Study of the Mechanical Characteristics of Sandwich Structures FDM 3D-printed

Rawabe Fatima Faidallah[1], Muammel M. Hanon[2*], Zoltán Szakál[3], István Oldal[3]

[1] Mechanical Engineering Doctoral School, Szent István Campus, MATE University, Páter Károly u. 1, 2100 Gödöllő, Hungary; faidallah.rawabe.fatima.2@phd.uni-mate.hu

[2] Baquba Technical Institute, Middle Technical University (MTU), Muasker Al-Rashid Street, 10074 Baghdad, Iraq; muammel.m.hanon@mtu.edu.iq

[3] Institute of Technology, Szent István Campus, MATE University, Páter Károly u. 1, 2100 Gödöllő, Hungary; szakal.zoltan@uni-mate.hu, oldal.istvan@uni-mate.hu

Abstract: Additive manufacturing (also known as 3D printing) technologies are successfully used in various applications as they offer many advantages in production, such as (i) less material consumption and (ii) shorter manufacturing times, factors that reduce costs. In recent years, experts in 3D printing have focused their studies on designing and printing cellular structures. This structure's advantages (high strength-to-weight ratio, high flexural stiffness-to-weight ratio, high-energy absorption capacity, thermal and acoustical insulation properties) make it widely used in aerospace, sustainable energy, marine, and automotive industries. This study aims to study the mechanical properties of sandwich structures manufactured using polylactic acid (PLA) material with rhombus and honeycomb core shapes as a single part by an FDM 3D printer. First, the functional properties of the sandwich structures were quantified by shape evaluations. Then, tensile, three-point bending, and compression tests were performed to determine the mechanical performance of the different samples. The results show that rhombus structures gave better mechanical behaviour as the tensile, bending, and compression strengths were 15.3%, 39.8%, and 35.1%, respectively, higher than the honeycomb, indicating their reliable core construction.

Keywords: Polylactic acid; 3D printing; Honeycomb core structure; Rhombus structure; Mechanical performance

1 Introduction

Additive manufacturing (AM), alias 3D printing, can be defined as the "process of joining materials to make parts from 3D model data" [1]. Often the products are constructed layer upon layer [1, 2], as opposed to formative and subtractive manufacturing methods. In the AM process, the material is fused, cooled, and solidified, thus obtaining 3D geometries without adopting complex moulds [3]. The product details are taken from a computer-aided design (CAD) file, which is later converted into a stereolithography (STL) file. The model is created in a 3D CAD software (e.g., Solidworks) and is approximated by triangles and cut into slices containing each layer's information to be printed [4]. Nowadays, additive manufacturing is involved in mechanical engineering applications for the research and development of different elements extending from simple constructs employed in daily life up to the complex components in aerospace implementations [5, 6].

Fused deposition modeling (FDM) is one of AM technologies that provides excellent mechanical, chemical, and thermal endurance and has become one of the most widely utilized in polymer additive manufacturing. The FDM technology approach is remarkably simple to use and set up in comparison with other AM technologies [7]. FDM is based on melting plastic filaments [8]. The FDM process is performed by extruding thermoplastic material, which has to be heated up to its melting point through a nozzle, then depositing the extruded layers of materials on top of each other [9]. The parameters of FDM process like raster angle, printing speed, printing orientation, and layer height have a significant influence on print characteristics. In order to determine the qualities of components, researchers studied the impacts of various process parameters on responses [7]. Currently, FDM is considered the most widely used technology of all types of 3D Printing techniques due to its low cost of printer devices, simplicity, and variety of inexpensive filaments [10, 11]. FDM 3D printers are commonly applied in different industries like aeronautics, construction, automotive, and medicine for rapid prototyping. Various thermoplastic polymers like acrylonitrile butadiene styrene (ABS), polylactic acid (PLA), and polyimide are used as the material of FDM technology in the shape of filament. ABS material has better elongation, ductility, and flexural strength than PLA material, but it emits an undesired smell in the printing process. On the other hand, PLA material is more environmentally friendly than ABS material because it degrades faster than ABS material and is produced from renewable resources [12].

Simultaneously, lightweight sandwich structures are being used in many industries such as automotive, sustainable energy, and building due to their effect on improving mechanical properties [13]. The frequent use in engineering applications of lightweight sandwich structures shows better results in acceleration, lower fuel and consumption (for aircraft), in addition, to lower life cycle costs due to lower operation costs for many applications. However, the traditional methods of manufacturing and structures include many stages that require the purchase of complex devices, which is usually difficult to use and make production expensive. Meanwhile, current processes in AM techniques offer great freedom for both realising and designing honeycomb structures with complex geometries. Because of that, a vast number of studies have been carried out to explore the correlations between specific mechanical properties and 3D-printed complex geometries [14-16]. Furthermore, recent research is focused on numerical analysis and geometric design, showing that geometric shape plays a crucial role in their mechanical response [15, 17, 18]. In addition, researchers [19] have been focused their studies on understanding the effect of different sandwich structures (triangle, square, and hexagon shapes) by considering numerous parameters such as the number of infill shapes, nature of the core shape, the orientation of cores, and the influence dynamic behaviour of sandwich structures.

