non-metallic inclusions can be recognized, and their shape, size, and distribution can be examined.

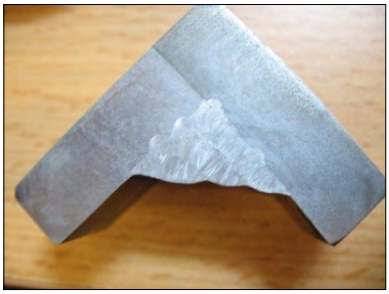


Figure 3. *Welded joint, polished and etched.*

6. Surface Etching

More detailed examinations are conducted by etching the surface following the recommendations of ISO 16060 [7]. The most used etchants are summarized in **Table 1** .

The etchant should be selected according to the quality of the material being examined. The etchant selectively attacks the sample surface to varying degrees at grain boundaries and grain surfaces, making them visible under a microscope. Altered microstructures resulting from heat treatment can also be effectively examined, especially in the case of tool steels **[7]**. Colour etching techniques provide more information and enhance recognition. The etchant forms a stable-coloured film on the sample surface, which can be an oxide, sulfide, chromate, or complex compound. Strict adherence to the mixing order and ratios is essential during etchant preparation.

Table 1. *Etchants.* **[1, 2, 3, 6]**

Name	Etchant	Application area
Nital	97 ml ethyl alcohol; 3 ml 69% nitric acid	Unalloyed and low-alloy steel
Pikral	100 ml ethyl alcohol; 4 g picric acid	Unalloyed and low-alloy steel
Hydrochloric acid-iron chloride	960 ml ethyl alcohol; 20 ml 35% hydrochloric acid; 50 g iron(III) chloride	Copper and its alloys
Keller	950 ml distilled water; 25 ml 69% nitric acid, 15 ml 35% hydrochloric acid; 10 ml 40% hydrofluoric acid	Titanium and its alloys, aluminium and its alloys
Kroll	10 ml 40% hydrofluoric acid; 30 ml 69% nitric acid; 960 ml distilled water	Aluminum and its alloys

For chemically resistant materials, electrolytic etching can be used. In conventional etching, the sample surface is immersed in the etchant and gently moved. The etching time also depends on the microscope resolution, with longer durations required for lower-resolution microscopes and shorter durations for higher-resolution microscopes. After etching, the sample is rinsed thoroughly with water, followed by rinsing with alcohol, and then dried with warm blown air. The microscopic image of an etched surface of a copper sample is shown in **Figure 4**.

During the welding process in the case of the austenitic microstructure steels (ex. X5CrNi18-10, 1.4301) some precipitation in the heat-affected zone can be established, which decreases the corrosion resistance **[8, 9]**. Precipitations formed at the grain boundaries can be identified by metallographic examination (**Figure 5**) **[10]**.

7. Conclusions

The macroscopic and microscopic examinations presented in this study are widely used in materials testing. **Figure 6** summarizes the steps involved in the examination of macro- and microstructures.

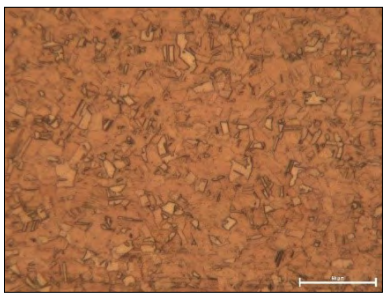


Figure 4. *Etched copper sample (etched with hydrochloric acid-iron chloride).*

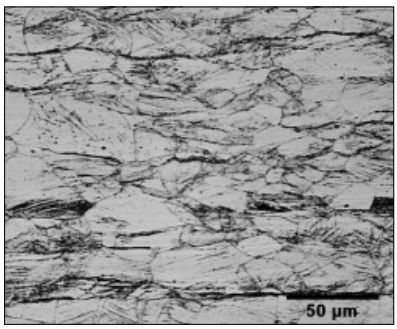


Figure 5. *Ausztenites rozsdamentes acél szemcsehatár-kiválások* **[10]**

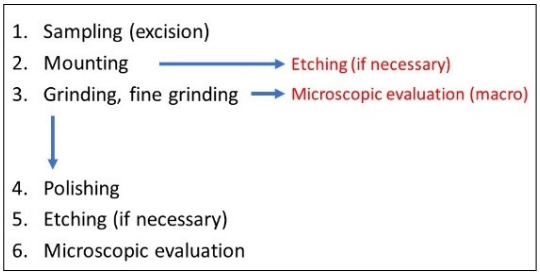


Figure 6. *Steps in macroscopic and microscopic examination*

These examinations must be conducted by properly trained personnel in materials testing who possess a high level of knowledge in both sample preparation and evaluation of the tests. The preparation of samples is typically a task performed at the laboratory technician level while conducting macroscopic examinations requires more knowledge and practice. Determining the microstructure of materials assumes a high level of expertise in materials science and practical experience. Naturally, traditional testing techniques can be complemented with modern methods offered by advanced equipment, such as scanning electron microscopy (SEM) or X-ray diffraction (XRD) for phase identification. For the examination of welded structures, the European Welding Federation (EWF) has established a specialized training system (EWF-627-07) for personnel responsible for the metallographic examination of structural materials and their joints produced by welding and related processes **[11]**.

Acknowledgment

The following work was carried out in the framework of the project registered under Erasmus+ 2021-1-RO01-KA220-VET-000025835, the authors want to declare acknowledgement for the support.

References

[1] Bramfitt B. L., Benscoter A. O.: *Metallographer's Guide, Practices and Procedures for Irons and Steels*. 1. kiadás. ASM International, Materials Park, 2002. 169–244.

[2] McCall J. L., Mueller W. M.: *Metallographic Specimen Preparation, Optical and Electron Microscopy*. 1. kiadás. Plenum Press, New York, 1973. 1–348.

[3] Vander Voort G. F.: *Metallography Principles and Practice*. 4. kiadás. ASM International, Materials Park, 2007. 1–258.

[4] MSZ EN ISO 17639:2022 Fémek hegesztett kötéseinek roncsolásos vizsgálatai. Varratok makroszkópos és mikroszkópos vizsgálata 2022.

[5] ISO 15614-1:2017 Fémek hegesztési utasítása és hegesztéstechnológiájának minősítése. A hegesztéstechnológia vizsgálata.

[6] ISO/TR 16060:2003 Fémes anyagok hegesztéseinek roncsolásos vizsgálata. Marószerek a makro- és a mikroszkópos vizsgálatokhoz.

[7] Tóth, L., Fábián, R.: *The Effects of Quenching and Tempering Treatment on the Hardness and Microstructures of a Cold Work Steel.* International Journal of Engineering and Management Sciences, 4/1. (2019) 286–294. https://doi.org/10.21791/IJEMS.2019.1.36.

[8] Dománková, M., Bártová, K., Slatkovsky, I., Pinke, P: *High Nitrogen Austenitic Stainless Steel Precipitation during Isothermal Annealing*. Materials Engineering - Materiálové inžinierstvo 23 (2016), 84–89.

[9] Dománková, M., Kocsisová, E., Pinke, P. Slatkovsky I.: *Effect of Deformation on Sensitisation Process in Austenitic Stainless Steel AISI 316*. Materials Science and Technology, 13 (2013) 1–10.

[10] Dománková, M., Adamech, M., Petzová, J., Bártová, K., Pinke, P.: *Microstructure Characteristics of Borated Austenitic Stainless Steel Welds*. Research Papers Faculty of Materials Science and Technology in Trnava, Slovak University of Technology in Bratislava, 26/43. (2018) 45–52. https://doi.org/10.2478/rput-2018-0029.

[11] EWF-627-07 Guideline (2007)

Műszaki Tudományos Közlemények vol. 18. (2023) 49–52.
DOI: Hungarian: https://doi.org/10.33895/mtk-2023.18.09
English: https://doi.org/10.33894/mtk-2023.18.09

FLOOD PROTECTION PARTIAL ANALYSIS OF THE OLT BRIDGE IN SÂNTIMBRU

Zsombor KISFALUDI-BAK,[1,2] F.-Zsongor GOBESZ[3]

[1] *Transylvanian Museum-Society, Department of Technical Sciences, Cluj-Napoca, Romania, kisfaludi.zsombor@eme.ro*

[2] *Technical University of Cluj-Napoca, Department of Structural Engineering, Cluj-Napoca, Romania, zsombor.kisfaludi@mecon.utcluj.ro*

[3] *Technical University of Cluj-Napoca, Department of Structural Engineering, Cluj-Napoca, Romania, go@mecon.utcluj.ro*

Abstract

In many places in Transylvania, we can observe the ruins and remains of buildings, from which many interesting conclusions can be drawn. As part of a more comprehensive research, in which the flood protection safety of old Transylvanian bridges is examined, in this case the rebuilt road bridge in Santimbru was investigated. Modelling and hydraulic calculations performed on the basis of in situ measurements can reveal the old hydraulic construction requirements used in the design of the bridge, and then the data were also considered in the light of today's standards.

Keywords: *technical history, bridge design, flood .*

1. Introduction

The examined bridge over the Olt River in Sântimbru (Csíkszentimre in Hungarian) is located on the county road number 133 in Harghita County, Romania, and probably adds to the list of bridges that were blown up in September 1916 by the retreating Austro-Hungarian army. The original design plans of the bridge are not known, but the structure is not unique on the Olt River, similar bridges were built in Sâncrăieni (Csíkszentkirály), Sfântu Gheorghe (Sepsiszentgyörgy) and other places. A postcard illustrating the original state of the bridge is shown in **Figure 1**. **Figure 2** proves that the railing at the bridgehead, which is completely different from the current railings on the bridge's superstructure, corresponds to the one shown on the old postcard.