Mechanical performance (bending, compressive, tensile, etc.) of sandwich structures (e.g., honeycomb and rhombus) correlates not only with structure geometry but is also strongly dependent on material properties [20-22]. An initial investigation on the manufacture and testing of mechanical performance (tensile and bending) of honeycomb core sandwich structures printed with different materials (PLA and ABS) and using different filling techniques [23, 24]. PLA is one of the most common materials used in thermoplastic FDM feedstock, which has good printability due to its crystalline phase at room temperature [25]. Also, PLA has excellent mechanical qualities and because of its biodegradability, can be used to replace petroleum-based polymers. The usage of PLA has significantly increased during the past ten years, mostly in the fields of biomedicine and packaging [26, 27]. A unique generation of green composites, or composite materials that are friendly to the environment, may be created with the aid of PLA. Furthermore, PLA is increasingly used in a number of components of composite sandwiches, including adhesion nonwovens, 3D-printed honeycombs, and skins on various composite panel configurations [28]. Additionaly, the potential of PLA-derived materials for use in automotive components has been proved by Nickels [29]. Researchers [30] have also focused their studies on estimating the modal parameters and the properties of damping through experimental dynamic analysis for PLA. However, limited studies have focused on using PLA material in the 3D printing of sandwich structures which could enhance its mechanical behaviour [31, 32].

In this study, the rhombus and honeycomb structures were designed and manufactured with PLA material by FDM printing technology. The material's mechanical performance in the case of sandwich structures was evaluated by conducting mechanical tests (tensile, three-point bending, and compression). The failure mode of specimens for each testing carried out was also analysed and highlighted. However, the static analysis was only investigated in the present work. The outcome of this study could be helpful in examining the capability of PLA sandwich structure to create lightweight materials with equitable mechanical properties and most importantly to support sustainable development. This will contribute to filling the gap in knowledge regarding the performance of sandwich structures desired in such applications because of their light weight and reliability.

2 Materials and Methods

2.1 Constructing of Sandwich Structure Specimens

The FDM technique was used to manufacture the test specimens with the commercial 3D printer WANHAO Duplicator 6 and PLA (polylactic acid) material. PLA was chosen as the preferred material because of its growing popularity in FDM printing and its sustainable environmental characteristics. The filament used was 1.75 mm in diameter and black in colour. As stated by the manufacturer, the filament's claimed qualities are tensile Young's modulus, ultimate tensile strength, strain at yield, and impact strength of 3.120 GPa, 70 MPa, 5%, and 3.4 KJ/m2, respectively, when tested according to ISO 527. WANHAO 3D printer has a 0.4 mm nozzle diameter and exists in the additive manufacturing laboratory at MATE (Szent Istvan University formerly), Gödöllő, Hungary. The test samples were printed with an infill density of 100% and a layer thickness of 200 µm at a print speed of 60 mm/s. The printing and building plate temperatures were set as 210 °C and 60 °C, respectively.

Sandwiches typically consist of two skins (outer surfaces) and a core (inner structure). The core's material may be the same or unlike the skins. In the current study, the material utilised for both outer surfaces and inner structure was the same (PLA), and the manufacturing, by 3D printing, was done in solely a single stage. Two sandwich structures, honeycomb and rhombus (see Figure 1) have been used. These structures were designed in accordance with the standards ASTM C393 and MIL-STD-401B. A three-dimensional designing software SOLIDWORKS 2016 was used to generate the 3D structure model of the samples and convert it into an "STL" file format. In order to slice the STL file (3D model), Ultimaker Cura 4.7 software was employed. The Cura program translates the digital model into a set of instructions (G-code) for the 3D printer, and through this software, the manufacturing parameters are established. The designed specimens, with the honeycomb and rhombus sandwich structures, were prepared for the tensile, bending, and compression testing, as shown in Figure 1a, Figure 1b, and Figure 1c, consecutively. Magnification for the sandwich structures' core shape is aside demonstrated in Figure 1: honeycomb core (left) and rhombus core (right). The specimens were built at an on-edge orientation, as illustrated in Figure 2d.

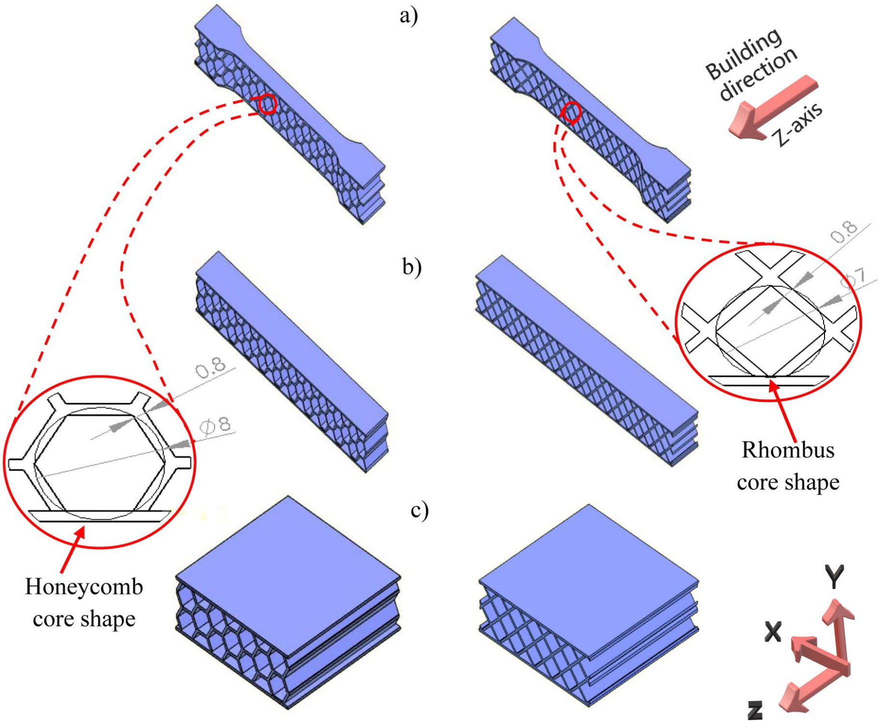


Figure 1
3D models of the sandwich structure specimens of the honeycomb (left) and rhombus (right) patterns for a) tensile test, b) three-point bending test, and c) compression test

According to the appointed standards, the tensile and three-point bending test specimens have the dimensions of 150 mm, 20 mm, and 15 mm for length, width, and height, respectively, with 0.75 mm as the thickness of the skin (see Figure 2a-b). Also, the compression test samples were designed with a length of 50 mm, a width of 50 mm, and a height of 15 mm, at a skin thickness of 0.75 mm (see Figure 2c). Three identical specimens were 3D-printed for each test condition. The cell wall thickness should be sufficient so that it can be easily printable by the FDM machine since too thin a cell wall thickness might make printing difficult and can deform the object. Therefore, the proper cell size was chosen to make the cell wall thickness thick enough to be easily printed. The chosen cell sizes were 8 mm and 7 mm for the honeycomb and rhombus, respectively. In addition, the cell wall size for both the honeycomb and rhombus was 0.8 mm (see core shape aside in Figure 1). However, it is worth noting that it is not easy to specify the accurate or optimal 3D printing parameters due to the anisotropic nature caused by the technology owing to the variety of its printers and materials [33]. Figure 2d shows the actual appearance of some of the manufactured samples. The tensile specimens were produced with structure support owing to their geometry as there is a wide space under their gauge section. However, there was no need for structural support while creating the bending and compression test pieces. The average density values for the specimens examined are tabulated in Table 1.

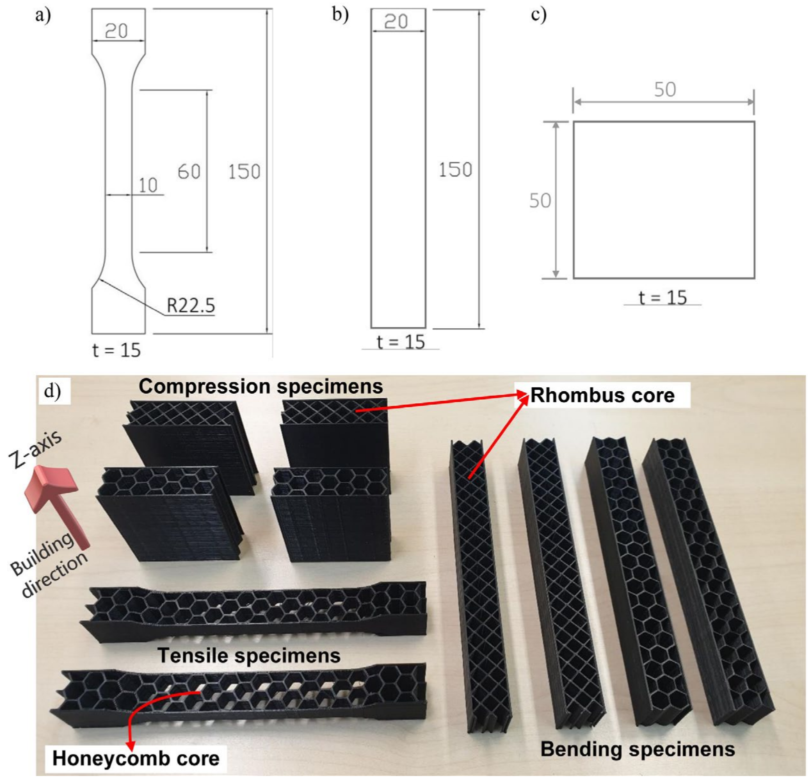


Figure 2
Specimens' dimensions of a) tensile test, b) bending test, c) compression test; and d) physical appearance of some of the test specimens after 3D printing

Table 1
Average values of mass and density (with its standard deviation (SD)) of specimens tested