According to the Hungarian text carved on a plaque on the bridgehead (**Figure 3**): „The renewal of the Olt bridge was carried out in 2016 by the town hall of the settlement, with the financial support of the Harghita county council,

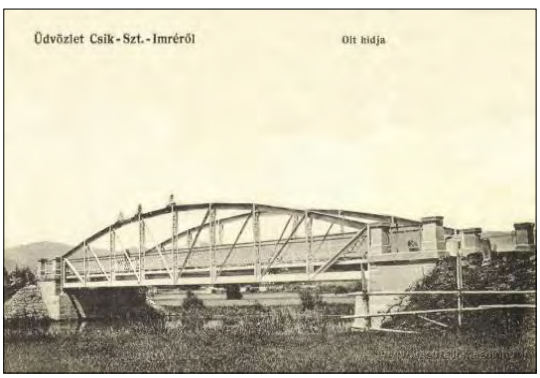


Fig. 1. *The image of the old Olt bridge in Santimbru on a postcard from 1912.* **[1]**

Fig. 2. *The current state of the old bridgehead.*

Fig. 3. *The plaque on the bridgehead.*

Fig. 4. *The road bridge built over the widened riverbed on the site of the old arched bridge.*

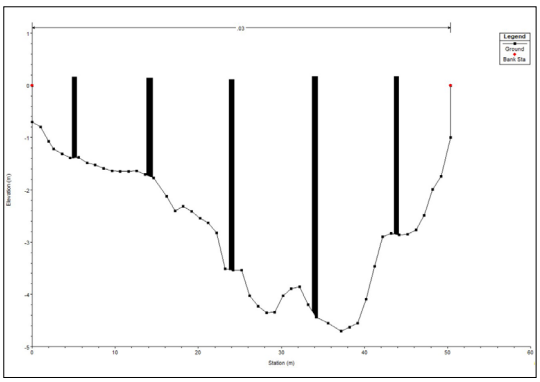


Fig. 5. *The current cross-section of the Olt River at the current bridge.*

the Santimbru commoners, the Santimbru local council" (translation by the authors). Additional information is engraved in the bridgehead: "Built in 1980". According to local residents, however, it was only renovated in 1980, the bridgehead was present much earlier. The original bridge (without piers) had a single span of approximately 25 meters.

The current bridge is 52 meters long and has 5 piers, being adjusted to the embankment built on this section of the Olt River after the widening of the riverbed (**Figure 4**).

2. Preliminary calculations

The subject of this investigation is the point in the Santimbru section of the Olt River where the bridge illustrated in **Figure 1** was located and where the current road bridge shown in **Figure 4** is located. The characteristics of this section of the river are shown in **Table 1** and **2**.

At a distance of approximately 4.2 km above the former bridge there is a hydrometric station, the known data of which greatly helped the accuracy of the calculations.

Using the above data, the values shown in **Table 3** resulted from the preliminary hydrological calculations, considering the maximum water flow of the 1% and 5% probability floods.

The current cross-section of the riverbed under the new bridge built after the riverbed widening works was modelled based on field measurements, as shown in **Figure 5.**

Table 1. *Long-term average flow.* **[2]**

River name	Cadastral number	Average flow
Olt	VIII.1	5.74 m^3/s

Table 2. *The parameters of the Olt River in the analysed cross-section* **[3]**

Parameter	Value
Length	62 km
Average slope	1.2 %
Sinuosity	1.27
Retention area	956 km^2
Average altitude above sea level	642 m

Table 3. *Ahe maximum flood flows of the Olt River in the cross-section below the bridge.*

Probability	Q_{max} p%
1 %	430 m^3/s
5 %	232 m^3/s

Under and around the current bridge, a large amount of sediment has been deposited in the riverbed. Based on the geometry of the undamaged bridge abutments and the bed level, the late cross-section assumed to be taken into account in the design of the bridge was modelled by ignoring the subsequent widening of the bed, as shown in **Figure 6.**

3. Design specifications

Based on the Romanian regulations, the selection of the flood yield used in the design of bridges is made according to the 4068/2 standard **[4]** depending on the importance classification of the construction. According to standard 4273 **[5]**, the bridge in Santimbru belongs to importance class 4, so the water flow value for design is the flood water yield with a 5% probability.

A 2010 government decision **[6]** stipulates that the target set on the basis of the national strategy for the future is that for rural settlements, the 1% probability flood water yield should be the design flow value, for increased safety.

During hydraulic design, when determining the water flow through the section of bridges, a certain free height must also be taken into account, which in our case means a height of 1.5 meters according to the Romanian PD-95 standard. **[7]**

4. Hydraulic calculations

Hydraulic modelling was carried out using the HEC-Ras program of the US Army Corps Engineers. The mentioned program uses Chézy-Manning hydraulic model.

With the help of a first model, it was examined to what extent the assumed cross-section in the case of the old bridge corresponds to the Romanian standards before 2010 and to the current standards.

Based on the results (**Figure 7**) we can state that the old bridge meets the requirements of Romanian design standards before 2010, since the model of the assumed original cross-section results in a 1.85-meter clear height at the maximum flow of a 5% probability flood. On the other hand, in the case of the maximum flow of a 1% probability flood, the free height is only 0.96 meters, so it does not meet the currently formulated national flood protection strategy requirement.

With the help of a second model, we examined the current cross-section of the riverbed in the case of the already mentioned maximum flood flows, under the existing road bridge.

As can be observed in **Figure 8** the current cross-section results in a free height of 11 cm in the case of a 1% probability maximum flow, and 90 cm in the case of a 5% probability maximum flow, so it does not provide the minimum 1.5-meter free height required by the PD-95 standard for flood discharges in case of the pre-2010 or post-2010 Romanian flood protection design standards.

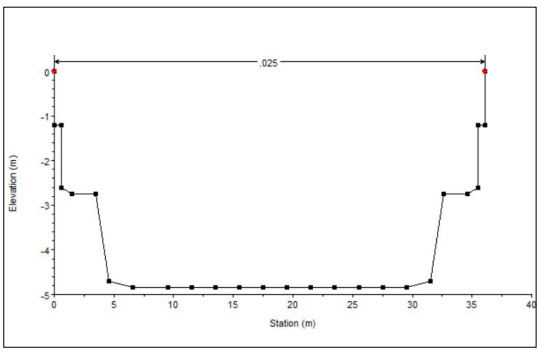


Fig. 6. *The supposed old cross-section of the riverbed under the bridge.*

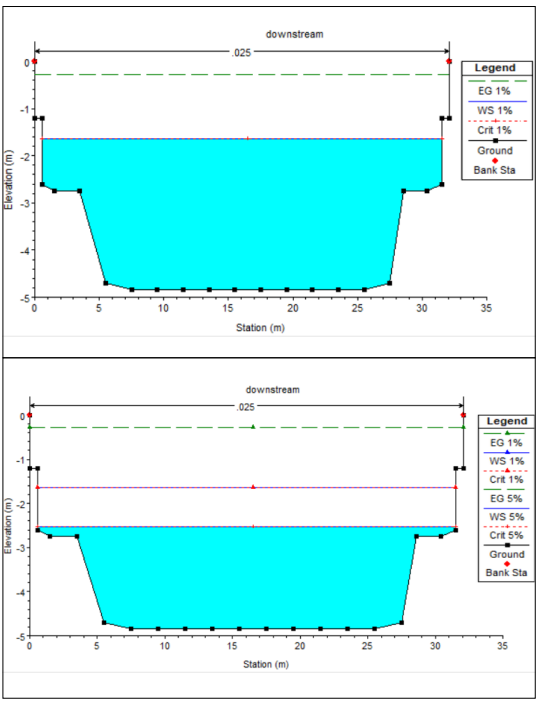


Fig. 7. *The assumed old cross-section of the riverbed with a 1% and 5% probability flood flow, under the old bridge.*

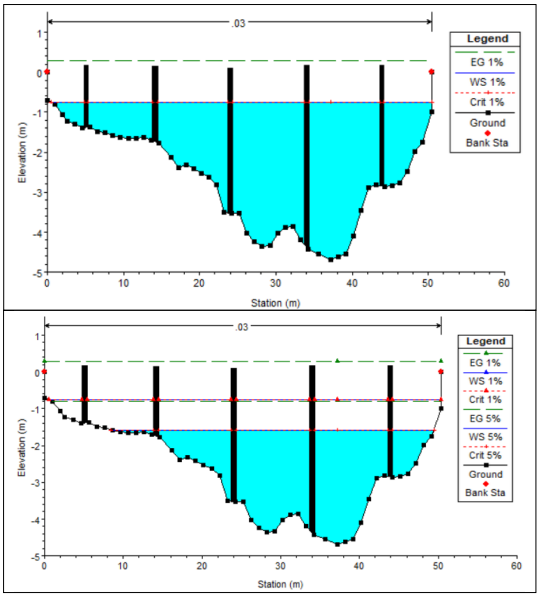


Fig. 8. *The current cross-section of the riverbed with a 1% and 5% probability flood flow, under the current road bridge,*

5. Conclusions and further plans

The hydraulic calculations carried out on the basis of measurements at the bridge in Santimbru indicate that the standards used in the design of the old bridge may have been stricter than the Romanian regulations valid until 2010, because the assumed cross-section of the riverbed at that time ensures the prescribed free flow height in case of the 5% probability flood water flow. This probability can be unequivocally substantiated only after studying the design regulations of the time, since the assumed cross-section was taken into account as an idealized model without alluvium and deposits.

Our second conclusion is that the current cross-section does not ensure the free height required by the standards, so it would be necessary to remove the deposited bed sediment (**Figure 9**) and stabilize the riverbed.

Further plans include the search and examination of the remains of Transylvanian bridges similar to the bridges in Sancraieni **[8]** and Santimbru in order to determine the original design requirements.

Fig. 9. *Deposited riverbed sediment.*

Acknowledgments

Thanks are due to the Department of Technical Sciences of the Transylvanian Museum Society for the support provided for the publication of this paper. We would also like to thank Zsuzsánna Kisfaludi-Bak for her help during the field measurements.

References

[1] Képeslapok. Erdélyi képeslapok a múltból, Pethő Csongor gyűjteménye. *Csikszentimre: Olt vasúti hídja*, 1912. https://kepeslapok.wordpress.com/2014/11/25/csik-2/csikszentimrehid/ (letöltés: 2022.10.11)

[2] Ministerul Mediului și Dezvoltării Durabile: Studii pentru cunoașterea resurselor de apă în vederea fundamentării planurilor de amenajare ale bazinelor / spațiilor hidrografice, *Bazinul hidrografic Olt*, 2008.

[3] Ministerul Mediului: Atlasul cadastrului apelor din România, Partea 1 – *Date morfo-hidrografice asupra rețelei hidrografice de suprafață*, 1992.

[4] STAS 4068/2-87: Debite și volume maxime de apă. Probabilități teoretice ale debitelor maxime în condiții normale și speciale de exploatare.

[5] STAS 4273-83: Construcții hidrotehnice. Încadrare în clase de importanță.

[6] HG 846-2010: Strategia naţională de management al riscului la inundaţii pe termen mediu şi lung.

[7] PD 95-2002: Normativ privind proiectarea hidraulică a podurilor și podețelor.