Structure	Mass (g)	Volume (mm^3)	Density (g/mm^3)	SD (±)
Tensile honeycomb	10.408	31500	$0.33*10^{-3}$	$6.73*10^{-6}$
Tensile rhombus	11.639	31500	$0.369*10^{-3}$	$4.93*10^{-7}$
Bend. honeycomb	14.955	45000	$0.332*10^{-3}$	$1.47*10^{-6}$
Bend. rhombus	17.228	45000	$0.382*10^{-3}$	$2.56*10^{-6}$
Comp. honeycomb	12.063	37500	$0.321*10^{-3}$	$1.22*10^{-6}$
Comp. rhombus	14.533	37500	$0.387*10^{-3}$	$1.15*10^{-6}$

2.2 Experimentations

The universal test machine (Zwick / Roell Z100, Germany) was used to test sandwich structure specimens for compression, tensile, and three-point bending. Three repeated tests have been accomplished on the samples that were constructed for each configuration of the sandwich structure's core (honeycomb and rhombus). The following sections clarify the conditions and details of the tests performed.

2.2.1 Tensile Testing

The ISO standard 527 for tensile testing of polymers [34] was used to assess the tensile properties of the built structures. The tensile behaviour determination included the tensile strength, tensile Young's modulus, and the failure form under the specified conditions. According to the ISO 527-2:2012 standard [35], the specimens were stretched at a steady speed of 1 mm/min along their main axis until they broke (Figure 3).

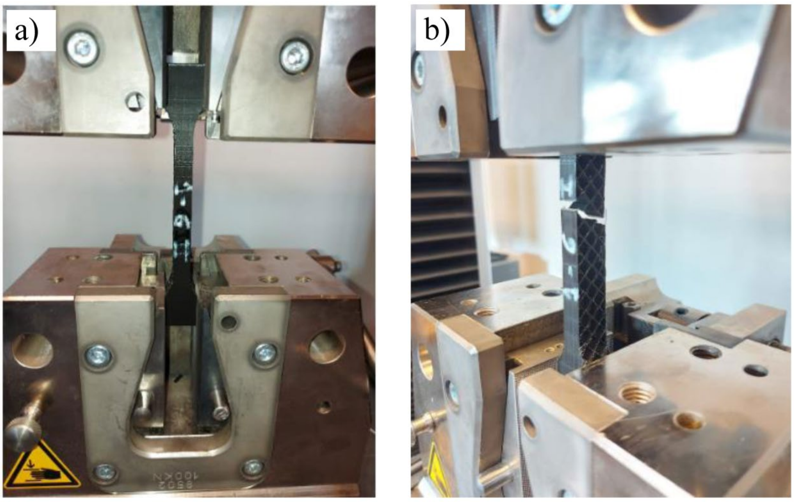


Figure 3
Tensile testing for one of the sandwich structure's specimens (rhombus), a) during the test and b) after breakage

Equation (1) was used to calculate the tensile strength (σ_t) of sandwich specimens

$$\sigma_t = P_t/A_t\,, \qquad (1)$$

where P_t stands for the ultimate load (N), while the sandwich specimen's cross-sectional area is represented by A_t (mm^2). In this study, the broken cross-section was chosen to calculate the tensile cross-sectional area, where it was assumed to be the weakest point (having the smallest cross-sectional area) within the gauge section of the specimens. For this, the Solidworks software was used to find and calculate the smallest cross-sectional area, as illustrated in Figure 4, where it was 30 mm^2 and 32 mm^2 for honeycomb and rhombus, respectively.

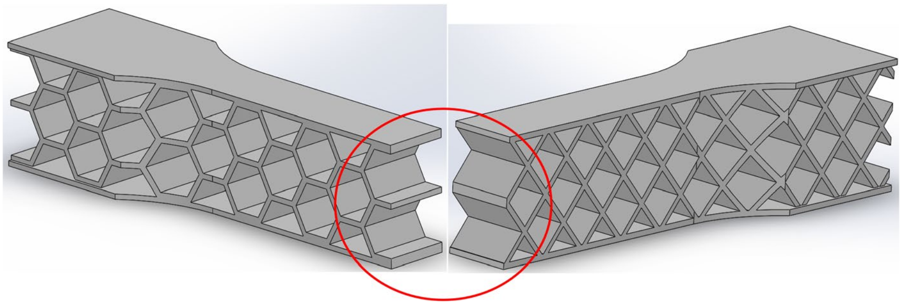


Figure 4
The cross-sectional area of tensile specimens determined by the Solidworks software

Tensile modulus (E_t) was calculated by Hooke's law

$$E_t = \sigma_t/\varepsilon \,, \tag{2}$$

as ε is the strain.

To get a more accurate value of Young's modulus (E_t), two points were fitted on the stress-strain curves of the sandwich specimens to draw the slope and determine the tensile modulus. These two points were specifically at 10% and 60% from each axis on the stress-strain curve (i.e., the values of σ_t and ε were picked at 10% and 60% and then substituted into the slop equation; $\sigma_{t(\text{at } 60\%)}$- $\sigma_{t\ (\text{at } 10\%)}$ / $\varepsilon_{(\text{at } 60\%)}$- $\varepsilon_{(\text{at } 10\%)}$.

2.2.2 Bending Testing

The bending tests (three-points) were carried out in accordance with ASTM C 393 [36]. The crosshead speed of the tests was 1 mm/min until the specimen broke. In the three-point bending, the radius of supports and punch was 15 mm, and the span length was 100 mm. Equations (3) and (4) were used to calculate the values of bending strength (σ_b) as well as the bending modulus (E_b) of sandwich samples [37]

$$\sigma_b = (3P.s)/(2b.d^2) \,, \tag{3}$$

$$E_b = (s^3.m)/(4b.d^3) \,, \tag{4}$$

where P stands for the force (N) at a particular point on the load-deflection curve, while s is the support span length (mm), in addition, b and d are the width (mm) and thickness (mm) of the sandwich specimen, respectively. Also, the tangent's slope to the load-deflection curve's initial straight-line component (N/mm) is m. The experimental setup of the three-point bending test is shown in Figure 5a.

2.2.3 Compression Testing

The compression of the sandwich structure specimens was carried out on a mechanical test machine (Figure 5b). The tests were performed at a 1 mm/min crosshead speed. For reliable results, each lightweight sandwich construction investigated (honeycomb and rhombus) was subjected to three tests.

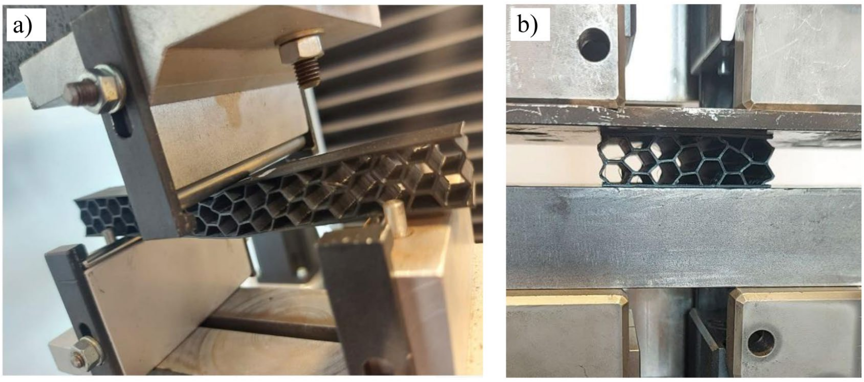


Figure 5
Experimental setup of testing one of the sandwich structures (honeycomb) for the tests of a) three-point bending and b) compression

The compressive strength (σ_c) value of the sandwich specimens was calculated using Equation (5)

$$\sigma_c=P_c/A_c, \tag{5}$$

where P_c represents the ultimate load (N) on the compression tests, and A_c is the sandwich specimens' cross-sectional area (mm^2).

Equation (6) was used to determine the compressive modulus (E_c)

$$E_c = (m.t)/A_c, \tag{6}$$

where m is the tangent's slope to the load-deflection curve's initial straight-line component (N/mm) and t denotes the core thickness (mm).

In order to calculate the cross-sectional area of the compression samples, and based on a solid structure, the average area (A_{av}) of the structure can be determined as follows, $A_{av} = V_{str}/t$ where V_{str} represents volume of sandwich structure and t is the thickness.

Since density (ρ) is the mass (M) over volume (V), thus,

$$\rho_{av}=M_{str}/V, \tag{7}$$

and

$\rho_{pla}=M_{pla}/V \rightarrow$

$$V= M_{pla}/\rho_{pla}, \tag{8}$$

where ρ_{av} and M_{str} denote the density and mass, respectively, of the core sandwich structure. Meanwhile, ρ_{pla} and M_{pla} are the density and mass of a solid structure for PLA polymer.

Substituting Eq (8) in Eq (7),

$\rho_{av}=(M_{str}/M_{pla})\times\rho_{pla} \rightarrow$

$$M_{str}=(\rho_{av}/\rho_{pla})\times M_{pla} \tag{9}$$

$\because V_{str}=A_{av}.t$, and $\rho_{pla}=M_{str}/V_{str}$,

$\therefore \rho_{pla}=M_{str}/(A_{av}.t) \rightarrow$

$$A_{av}=M_{str}/(\rho_{pla}.t) \tag{10}$$

Putting (9) in (10),

$$A_{av}=(M_{pla}/\rho_{pla}.t)\times(\rho_{av}/\rho_{pla}) \tag{11}$$

$\because \rho_{pla}=M_{pla}/(A_{pla}.t)$

$$\therefore A_{pla}=M_{pla}/(\rho_{pla}.t) \tag{12}$$

From (11) and (12),

$$A_{av}= A_{pla}\times(\rho_{av}/\rho_{pla}) \tag{13}$$

The density (ρ_{av}) of the sandwich structure compression specimens is $0.32*10^{-3}$ (g/mm^3) for honeycomb and $0.39*10^{-3}$ (g/mm^3) for rhombus, as listed in Table 1. Additionally, the density of the PLA material (ρ_{pla}), in case of a solid bulk, is $1.252*10^{-3}$ (g/mm^3).