[8] Kisfaludi-Bak Zs., Gobesz F.-Zs.: *A csíkszentkirályi lebombázott Olt-híd tervezésének részleges vizsgálata.* Műszaki Tudományos Közlemények, 16. (2022) 43–46. https://doi.org/10.33895/mtk-2022.16.08

Műszaki Tudományos Közlemények vol. 18. (2023) 53–57.
DOI: Hungarian: https://doi.org/10.33895/mtk-2023.18.10
English: https://doi.org/10.33894/mtk-2023.18.10

INVESTIGATION OF THE EFFECT OF A SUSPENDED WORKING MACHINE ON TRACTION FORCE

Judit PÁSZTOR,[1] Előd MIKLÓS,[2] Rudolf-László FARMOS,[3] Izolda POPA-MÜLLER,[4] Erzsébet EGYED-FALUVÉGI[5]

[1] *Sapientia Hungarian University of Transylvania, Faculty of Technical and Human Sciences, Târgu Mureş, Department of Mechanical Engineering, Romania, pjudit@ms.sapientia.ro*

[2] *Knorr-Bremse Development and Research Center, Budapest, Hungary, Elod.Miklos@knorr-bremse.com*

[3] *Sapientia Hungarian University of Transylvania, Faculty of Technical and Human Sciences, Târgu Mureş, Department of Mechanical Engineering, Romania, farmos_rudolf@ms.sapientia.ro*

[4] *Sapientia Hungarian University of Transylvania, Faculty of Technical and Human Sciences, Târgu Mureş, Department of Mechanical Engineering, Romania, ipmuller@ms.sapientia.ro*

[5] *Sapientia Hungarian University of Transylvania, Faculty of Technical and Human Sciences, Târgu Mureş, Department of Mechanical Engineering, Romania, faluvegi.erzsebet@ms.sapientia.ro*

Abstract

It is always an aim in agricultural machine operation to understand the tractor-machine-soil relationship precisely. The working machine is often connected to the tractor with a three-point linkage system. The three-point linkage system coordinates the movement of the tractor-work machine group, increases the traction force, and also enables the adjustment of the coupled work machine. Understanding how it works is very important. In this paper, we examine the effect of the connection points of the three-point linkage system on the change in the adhesion weight force in a specific case, in the case of U532DT tractor and PP2-30 suspended plough.

Keywords*: three-point linkage system, adhesion force, traction force, suspended plough.*

1. Introduction

1.1. Traction force of an agricultural power machine

The tractor is the energy source of agriculture, its task is to tow and operate the working machine. A good use of the traction power of a tractor contributes to the reduction of agricultural environmental impact [1], [2]. The traction force is the adhesion force appearing on the driven wheels of the power machine. Its maximum value depends on the quality of the surfaces between the soil and the driven wheels and the wheel load:

$$F_{vmax} = \mu \cdot G_{adh} \text{ [N]}, \qquad (1)$$

where: μ adhesion coefficient; G_{adh} adhesion weight force [N].

The adhesive weight force is the weight force exerted on the driven wheels of the tractor. Generally, 2/3 of the tractor's weight is considered adhesive weight force [3], but in the case of four-wheel drive, the entire tractor's weight is adhesive weight force.

The maximum traction force can be increased by increasing the adhesion coefficient and the adhesion weight force. There are several options for increasing the adhesion weight force. In this paper, we examine the additional weighting effect of the work machine suspended on the power machine.

In the paper, we study and measure the load of the weight of the suspended work machine on the driven wheel. This increases the adhesive weight force, thus causing an increase in traction force when towing. The work machine is attached to the power machine with the three-point linkage system.

1.2. Methods of determining traction force

Traction force is measured directly with measuring equipment or determined indirectly [4].

1.2.1. Direct measuring methods

In this case, the forces occurring at different points of the tool are measured during work and the traction force is determined from that. This can happen:

– by measurement in laboratory conditions;
– by measuring in field conditions.

Laboratory measurements take place in a soil bin. The main parts of the soil bin: soil modelling soil tank, with rails on both sides; tillage tool holder that moves on rails; power source and propulsion system that models the tractor; measurement/data collection and analysis system. The soil bin is large, takes up a lot of space and has a high investment cost.

A force measuring frame is used for outdoor measurements. The force measuring frame is placed between the mounted implement and the tractor's three-point hitch. Force-measuring sensors are attached to the force-measuring frame, which measure the traction force/traction resistance in the horizontal plane and the lifting/loading forces in the vertical plane. The size of the force measuring frame is large and robust. In general, each research centre and factory prepare its own force measuring frame. Not commercially available.

1.2.2. Indirect determination method

In this case, measurements are made in experimental conditions, and from these values the actual traction force during work is deduced.

A working machine connected with a three-point linkage system causes an increase in traction force by increasing the adhesive weight force [4]. Thus, the forces measured on the suspended working machine can be suitable for indirectly determining the change in traction force. With dynamic modelling, the forces acting on the group of machines, which are essential in terms of the phenomenon, can be revealed. The forces can be analysed and the factors influencing the traction force can be selected based on this [5].

2. The investigation

2.1. Theoretical background of the investigation

Part of the weight of the suspended working machine is loaded on the power machine, thus influencing the G_{adh} adhesive weight force, and indirectly the traction force. We examine the effect of the additional weight of a suspended working machine.

The measurements were carried out on the U523DT universal horticultural tractor and PP2-30 suspended, two-plough-body plough machine group. This test can also be performed when the machine group is stationary.

With dynamic modelling, we reveal the forces acting on the group of machines, which are essential from the point of view of the phenomenon.

In the vertical plane the forces on the tractor-suspended plough machine group are in balance [6]:

$$R_E + R_H - G_T - G_E + R_K = 0 \text{ [N]} \quad (2),$$

where:

– R_E and R_H are soil resistance forces under the wheels of the power machine [N];
– G_T is the weight force of the tractor [N];
– G_E is the weight force of the plough [N];
– R_K is the soil resistance force under the depth limiting wheel [N].

Plowing requires a large tractive force, for this reason the 4×4-es drive is mostly used on the power machine, so the G_{adh} can be calculated with the relationship (2):

$$G_{adh} = R_E + R_H = G_T + G_E - R_K \text{ [N]}. \quad (3)$$

The relation (3) indicates that the R_K force affects the value of G_{adh}.

Therefore, the soil resistance force under the depth- limiting wheel of the suspended working machine is suitable for characterizing the change in traction force.

To examine and analyse the R_K force, we write down the moment balance of the forces acting on the plough at the vertical instantaneous centre of rotation P [5]:

$$-R_K \cdot (x + l_K + l) + G_E \cdot (x + l + l_G) = 0, \quad (4)$$

where: x is the distance of the vertical instantaneous center of rotation from the wheel center [m]; the distances l_K, l, l_G can be seen in Figure 1 [m].

The R_K soil resistance force can be calculated with the relation (4):

$$R_K = \frac{G_E \cdot (x + l + l_G)}{(x + l_K + l)} \text{ [N]}. \quad (5)$$

The relationship (5) indicates that the magnitude of the R_K is influenced by the distance x. The magnitude of the distance x is influenced by the connection point of the three-point linkage system (Figure 1).

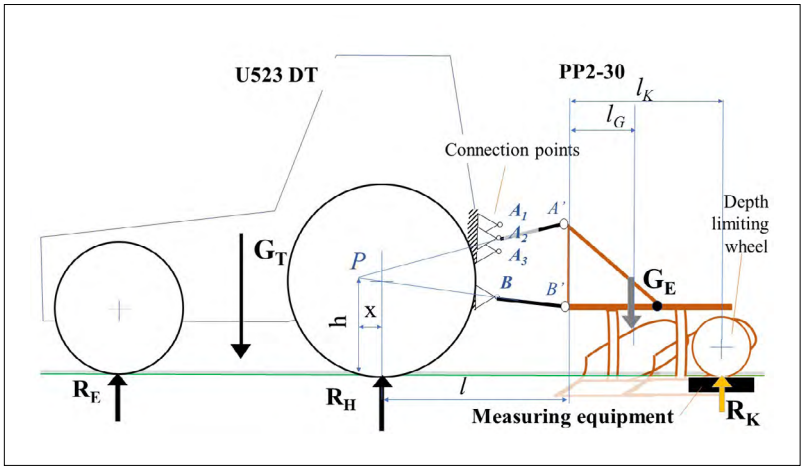


Fig. 1. *Dynamic model of a tractor-suspended working machine.*

2.2. The method of investigation

It is assumed that the connection points of the upper support rod of the three-point linkage system have an effect on the adhesion weight force G_{adh}.

We measure the soil resistance under the depth-limiting wheel in three cases, at the three connection points of the upper support rod: in points: A_1, A_2, A_3 (**Figure 1**).

2.3. The measuring equipment

The measuring equipment is actually a scale that is placed under the depth limiter wheel. Its structure can be seen in **Figure 2**:

- mechanical part: holder cassette, cover;
- electronic part: control-data collector, indicator, printed wiring board, buttons;
- force measuring sensor.

2.4. The measurements

The measurements were carried out in June 2022 in The Faculty of Technical and Human Scientist Târgu-Mureș Medicinal Plant Garden of Sapientia Hungarian University of Transylvania.

We assembled the tractor-plough machine group and we adjusted the plough: the longitudinal horizontality with the screw spindle of the upper support arm; the transverse horizontality with the lift links; the working depth with the depth-limiting wheel to 21 cm.

The state of the tractor-plow machine group during operation can be seen in **Figure 3**. The tractor's right wheel is at the bottom of the previously plowed furrow; thus, the tractor's axis is not horizontal, we paid attention to this when creating the measurement conditions.