3 Results and Discussion

3.1 Tensile Performance of Sandwich Structures

Figure 6a shows the load-displacement curves that were obtained from the tensile test of the honeycomb and rhombus sandwich structures. Obviously, the rhombus core sandwich samples had the greatest maximum load of 714 N in the load-elongation curves. In terms of the tensile strength, which ranged between 19.49 and 23.01 MPa, the better values were from rhombus core sandwich structures (see Figure 6b). The increased tensile strength of these specimens is due to their reliable core structure, which had more contact sites at the fracture area under tensile stress than the honeycomb core construction (see Figure 7). Therefore, the applied load was distributed across a larger area, resulting in higher resistance to failure. In contrast, honeycomb sandwich structures had the best tensile Young's modulus of 599 MPa compared to rhombus (440 MPa), as shown in Figure 6c. Zaharia et al. [38] have studied the mechanical properties of different sandwich structures (honeycomb, diamond-celled (resembles the rhombus in the current work), and corrugated). They reported a higher load (required to fracture) and tensile strength for the diamond structure than the honeycomb, which is in good agreement with the present study findings.

Figure 7 shows that sandwich specimens seem to be failed first due to the yielding of the face sheet (shell), thereafter the core (inner structure) shear failure. One of the outer surfaces of the sandwich specimens yielded, the fracture progressed through the whole core, and finally terminated at the lower level, causing cracking of the other outer surface. The specimens, as shown in Figures 7a and 7b, had a complete fracture of the whole structure, beginning with the first outer surface (shell), then the core, and lastly the second outer surface, for both (honeycomb and rhombus) tested sandwich structures.

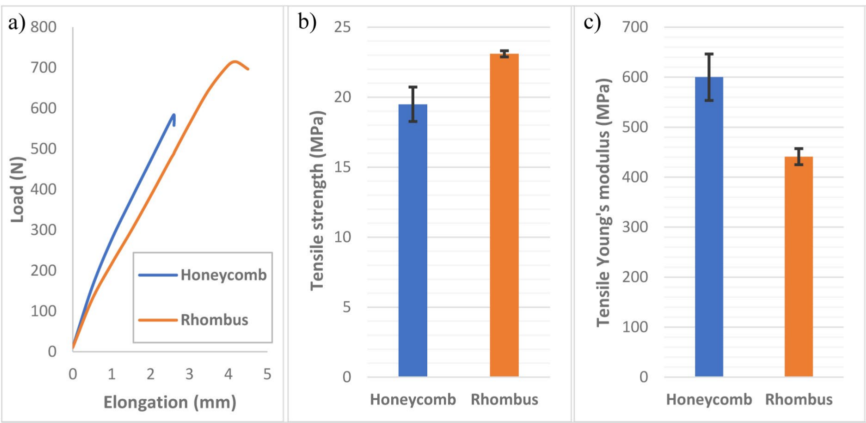


Figure 6

Tensile test result of sandwiches structure specimens, a) load-displacement curves, b) tensile strength values, and c) Young's modulus values

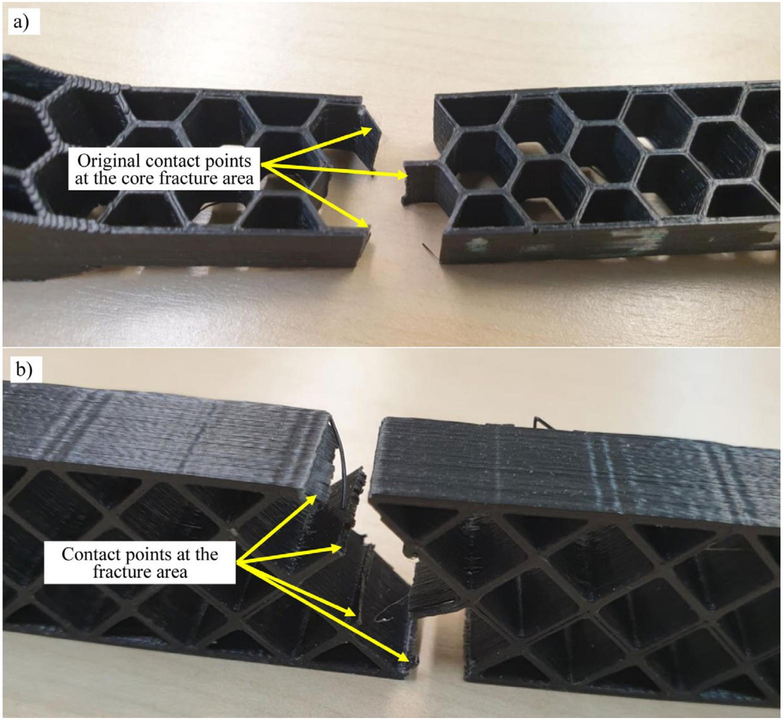


Figure 7
Fracture shape and contact sites at the core structure failure area after the tensile tests of a) honeycomb, and b) rhombus sandwich specimens