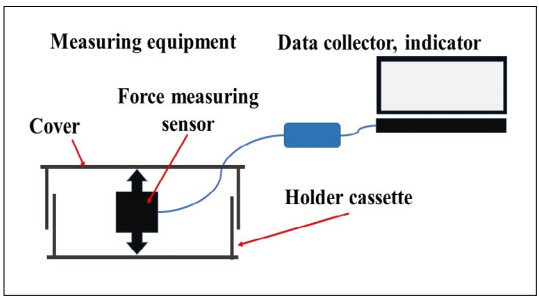


Fig. 2. *Construction principle of the measuring equipment.*

Fig. 3. *The measuring location.*

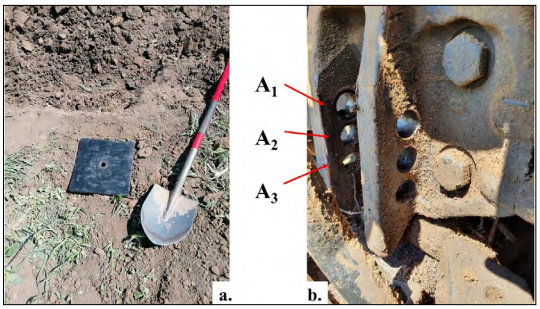


Fig. 4. *The measurement conditions: a) Buried measuring equipment; b) Connection points.*

Table 1. *Averages of measurement data*

	Measurement I.			Measurement II.		
Connection points	A_1	A_2	A_3	A_1	A_2	A_3
R_K average [N]	1002.93	978.27	1028.23	904.87	849.33	952.53
***s*, deviation**	0.12	0.48	9.11	2.79	1.38	3.69

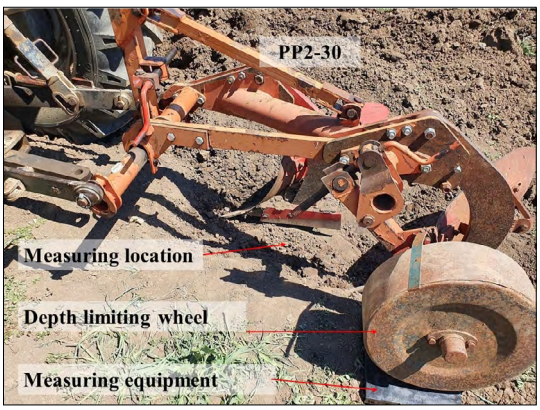


Fig. 5. *Measurement at the measuring location*

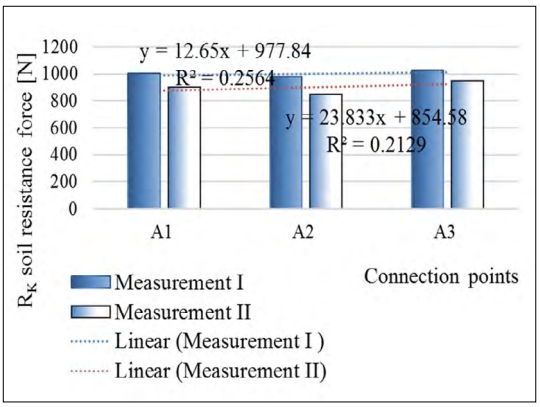


Fig. 6. *Soil resistance force under the depth-limiting wheel of the suspended PP2-30.*

We dug the measuring location of the plough bodies at 21 cm depth.

We determined the location of the depth-limiting wheel with trial pulls, and there we dug in the measuring equipment up to the soil surface (**Figure 4.a.**).

We calibrated the measuring device with a known mass.

We fixed the upper support arm of the three-point linkage system at point A1 (**Figure 4.b.**) and we performed the measurement, then we proceeded similarly for the connection point A3.

We repeated the series of measurements (**Figure 5**).

2.5. Processing of measurement data

The averages of measurement data are described in **Table 1**.

The data were displayed (**Figure 6**).

A moderate correlation between the connection points of upper rod and the soil resistance under the depth-limiting wheel can be shown, $R_I = 0.5$, $R_{II} = 0.48$ **[7]**.

The soil resistance under the depth-limiting wheel is the highest in the lower connection point A_3, for both measurements. The high soil resistance means that the plow rests on the soil, so the weight on the driven axle of the power machine is less, i.e. the additional weighting effect of the suspended PP2-30 machine is smaller.

We verified with a hypothesis test that the connection points have an effect on the ground resistance under the depth-limiting wheel at the $p = 0.1$ level. **[7]**

The values of G_{adh} calculated with relation (3) in the versions with and without suspended work machine are summarized in **Table 2** and shown in **Figure 7.**

The machine data used in the context: tractor weight force $G_T = 27566.1$ N, plough weight force $G_{eke} = 2452.5$ N.

The suspended working machine increases the value of G_{adh} so it has effect of increasing traction force. The additional weighting effect was checked with a hypothesis test, the difference is very significant at the $p = 0.1$ level **[7]**.

Table 2. *Traction increasing effect of suspended PP2-30 work machine*

Connection points	A_1	A_2	A_3
R_K average [N]	953.90	913.80	990.38
G_{adh} with suspended work machine	29064.70	29104.80	29028.22
G_{adh} without suspended work machine	27566.10	27566.10	27566.10
Traction force increase %	**5.44**	**5.58**	**5.30**

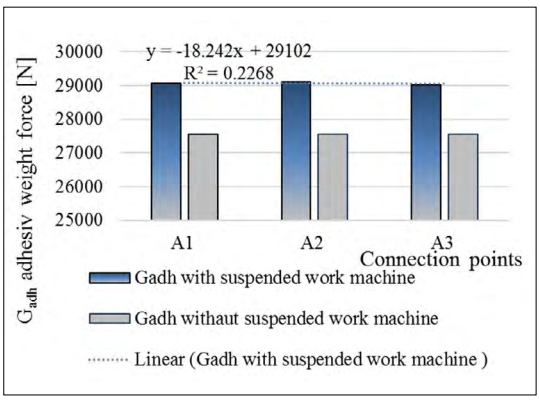


Fig. 7. *Changes in adhesion weight force due the suspended work machine and the connection points/*

3. Conclusions

We verified by measurement that the suspended working machine increases the adhesion weight force, thus indirectly affecting the traction force.

We proved that the connection points of the three-point linkage system influence the magnitude of the adhesion weight force. The lower attachment point causes the smallest additional weighting effect, that is here the effect on the traction force is the smallest.

The trend of the measurement data confirms the highest additional weighting effect can be expected at the upper connection point. The results of the measurement were influenced by the conditions of the measurement: the preparation of the measuring location, the placement of the plow on the measuring equipment, the effect of the connection points on the horizontality of the machine. This explains the measurement data large deviation and the discrepancy from the expected values in the middle connection point.

Based on our measurements, the PP2-30 mounted implement results in an average increase in traction force of 5.44 %.

References

[1] Matache M. G., Cristea M., Găgeanu I., Zapciu A., Tudor E., Carpus E., Popa L. D.: *Small Power Electric Tractor Performance During Ploughing Works*. In: INMATEH – Agricultural Engineering, 6/1. (2020) 123–128. https://doi.org/10.35633/inmateh-60-14

[2] Cârdei P., Nuţescu C., Matache M., Cristea O.: *Optimum Working Conditions for Variable Width Ploughs*. In: INMATEH – Agricultural Engineering, Vol. 65 (3). (2021). 248–254. https://doi.org/10.35633/inmateh-65-26.

[3] Vasilache L., Nastasoiu M., Nastasoiu S., Ispas N., Simeon C.: *Researches on the Mathematical Modelling of the Kinematics of Three-Point Hitch Couplers Used at Agricultural Tractors*. In: CONAT 2010 International Automotive Congress, Braşov, Romania, CONAT 5011. (2010). 47–54. https://hdl.handle.net/123456789/149

[4] Kumar G. V. P.: *Geometric Performance Parameters of Three-Point Hitch Linkage System of a 2WD Indian Tractor*. Research in Agricultural Engineering, 61/1. (2015) 47–53. https://doi.org/10.17221/79/2012-RAE

[5] Pásztor J., Popa-Müller I.: *Study of Three-Point Linkage of Power Machine*. In: Műszaki Tudományos Közlemények, Vol. 14. (2021). 60–64. https://doi.org/10.33895/mtk-2021.14.09

[6] Máté M.: *Műszaki mechanika – kinematika*. Erdélyi Múzeum-Egyesület, Kolozsvár, 2010. https://doi.org/10.36242/mtf-10

[7] Harangus K., Kakucs A.: *Valószínűségszámítás és statisztika a mérnöki gyakorlatban*. Erdélyi Múzeum-Egyesület, Kolozsvár, 2021. https://doi.org/10.36242/mtf-15.

Műszaki Tudományos Közlemények vol. 18. (2023) 58–62.
DOI: Hungarian: https://doi.org/10.33895/mtk-2023.18.11
English: https://doi.org/10.33894/mtk-2023.18.11

KINEMATICS ANALYSIS OF CRANK MECHANISM WITH CONSTRAINT EQUATIONS

Izolda POPA-MÜLLER,[1] István PAPP[2]

[1] *Sapientia Hungarian University of Transylvania, Faculty of Technical and Human Sciences,, Department of Mechanical Engineering. Târgu Mureș, Romania, ipmuller@ms.sapientia.ro*

[2] *Transylvanian Museum-Society, Department of Technical Sciences, Cluj-Napoca, Romania, pappistvan1944@yahoo.com*

Abstract

The equation system, described for kinematic pairs of Crank mechanism, consist of general constraint equations. For Crank mechanism, the position of the center of gravity for the components, and the directions of the main axis of inertia can be determined for the entire kinematic cycle. To determine the positions, we use a coordinate system fixed to the center of gravity of the mechanism. The coordinates of the center of gravity of the members of the mechanism are always constant in relation to the geometric system attached to the component. The vector contour created in this way, defined by position vectors, must always be closed.

Keywords*: constraints, kinematic pairs, constraint equations.*

1. Crank Mechanism

The crank mechanism consists of (**Figure 1**) :
– engine fuselage (1);
– crankshaft (2);
– crank arm (3);
– piston (4).

OXYZ is the fixed system, $G_iX_iY_i,Z_i, i\in\overline{1,4}$ are the crankshaft, crank arm, piston central axis of inertia system, $O_iX_i^*Y_i^*Z_i^*, i\in\overline{1,4}$, is the auxiliary systems of the crankshaft components.

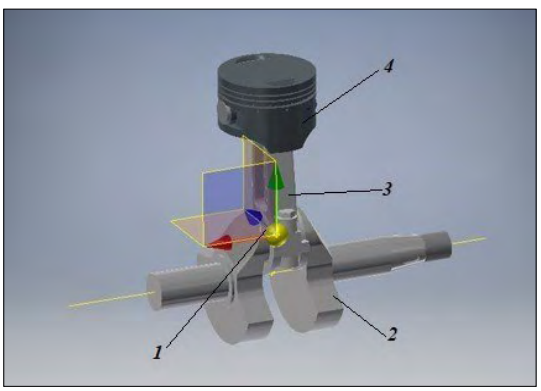


Fig. 1. *3D model of the crank mechanism and visibility of its center of gravity*

The vector contour defined by position vectors, must always be closed (**Figure 2**).

$$\vec{r}_{12}-\vec{r}_{21}+\vec{r}_{23}-\vec{r}_{32}+\vec{r}_{34}-\vec{r}_{43}+\vec{r}_{41}-\vec{r}_{14}=0 \quad (1)$$

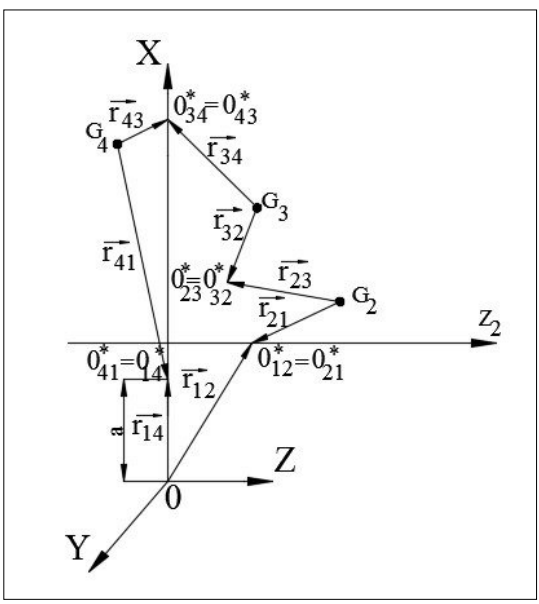


Fig. 2. *Vector contour of crank mechanism.*

1.1. Determination of mechanism constants

The center of gravity of the mechanism, in the auxiliary system:

$$G_1 = \begin{pmatrix} 50 \\ 0 \\ 0 \end{pmatrix} \cdot 10^{-3} \tag{2}$$

Directional factors of the main axis of inertia of the mechanism, in relation to the auxiliary system:

$$A_1 = \begin{pmatrix} 0.999277 & 0 & 0 \\ 0 & 0.999097 & 0 \\ 0 & 0 & 0.99978 \end{pmatrix} \tag{3}$$

The direction factors of the auxiliary system, in relation to the coordinate system of the main axis of inertia of the mechanism [1]:

$$\begin{pmatrix} \cos\alpha_{1i}^0 & \cos\beta_{1i}^0 & \cos\gamma_{1i}^0 \\ \cos\alpha_{2i}^0 & \cos\beta_{2i}^0 & \cos\gamma_{2i}^0 \\ \cos\alpha_{3i}^0 & \cos\beta_{3i}^0 & \cos\gamma_{3i}^0 \end{pmatrix} = A_1^T \tag{4}$$

The $\vec{r}_{12}$ position vector defines the coordinates of the point O_{12}^* in relation to the auxiliary system placed in the center of gravity of the mechanism:

$$\vec{r}_{12} = O_{12}^* - G_1 \tag{5}$$

The $\vec{r}_{12I}$ position vector expresses the coordinates of the point O_{12}^* in relation to the coordinate system of the central axes of inertia placed in the center of gravity of the mechanism:

$$\vec{r}_{12I} = A_1^T \cdot \vec{r}_{12} \tag{6}$$

The $\vec{r}_{14}$ position vector expresses the coordinates of the point O_{14}^* in the auxiliary system placed in the center of gravity of the mechanism:

$$\vec{r}_{14} = O_{14}^* - G_1 \tag{7}$$

The fixed point O_{14}^* on the mechanism that coincides with the origin of the auxiliary system fixed on the mechanism and is located on the axis of the 4th piston:

$$O_{14}^* = \begin{pmatrix} 30 \\ 0 \\ 0 \end{pmatrix} \tag{8}$$

1.2. Determination of the crankshaft constants

From the Inventor window (Figure 4) you can read the moments of inertia of the crankshaft and the coordinates of its centre of gravity.

The O_{21}^* point on the main axis of the mechanism (Figure 3) the position vector relative to the auxiliary system placed at the center of gravity G_2 can be written:

$$\vec{r}_{21} = O_{21}^* - G_2, \; O_{21}^* = \begin{pmatrix} 0.03 \\ 0 \\ -0.0125 \end{pmatrix} \tag{9}$$

The center of gravity of the axle, in relation to the auxiliary system on the main axle (Figure 4):

$$G_2 = \begin{pmatrix} 1.677 \\ -0.004 \\ -5.574 \end{pmatrix} \cdot 10^{-3} \tag{10}$$

Directional factors of the main axes of inertia in relation to the auxiliary system on main axis 2:

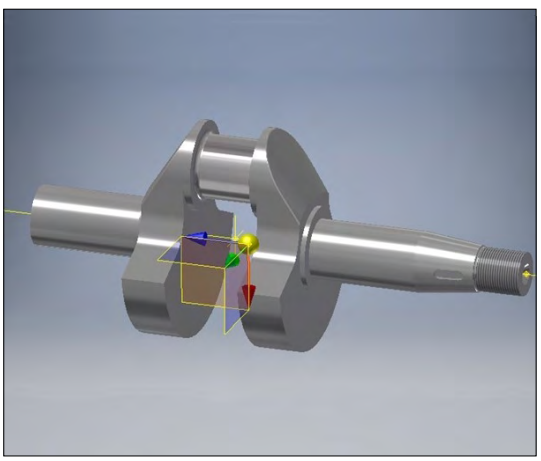


Fig. 3. *3D model of the crankshaft and visibility of its center of gravity.*

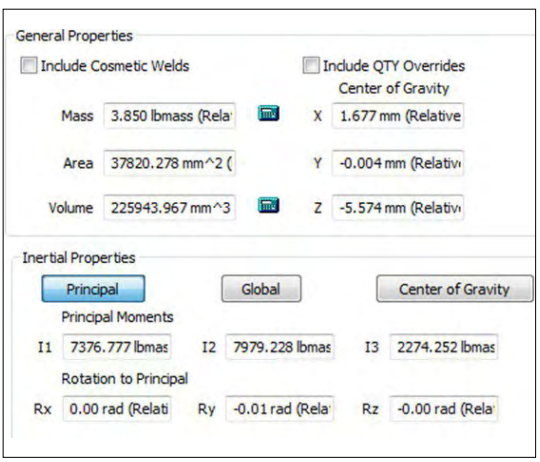


Fig. 4. *Data of inertia and center of gravity of crankshaft in Inventor.*

$$A_2 = \begin{pmatrix} 0.99997563 & 0 & 0 \\ 0 & -1.744444 \cdot 10^{-4} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (11)$$

1.3. Determination of the crank arm constants

The coordinates of the center of gravity relative to the auxiliary system taken on the axis line of the upper hole of the crank arm connected to the crankshaft (**Figure 5**):

$$G_3 = \begin{pmatrix} -61.158 \\ -0.004 \\ -2.586 \end{pmatrix} \cdot 10^{-3} \quad (12)$$

The coordinates of the point O_{34}^* taken on the axis line of the lower hole of the crank arm relative to the auxiliary system of the crank arm. The point coincides with the point O_{43}^* on the piston pin.

$$O_{34}^* = \begin{pmatrix} 89.3 \\ 0 \\ 0 \end{pmatrix} \cdot 10^{-3} \quad (13)$$

1.4. Determination of the piston constants

Direction factors of the main inertial axes relative to the auxiliary system attached to the piston (**Figure 6**):

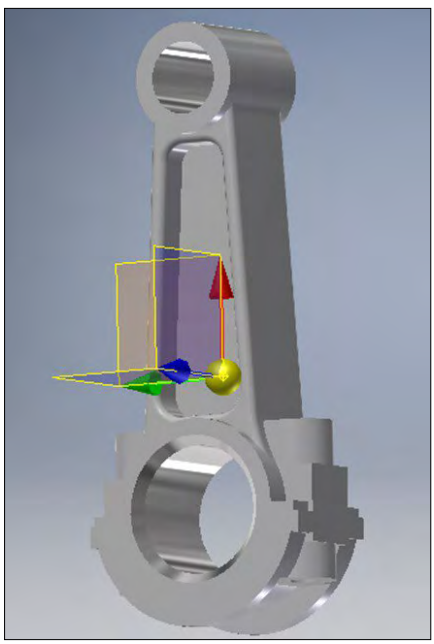


Fig. 5. *3D model of the crank arm and visibility of its center of gravity.*

$$A_4 = \begin{pmatrix} 0.99979172 & 0 & 0 \\ 0 & 0.99999998 & 0 \\ 0 & 0 & 0.99999945 \end{pmatrix} \quad (14)$$

The coordinates of the point O_{43}^* taken on the axis line of the piston pin coincide with the point taken on the crank arm during operation.

$$O_{43}^* = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \cdot 10^{-3} \quad (15)$$

2. Incorporating the constraint equations into a system

We define the components in 37 positions. The driving parameter of the mechanism ψ_2^*, knowing its direction factors formed with the inertial crankshafts for the components' auxiliary systems and its rotation angles. **[1, 2].**

2.1. Constraint equations of the crank mechanism discussed as a space mechanism

The number of constraint equations depends on the type of joint, in the case of a rotational joint we have 5 constraint equations.

The constraint equation written on the rotating joint of the mechanism **[3–5]**:

$$\begin{aligned} & x_{G_1} + x_{12} \cdot \alpha_{110} + y_{12} \cdot \alpha_{210} + z_{12} \cdot \alpha_{310} \\ & -x_{G_2} - x_{21} \cdot \alpha_{12}(\psi_2, \theta_2, \phi_2) + \\ & -y_{21} \cdot \beta_{12}(\psi_2, \theta_2, \phi_2) - z_{21} \cdot \gamma_{12}(\psi_2, \theta_2)) = 0 \end{aligned} \quad (16)$$

$$\begin{aligned} & y_{G_1} + x_{12} \cdot \beta_{110} + y_{12} \cdot \beta_{210} + z_{12} \cdot \beta_{310} \\ & -y_{G_2} - x_{21} \cdot \alpha_{22}(\psi_2, \theta_2, \phi_2) - y_{21} \cdot \beta_{22}(\psi_2, \theta_2, \phi_2) + \\ & -z_{21} \cdot \gamma_{22}(\psi_2, \theta_2) = 0 \end{aligned} \quad (17)$$

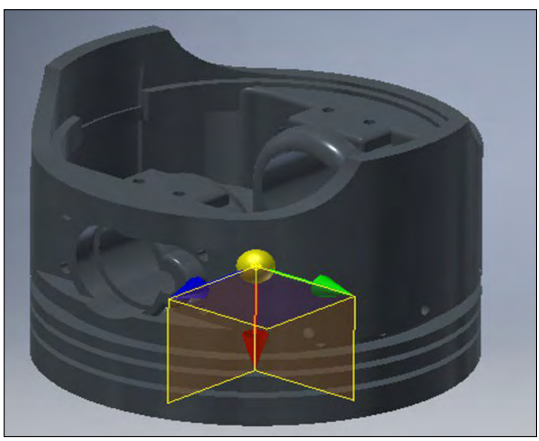


Fig. 6. *3D model of the piston and visibility of its center of gravity.*

$$z_{G_1} + x_{12} \cdot \gamma_{110} + y_{12} \cdot \gamma_{210} + z_{12} \cdot \gamma_{310} - z_{G_2} - x_{21} \cdot \alpha_{32}(\psi_2, \phi_2) - y_{21} \cdot \beta_{32}(\psi_2, \phi_2) - z_{21} \cdot \gamma_{32}(\psi_2, \theta_2) = 0 \quad (18)$$

$$b_{131} - \begin{pmatrix} \alpha_{12}(\psi_2, \theta_2, \phi_2) \cdot \gamma_{120} + \beta_{12}(\psi_2, \theta_2, \phi_2) \cdot \gamma_{220} + \\ \gamma_{12}(\psi_2, \theta_2) \cdot \gamma_{320} \end{pmatrix} = 0 \quad (19)$$

$$b_{231} - \begin{pmatrix} \alpha_{22}(\psi_2, \theta_2, \phi_2) \cdot \gamma_{120} + \beta_{22}(\psi_2, \theta_2, \phi_2) \cdot \gamma_{220} + \\ \gamma_{22}(\psi_2, \theta_2) \cdot \gamma_{320} \end{pmatrix} = 0 \quad (20)$$