3.2 Three-Point Bending Performance of Sandwich Specimens

The flexural performance of the lightweight sandwich constructions (honeycomb and rhombus), including bending strength, bending modulus, and stress-displacement characteristics, was investigated using this test approach. The load-displacement curves of test specimens for three-point bending have two main stages, as shown in Figure 8a: a steady increase between the load applied and the displacement towards the curve's maximum, but then when the specimens broke, there was a sudden decline from the maximum load. Sugiyama et al. [15] attributed this non-linearity or sharp decline exhibited in the load curve to the gradual failure progression. They mentioned that in the fracture mode, the crack initially happens at the upper skin while wrinkling occurs at the core, causing the curve to drop. Using Equations (3) and (4) and the sandwich specimen dimensions, the test machine software automatically determined the bending modulus and bending strength, which are the most essential three-point bending aspects.

Three-point bending results were better for the rhombus than the honeycomb core sandwich structure. The maximum force reported was about 381 N at a displacement of 4.47 mm of rhombus core specimens, according to the curves shown. In terms of the bending strength, Figure 8b exhibits that the rhombus core sandwich models had a bending strength average value of 40% higher than that of the honeycomb core sandwich samples. Furthermore, the bending Young's modulus of rhombus specimens was twice as high as that of honeycomb specimens (see Figure 8c).

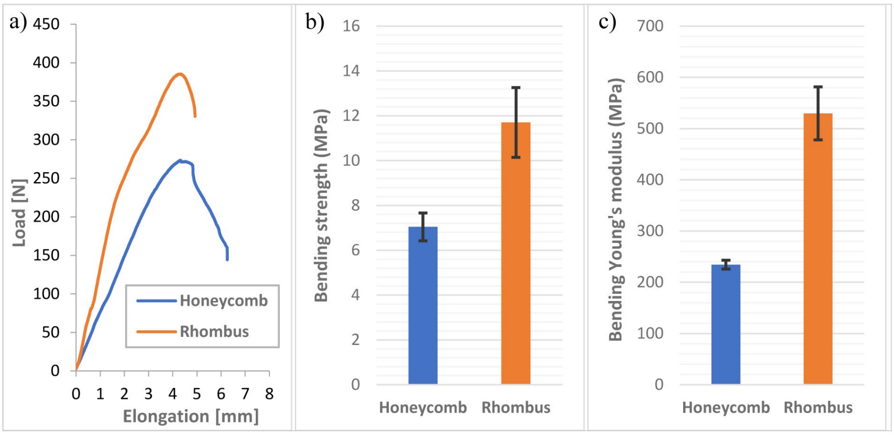


Figure 8

Three-point bending test results of sandwich structures, a) load-displacement curves, b) bending strength, c) Young's modulus of bending

The primary failure scenarios of sandwich structures that occurred during three-point bending tests are shown in Figure 9. The fracture point pictured in Fig. 9a was not at the same loading point while testing. This is due to, at the fracture site, the core being at an angle perpendicular to the longitudinal axis, making fracture/buckling easy [15]. Sandwich specimens with thin skins may readily fail in a skin yield mode because rhombus core sandwich faces/skins practically withstand all tensile and compressive loads in bending [38]. However, an indentation mode occurs initially when the core sandwich walls are thick enough, but given sufficient impact energy, the sandwich structures will eventually collapse in a skin yield fracture mode on the top face [39]. Then, cracks are formed in the rhombus core, followed by propagation of the failure up to the lower face. On the other hand, the bending test results of the honeycomb core structure specimen showed that a tension failure in the upper skin takes place like compression face buckling/wrinkling (see Figure 9b). Local short-wavelength wrinkling of skins is another name for this sort of failure mechanism [40].