Constraint equation written on the rotating joint of the crankshaft:

$$x_{G_2} + x_{23} \cdot \alpha_{12}(\psi_2, \theta_2, \phi_2) + y_{23} \cdot \beta_{12}(\psi_2, \theta_2, \phi_2) + z_{23} \cdot \gamma_{12}(\psi_2, \theta_2) - x_{G_3} - (x_{32} \cdot \alpha_{13}(\psi_3, \theta_3, \phi_3) + y_{32} \cdot \beta_{13}(\psi_3, \theta_3, \phi_3) + z_{32} \cdot \gamma_{13}(\psi_3, \theta_3)) = 0 \quad (21)$$

$$y_{G_2} + x_{23} \cdot \alpha_{22}(\psi_2, \theta_2, \phi_2) + y_{23} \cdot \beta_{22}(\psi_2, \theta_2, \phi_2) + z_{23} \cdot \gamma_{22}(\psi_2, \theta_2) - y_{G_3} - (x_{32} \cdot \alpha_{23}(\psi_3, \theta_3, \phi_3) + y_{32} \cdot \beta_{32}(\psi_3, \theta_3, \phi_3) + z_{32} \cdot \gamma_{32}(\psi_3, \theta_3)) = 0 \quad (22)$$

$$z_{G_2} + x_{23} \cdot \alpha_{32}(\theta_2, \phi_2) + y_{23} \cdot \beta_{22}(\theta_2, \phi_2) + z_{23} \cdot \gamma_{32}(\theta_2) - z_{G_3} - (x_{32} \cdot \alpha_{33}(\theta_3, \phi_3) + y_{32} \cdot \beta_{33}(\theta_3, \phi_3) + z_{32} \cdot \gamma_{33}(\theta_3)) = 0 \quad (23)$$

$$(\alpha_{12}(\psi_2, \theta_2, \phi_2) \cdot \gamma_{120} + \beta_{12}(\psi_2, \theta_2, \phi_2) \cdot \gamma_{220} + \gamma_{12}(\psi_2, \theta_2) \cdot \gamma_{320}) - (\alpha_{13}(\psi_3, \theta_3, \phi_3) \cdot \gamma_{130} + \beta_{13}(\psi_3, \theta_3, \phi_3) \cdot \gamma_{230} + \gamma_{13}(\psi_3, \theta_3) \cdot \gamma_{330}) = 0 \quad (24)$$

$$(\alpha_{22}(\psi_2, \theta_2, \phi_2) \cdot \gamma_{120} + \beta_{22}(\psi_2, \theta_2, \phi_2) \cdot \gamma_{220} + \gamma_{22}(\psi_2, \theta_2) \cdot \gamma_{320}) - (\alpha_{23}(\psi_3, \theta_3, \phi_3) \cdot \gamma_{130} + \beta_{23}(\psi_3, \theta_3, \phi_3) \cdot \gamma_{230} + \gamma_{23}(\psi_3, \theta_3) \cdot \gamma_{330}) = 0 \quad (25)$$

Constraint equation written on the rotating joint of the crank arm:

$$x_{G_3} + x_{34} \cdot \alpha_{13}(\psi_3, \theta_3, \phi_3) + y_{34} \cdot \beta_{13}(\psi_3, \theta_3, \phi_3) + z_{34} \cdot \gamma_{13}(\psi_3, \theta_3) - (x_{G_4} + x_{43} \cdot \alpha_{14}(\psi_4, \theta_4, \phi_4) + y_{43} \cdot \beta_{14}(\psi_4, \theta_4, \phi_4) + z_{43} \cdot \gamma_{14}(\psi_4, \theta_4)) = 0 \quad (26)$$

$$y_{G_3} + x_{34} \cdot \alpha_{23}(\psi_3, \theta_3, \phi_3) + y_{34} \cdot \beta_{23}(\psi_3, \theta_3, \phi_3) + z_{34} \cdot \gamma_{23}(\psi_3, \theta_3) - (y_{G_4} + x_{43} \cdot \alpha_{24}(\psi_4, \theta_4, \phi_4) + (y_{G_4} + x_{43} \cdot \alpha_{24}(\psi_4, \theta_4, \phi_4) + y_{43} \cdot \beta_{24}(\psi_4, \theta_4, \phi_4) + z_{43} \cdot \gamma_{24}(\psi_4, \theta_4)) = 0 \quad (27)$$

$$z_{G_3} + x_{34} \cdot \alpha_{33}(\theta_3, \phi_3) + y_{34} \cdot \beta_{33}(\theta_3, \phi_3) + z_{34} \cdot \gamma_{33}(\theta_3) - (z_{G_4} + x_{43} \cdot \alpha_{34}(\theta_4, \phi_4) + y_{43} \cdot \beta_{34}(\theta_4, \phi_4) + z_{43} \cdot \gamma_{34}(\theta_4)) = 0 \quad (28)$$

$$(\alpha_{13}(\psi_3, \theta_3, \phi_3) \cdot \gamma_{130} + \beta_{13}(\psi_3, \theta_3, \phi_3) \cdot \gamma_{230} + \gamma_{13}(\psi_3, \theta_3) \cdot \gamma_{330}) - (\alpha_{14}(\psi_4, \theta_4, \phi_4) \cdot \gamma_{140} + \beta_{14}(\psi_4, \theta_4, \phi_4) \cdot \gamma_{240} + \gamma_{14}(\psi_4, \theta_4) \cdot \gamma_{340}) = 0 \quad (29)$$

$$(\alpha_{23}(\psi_3, \theta_3, \phi_3) \cdot \gamma_{130} + \beta_{23}(\psi_3, \theta_3, \phi_3) \cdot \gamma_{230} + \gamma_{24}(\psi_3, \theta_3) \cdot \gamma_{330}) - (\alpha_{24}(\psi_4, \theta_4, \phi_4) \cdot \gamma_{140} + \beta_{24}(\psi_4, \theta_4, \phi_4) \cdot \gamma_{240} + \gamma_{24}(\psi_4, \theta_4) \cdot \gamma_{340}) = 0 \quad (30)$$

Constraint equation written on the rotating joint of the piston:

$$x_{G_4} + x_{41} \cdot \alpha_{14}(\psi_4, \theta_4, \phi_4) + y_{41} \cdot \beta_{14}(\psi_4, \theta_4, \phi_4) + z_{41} \cdot \gamma_{14}(\psi_4, \theta_4) + a \cdot b_{114}(\psi_4, \theta_4, \phi_4) - (x_{G_1} + x_{14} \cdot \alpha_{110} + y_{14} \cdot \alpha_{210} + z_{14} \cdot \alpha_{310} = 0 \quad (31)$$

$$y_{G_4} + x_{41} \cdot \alpha_{24}(\psi_4, \theta_4, \phi_4) + y_{41} \cdot \beta_{24}(\psi_4, \theta_4, \phi_4) + z_{41} \cdot \gamma_{24}(\psi_4, \theta_4) - (y_{G_1} + x_{14} \cdot \beta_{110} + y_{14} \cdot \beta_{210} + z_{14} \cdot \beta_{310}) = 0 \quad (32)$$

$$z_{G_4} + x_{41} \cdot \alpha_{34}(\theta_4, \phi_4) + y_{41} \cdot \beta_{34}(\theta_4, \phi_4) + z_{41} \cdot \gamma_{34}(\theta_4) - (z_{G_1} + x_{14} \cdot \gamma_{110} + y_{14} \cdot \gamma_{210} + z_{14} \cdot \gamma_{310}) = 0 \quad (33)$$

$$\alpha_{14}(\psi_4, \theta_4, \phi_4) \cdot \alpha_{140} + \beta_{14}(\psi_4, \theta_4, \phi_4) \cdot \alpha_{240} + \gamma_{14}(\psi_4, \theta_4) \cdot \alpha_{340} - b_{111} = 0 \quad (34)$$

$$\alpha_{24}(\psi_4, \theta_4, \phi_4) \cdot \alpha_{140} + \beta_{24}(\psi_4, \theta_4, \phi_4) \cdot \alpha_{240} + \gamma_{24}(\psi_4, \theta_4) \cdot \alpha_{340} - b_{211} = 0 \quad (35)$$

3. Conclusions

The following can be determined most effectively with the method of constraint equations:

-the position of the centre of gravity for each component of the crank mechanism;

-the positions of inertial crankshafts.

The values were determined relative to a fixed system.

This advantageous method can be used to determine the position of the main axis of inertia of the components.

References

[1] Papp I.: *Mechanizmusok elmélete.* Scientia Kiadó, Kolozsvár, 2010. 226 oldal.

[2] Popa-Müller I., Papp I.: *Négyoldalú térmechanizmus kinematikai vizsgálata kényszeregyenletek segítségével.* XXII OGÉT Nemzetközi Gépész Találkozó, Nagyszeben, Románia, 2014, ISSN 2068-1267, 301–305.

[3] Popa-Müller I., Papp I.: *Ötoldalú térmechanizmus kinematikai vizsgálata.* Műszaki Tudományos Közlemények 6. (2017) 147–150. https://doi.org/10.33895/mtk-2017.06.16

[4] Popa-Müller I., Papp I.: *Ötoldalú térmechanizmus, sebesség meghatározás.* XXV. OGÉT Nemzetközi Gépész Találkozó, Kolozsvár, Románia, 2017, ISSN 2068-1267, 340–344.

[5] Popa-Müller I., Papp I.: *Ötoldalú térmechanizmus, gyorsulás meghatározás*. Műszaki Tudományos Közlemények, 9. (2018) 199–202. https://doi.org/10.33895/mtk-2018.09.45

Műszaki Tudományos Közlemények vol. 18. (2023) 63–69.
DOI: Hungarian: https://doi.org/10.33895/mtk-2023.18.12
English: https://doi.org/10.33894/mtk-2023.18.12

ANOMALY DETECTION WITH ARTIFICIAL INTELLIGENCE METHODS – AN OVERVIEW

Sebestyén Pál György,[1] Hangan Lia-Anca,[2] Czakó Zoltán[3]

[1] *Technical University of Cluj-Napoca, Faculty of Automation and Computers, Department of Computers. Cluj-Napoca, Romania, gheorghe.sebestyen@cs.utcluj.ro*

[2] *Technical University of Cluj-Napoca, anca.hangan@cs.utcluj.ro*

[3] *Technical University of Cluj-Napoca, zoltan.czako@cs.utcluj.ro*

Abstract

Nowadays, more and more human activities depend on computer-based automated systems. Fully automated (robotized) production lines, energy distribution infrastructures and other urban services or environmental surveillance systems are just some examples of cyber-physical systems that depend entirely on automated control systems. In these cases a significant challenge is to identify abnormal behaviors of the supervised or controlled systems, in order to avoid malfunction or sometimes catastrophic events. Our main research goal was to evaluate the potential of adapting and using AI techniques in the field of anomaly detection. We also developed a platform, called AutomaticAI, which can help specialists in different domains to identify the best approaches to solve a given anomaly detection problem. The platform can select the best AI algorithm and parameter configuration for a given set of data containing normal and abnormal data. The tool was used successfully in a variety of domains, from cyber-physical systems to the medical domain.

Keywords*: anomaly detection, artificial intelligence, outlier detection.*

1. Introduction

1.1. Context and motivation

In the times of the forth industrial revolution (Industry 4.0) we are surrounded by different kinds of cyber-physical systems, such as robotized industrial processes, energy and other urban service distribution infrastructures, autonomous transportation, remote healthcare systems, intelligent buildings and cities. All these systems are supervised and controlled by computers, computer programs or computer-based devices. Sometimes services and functionalities critical for the safety of our lives (e.g. autonomous cars) are totally dependent on such autonomous systems. But what happens when something goes wrong, for instance a sensor is not providing the correct measured value, a control device responsible for a critical functionality is broken, or the behaviour of the controlled system is changing in a manner not foreseen in the designing phase? How does an automated (e.g. computer-based) supervising system detect and then react to such an anomaly?

In the past when supervision tasks were performed by humans (e.g. process operators, or simple beneficiaries of some services or devices) a basic training was enough to detect and react to such cases. In an automated system the anomaly detection should be part of the control system. Therefore, in recent years anomaly detection has become an important research subject, and a number of solutions have been proposed and tested.

Anomalies may come in different shapes and forms, from very simple ones (e.g. a value that exceeds an allowed interval) that can be solved with simple thresholding or filtering methods towards complex ones that are hard to define and detect even by a human observer (e.g. climate change effects). In this last case artificial intelligence methods can be used to model and recognize abnormal behaviour.

1.2. Anomaly types

A simple definition for “anomaly” would be a data or a (system) behavior that is very different

from other: data or detected behaviour. A more elaborate definition was given by Hawkins in 1980: "anomaly is an observation that differs so much from other observations that raises suspicion that it was produced by a different mechanism" (e.g. a malfunction in the system or an artificial intervention in a financial process).

There are some other words that have similar or close meaning with the anomaly word: abnormal (behaviour), outlier or deviant.

There are a number of issues regarding the recognition of an anomaly:

- how big the deviation should be in order to classify it as an anomaly?
- natural noise present in some data should not be confused with anomaly
- an anomaly may be a question of viewpoint, or might depend on the context in which it was produced (e.g. a European in an Asian population may be considered an anomaly)
- humans, based on education and experience have a natural ability to recognize an anomaly, even there is no rational explanation for it; computers do not have this "natural" ability.

Regarding the types of possible anomalies we could mention some of them:

- a data/value in the evolution of a signal that exceeds some interval considered normal;
- a change in the linear evolution of a signal;
- some stochastic values that don't follow a known probability distribution (e.g. Gaussian distribution);
- some multi-parameter observations/objects that don't fit in some pre-defined classes considered normal cases;
- some deviations in the periodicity of a signal (e.g. arithmias in an ECG signal).

1.3. Domains for automated anomaly detection

Automated recognition of anomalies is an important task [1–4] in the context of new trends such as: IoT, IioT, and Industry 4.0. It is important for safety critical cyber-physical systems where the system must contain fault tolerant and self healing solutions.

In healthcare, different diseases and their symptoms may be considered anomalies of the human body. Early detection of such changes in the physiological parameters of a patient may prevent later evolution of more serious diseases.

Anomaly detection is also useful in any application which deals with huge amounts of data, catalogued as „big data". In this case manual analyses are not feasible, but some outlier data may influence the outcome of the automated analysis. Therefore, anomaly detection is used in order to eliminate erroneous data.

Automated anomaly detection is needed also in cases when changes are so subtle that the human senses don't detect them. For instance, in industry early detection of future defects is the bases for preventive maintenance.

Examples of domains where anomaly detection may play a significant role:

- in economic and financial applications, for the detection of frauds or economical tendency changes;
- in industrial processes, for detection of malfunctioning devices or infrastructures, preventive maintenance, alarms generation;
- in medicine for the detection of diseases and symptoms or for abnormal behaviour detection in case of psychologically impaired persons;
- in environment monitoring for detection of parameter changes, contaminations, air pollution;
- in informatics systems for the detection of cybernetic attacks on computing devices, networks of software.

The paradox is that the same anomaly detection technique may be applied for a wide variety of domains and application types. Therefore, we consider that there should be a general anomaly detection tool that may be adapted for different particular domains.

2. Taxonomy of anomaly detection techniques

Because there are many types of anomalies and anomaly detection techniques we considered that it would be useful to define a kind of classification or a taxonomy [5]. Analysing the literature in the field we identified a number of criteria that can be used for classification:

- based on the nature of the anomaly;
- based on the nature of the analysed data;
- based on the evaluation methods;
- based on the nature of the applications.

According to the first criterion the following types of anomalies can be identified:

- single point anomalies or outliers – a value that significantly differ from the rest of the values; examples are: wrong measurements, erroneous data transmission; [1]

– contextual anomaly – where a value/data is very different from a closer context – a more complex anomaly that take into consideration a value's neighborhood; examples are: deviations on the stock market, pixels in an image affected by noise; [2]
– anomalies in time series – where the anomaly is detected as a deviation form a normal trend; e.g. artefacts in an ECG signal. [4]

Based on the nature of analysed data we can take into consideration the number of attributes contained in an observation and the type of correlation existing between observations. From the first point of view we have single variable observations (e.g. temperature variations) and multivariable observations (e.g. EEG signals). Based on the second view there may be:

– data with statistical correlations between observations (e.g. the marks for a discipline follow a Gauss distribution);
– time-based correlation – where the values sampled in time have some kind of correlation (e.g. process parameters);
– spatial correlation – where data measured in neighbouring nodes are somehow correlated (e.g. temperature measurements made by a number of sensors spread in a given region);
– functional correlation – where the physical and chemical law determine some rigorous correlation between measured values (e.g. electrical parameters such as voltage, current and power in an electrical infrastructure).

The anomaly detection methods applied for a given case should take into consideration these types of correlations present in the input dataset.

We can also define different informatics approaches for solving the anomaly detection problem. Some of them are based on system analysis and signal processing methods such as min-max, thresholding, interpolation, system identification, Fourier or Laplace transforms. Other methods are inspired from artificial intelligence such as classification and clustering (e.g. KNN, K means, SVM, random forest, etc.) or neural networks. In this case through a process of training a given classifier may distinguish between normal and abnormal observations.

And finally, based on the type of the applications we may distinguish between techniques that may be applied in off-line mode, where there are no real-time restrictions and any complex method can be applied and techniques for on-line mode, where strict time restrictions limit the use of complex methods that are time consuming.

3. A software platform for anomaly detection

In order to offer a generic tool for anomaly detection we developed a platform which is a collection of procedures useful in the process of data analysis and anomaly detection [6, 7].

The motivations for such an approach were:

– there are many detection techniques and it is difficult to decide from the start which will be the best one for a given dataset;
– beside anomaly detection usually we need other data preprocessing and visualization tools which should be part of a platform;
– specialists from different domains (physics, chemistry, biology, environmental or earth sciences) may not have enough knowledge in artificial intelligence or signal processing domains but they need anomaly detection tools.

The platform is a configurable tool in which a user can define a data processing pipeline containing different preprocessing, classification and visualization procedures that are adapted for a given dataset or application. The platform was initially developed for anomaly detection but it can be useful for other purposes where datasets must be classified.

3.1 AutomaticAI – a tool for selection of best classifier algorithm

If we consider just the artificial intelligence methods and their variations , there is a huge variety of choices and unfortunately there is no unique best solution for all possible applications. In a classic research approach, a limited number of methods are tested and compared and from that set the best performing one is selected. But there may be other methods (untested) which may perform even better.

There is also the problem of setting the optimal parameters for a given method; the obtain results may vary drastically with the values of the setting parameters. The researcher should select the best method form a multidimensional search space, which may be a very time consuming process. [8].

In order to solve this selection and parameters' setting problem we proposed and implemented an automatic method that try to find the best performing method from a wide variety of possibilities. [9]

Our solution is based on an optimization technique built upon particle swarm optimization and simulated annealing methods. [10, 11]

In principle the method works as follows:

– in the first step a number of particles are defined, that represent different classification techniques in different settings;
– in the next repeating step, through a number of epochs the best algorithm and its best setting is selected using the particle swarm optimization technique; each algorithm is trained on the given dataset and a performance function determines which has the best quality parameters. In the next epoch the best performing algorithm set is preserved and through mutations (e.g. parameter changes) new particles are created. Through simulated annealing the mutation degree (e.g. parameter changes) is reduced as the simulation epochs pass;
– finally the method presents the best performing algorithm and their settings together with their quality measures (e.g. accuracy, precision, recall, F1, confusion matrices.

The experiments made on different datasets from different domains showed the efficiency of the proposed method; in each case the automated method found a solution in a reasonable time that gave quality parameters comparable with methods proposed by other research groups. The experiments also showed that there is no unique best solution for all the studied cases. Therefore, the process of finding the best solution is a necessary step.

4. Experiments with the AutomaticAI platform

In this chapter we present a number of experiments in different domains performed in order to validate the efficiency of the proposed AutomaticAI tool.

4.1. Anomaly detection in water infrastructures

In this experiment the goal was to find a method for determining the quality of the water in a distribution infrastructure and generate alerts in case of contamination.

In this experiment we used a dataset called GECCO 2017 [12], that contains approximately 100.000 observations/probes, and each probe include 9 parameter values, that are relevant for the quality of the water (e.g. temperature, pH, conductivity, chlorine content, turbidity).

After applying the AutomaticAI procedure we obtained the best performing algorithms with their best setup. Table 1 shows the results for the first 10 classification algorithms. It seems that in this experiment the Random forest offer very high rate of anomaly (contamination) recognition and the Ridge classifier showed the poorest results.

Table 1. *Algorithms selected for water quality classification. (set parameters: class_weight= 'balanced', max_depth=42, n_estimators=130)*

Algorithm	F1-Measure
RANDOM FOREST	99.92%
EXTRA TREES CLASSIFIER	99.81%
DECISION TREE	99.19%
MLP	99.49%
KNN	99.47%
One-Class SVM	81.46%
SGD CLASSIFIER	50.36%
LOGISTIC REGRESSION	49.66%
PASSIVE AGRESSIVE CLASSIFIER	45.36%
RIDGE CLASSIFIER	37.34%

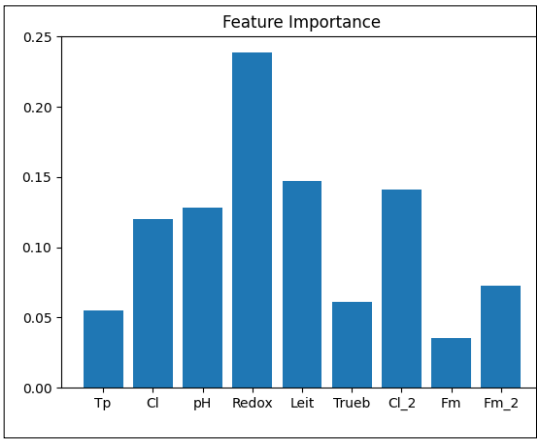


Fig. 1. *Feature importance in the classification of water probes.*

Through other methods included in the platform it could be calculated which parameter had the highest influence in the decision. Figure 1 show this dependence.

4.2. Corona virus infection as anomaly detection

In this experiment our goal was to detect corona virus infection from usual blood analysis without the need for specialized tests (e.g. PCR tests). The need for such an approach was evident in the first phase of the pandemic when PCR tests were rare [13].

In this experiment our goal was to detect corona virus infection from usual blood analysis without

the need for specialized tests (e.g. PCR tests). The need for such an approach was evident in the first phase of the pandemic when PCR tests were rare.

4.3. Cancer detection from colorectal images

In this experiment the goal was to classify colorectal images as containing cancer polyps or not. Traditionally through colonoscopy a trained medical person can distinguish between images with and without cancer formations (polyps). **Figure 2** shows the difficulty of separating normal colon images from those with cancer.

In this experiment we used a pre-trained convolution neural network called ResNet50 in order to extract relevant features from images. Then, using AutomaticAI we determined which classifying algorithm performs best in distinguishing between images containing cancer formations of not. In this experiment the Ridge classifier proved to be the best one with Accuracy=98.33%, Precision=100% and Recall=76.64%.

It is interesting to mention that in the first presented experiment the Ridge Classifier had the poorest results.

Fig. 2. *Colonoscopy images; first two with cancer and the last without.*

4.4. Anomaly detection in esophageal measurements

In this experiment **[14, 15]** we analysed acoustic 2D images from a device called High Resolution Manometry (HRM). (**Figure 3**). This device is used for the detection of abnormal functioning of esophagus in the process of swallowing.

The goal in this case was twofold:

– to identify wrong positioning of the probe in the esophagus;
– classify diseases related with the swallowing process.

In each case a set of images obtained from a HRM device were labelled by trained specialists as good of wrong (bed positioning) images and images representing different diseases.

The difficulty in this case was the relatively small number of labelled images available for training and that sometimes, even for the human observer, it was difficult to discriminate between wrong positioning of the probe and a given disease. Because of this the quality parameters are not as good as in the previous experiments, but still very good in this context. In the case of probe positioning error detection, the precision obtained was 90,67% and the F1 parameter was 84.21%. For classification of diseases we used a classical algorithm present in the medical literature.

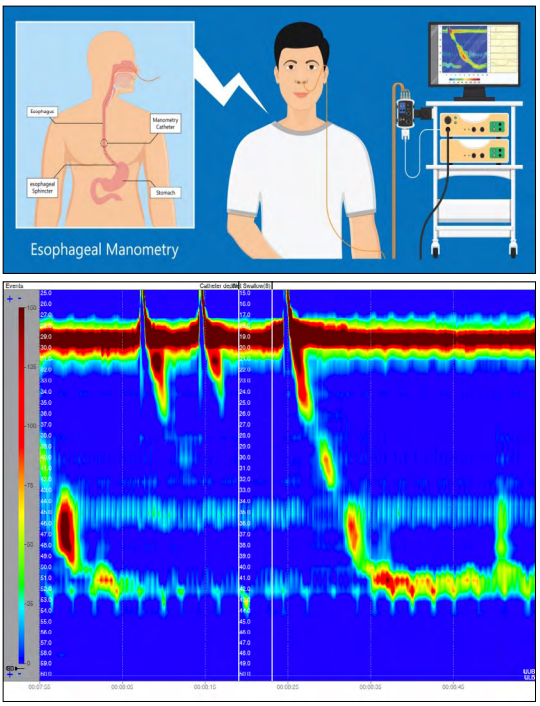


Fig. 3. *HRM measurement- device and obtained 2D image.*

5. Future work

In order to provide access to our anomaly detection platform for a wider category of specialists from different domains we intend to deploy the platform on a cloud infrastructure. A first version was successfully deployed on a private cloud, but for open access services a public cloud is needed. The cloud deployment can benefit from higher computing resources, that may contribute to the reduction of time needed for the automatic selection and training of anomaly detection algorithm.

We also intend to diversify the domains in which anomaly detection and our platform is used. We also intend to add to the platform new artificial intelligence algorithms as well as signal processing detection methods.

6. Conclusions

As showed in this paper, anomalies play an important role in the lifetime of an application. Automatic anomaly detection is more and more a mandatory functionality of a system, mainly for those cases when systems have an autonomous life. It is also important for safety-critical, fault-tolerant and self –healing systems. Anomaly detection techniques can be used also in different medical areas in order to identify diseases as deviations from the normal physiological behaviour of human body.

A specialist in a given domain needs an efficient tool in order to identify and separate or eliminate anomalies for datasets. For this purpose, we developed a configurable platform that contains multiple procedures necessary in the process of data analysis and anomaly detection.

The most important part of the platform is the AutomaticAI tool that automatically selects the best performing classification algorithm for a given dataset. The efficiency and versatility of this automatic method was proved through a number of experiments performed in a variety of domains: water infrastructures, medical imagery, covid detection, etc.

References

[1] Chandola V. , Banerjee A., Kumar V.: *Anomaly Detection: A Survey.* ACM Computing Surveys, 41/3. (2009) 1–58. https://dl.acm.org/doi/10.1145/1541880.1541882

[2] Agrawal S., Agrawal J.: *Survey on Anomaly Detection using Data Mining Techniques.* Procedia Computer Science, 60/1. (2015) 708–713. https://doi.org/10.1016/j.procs.2015.08.220

[3] Gupta M., Gao J., AggarwalC. C. , Han J.: *Outlier Detection for Temporal Data: A Survey*. IEEE Transactions on Knowledge and Data Engineering, 26/9. (2014) 2250–2267. https://doi.org/10.1109/TKDE.2013.184

[4] Hodge V. J., Austin J.: *A Survey of Outlier Detection Methods.* Kluwer Academic Publishers, 2004.

[5] Czako Z., Sebestyén Gy., Hangan A.: *Colorectal image classification with transfer learning and auto-adaptive artificial intelligence platform*. In: Trends and Innovations in Information Systems and Technologies 28. Springer International Publishing, 2020. 534–543.

[6] Czako Z., Sebestyén Gy., Hangan A.: *AutomaticAI–A hybrid approach for automatic artificial intelligence algorithm selection and hyperparameter tuning.* Expert Systems with Applications, 182. 2021. https://doi.org/10.1016/j.eswa.2021.115225

[7] Czako Z., Sebestyén Gy., Hangan A.: *Artificial Intelligence Algorithms Selection and Tuning for Anomaly Detection.* Computational Intelligence: International Joint Conference, IJCCI 2018 Sevilla, Spain. Revised Selected Papers, 2021.

[8] Adankon M. M., Cherirt M.: *Model Selection for LS-SVM: Application to Handwriting Recognition.* Pattern Recognition, 42/12. (2009) 3264–3270. https://doi.org/10.1016/j.patcog.2008.10.023

[9] Sebestyén Gy., Hangan A., Czako Z., Kovács Gy.: *A taxonomy and platform for anomaly detection.* 2018 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR), 2018. https://doi.org/10.1109/AQTR.2018.8402710

[10] Pradeepmon T., Panicker V., Sridharan R..: *Parameter selection of discrete particle swarm optimization algorithm for the quadratic assignment problems.* Procedia Technology, 25. (2016) 998–1005. https://doi.org/10.1016/j.protcy.2016.08.199

[11] Simsek A., Kara R.: *Using Swarm Intelligence Algorithms to Detect Influential Individuals for Influence Maximization in Social Networks*. Expert Systems with Applications, 114. (2018) 224–236. https://doi.org/10.1016/j.eswa.2018.07.038

[12] Sebestyén Gy., Hangan A., Czako Z.: *Anomaly detection in water supply infrastructure systems.* 23rd International Conference on Control Systems and Computer Science (CSCS), 2021. https://doi.org/10.1109/CSCS52396.2021.00064

[13] Czako Z., Sebestyén Gy., Hangan A.: *COVID-19 Preliminary Patient Filtering based on Regular Blood Tests using Auto-Adaptive Artificial Intelligence Platform.* IEEE 16th International Conference on Intelligent Computer Communication and Processing (ICCP), 2020. https://doi.org/10.1109/ICCP51029.2020.9266277

[14] Popa Ș. L., Surdea-Blaga T., Dumitrașcu D. L., Sebestyén Gy., et al.: *Automatic Diagnosis of High-Resolution Esophageal Manometry Using*

Artificial Intelligence. Journal of Gastrointestinal & Liver Diseases, 31/4. (2022) 383–389. https://doi.org/10.15403/jgld-4525

[15] Surdea-Blaga T., Sebestyén Gy., Czako Z., Hangan A., et al.: *Automated Chicago Classification for Esophageal Motility Disorder Diagnosis Using Machine Learning.* Sensors, 22/14. (2022). https://doi.org/10.3390/s22145227

SZERZŐK JEGYZÉKE
LIST OF AUTHORS