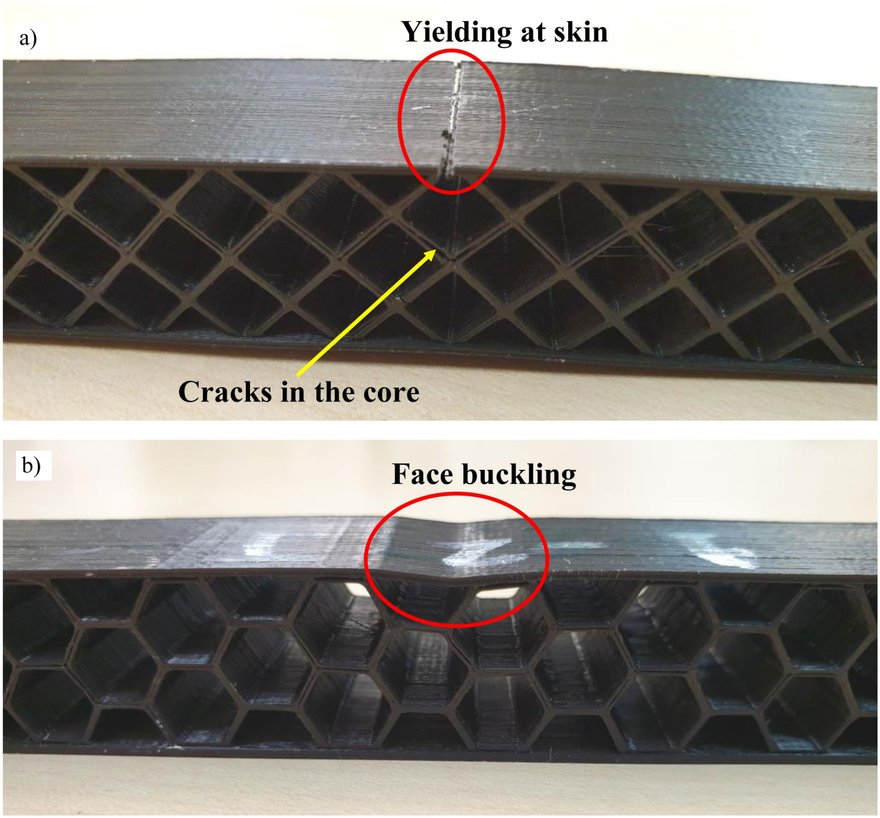


Figure 9
Three-point bending test's failure modes of sandwich specimens; a) rhombus, and b) honeycomb core structures

3.3 Compression Behaviour of Sandwich Structures

The present section discusses the compression test results obtained for 3D-printed rhombus and honeycomb structure specimens. Figure 10 exhibits the compression behaviour gained from the compression tests for each examined type of sandwich structure. It is evident from Figure 10a that the load-displacement responses are generally linear until the core shear starts, at which point there is a dramatic reduction in load. The maximum force (roughly 5850 N) was found in the rhombus core structures (at an elongation of 1.6 mm) until unredeemable damage in this sandwich structure had occurred. However, for the honeycomb structure, irreversible damage occurred when the load force reached 2820 N at 1.35 mm elongation.

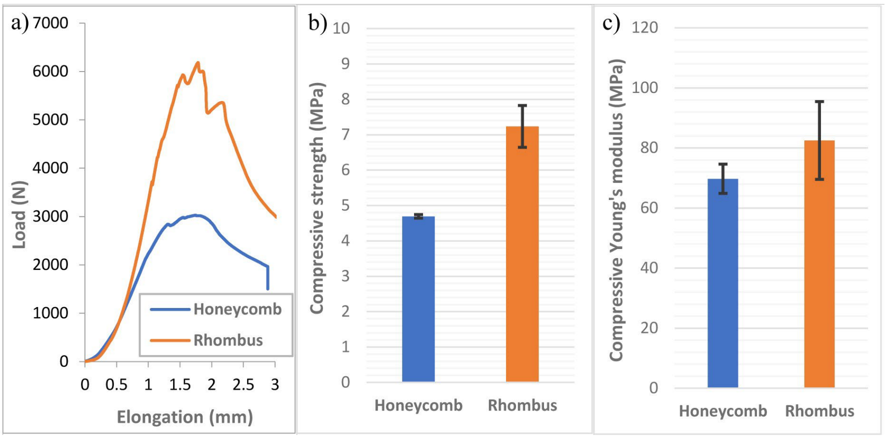


Figure 10
Compression test results, a) load-elongation curves, b) compressive strength, and c) Young's modulus of compression

The results of compressive stress and modulus are represented in Figure 10b and Figure 10c, respectively. The cross-section area was calculated using Equation (13), as mentioned in section 2.2.3.

Again, the rhombus core sandwich structure offered the best performance as the average of its compressive strength (7.23 MPa) was 35.1% higher than that of the honeycomb (4.69 MPa) and its compressive modulus (82.47 MPa) was 15.4% better than the honeycomb (69.73 MPa). The dense network of its structures, which also caused these specimens to weigh more (as shown in Table 1), is responsible for the outstanding performance of rhombus core specimens.

Figure 11 shows the specimens after the compression. For the honeycomb specimen (Figure 11a), deformations develop in the structure's core as the breakdown was the buckling and then shearing of the core of the sandwich. This may be explained by the fact that when sandwich structures are compressed, the skins are too thick and robust to be crushed, resulting in a core buckling failure mode. However, in the case of rhombus structure (Figure 11b), the extruded filament layers were debonded because the sandwich's core had a high degree of flexibility.

Comparing Figures 6 and 10, it can be noticed that the tensile Young's modulus (E_t) is much higher (almost 8 times for honeycomb and 5 times for rhombus) than the compressive modulus (E_c). This significate difference in the magnitudes can be attributed to the influence of test load direction with respect to the sandwich structure. In the case of tensile, the load is parallel to the sandwich structure, and the shell (the outer skins) can support the core structure significantly to provide much more strength. However, for the compression, and due to the loading direction as well, the whole load will be applied on the core sandwich structure, causing failure in it without any support from the outer shells, as shown in Fig. 11, as the skins were not affected. This would make the specimen's compressive strength, and implicitly the compression modulus, much weaker than the tensile ones.

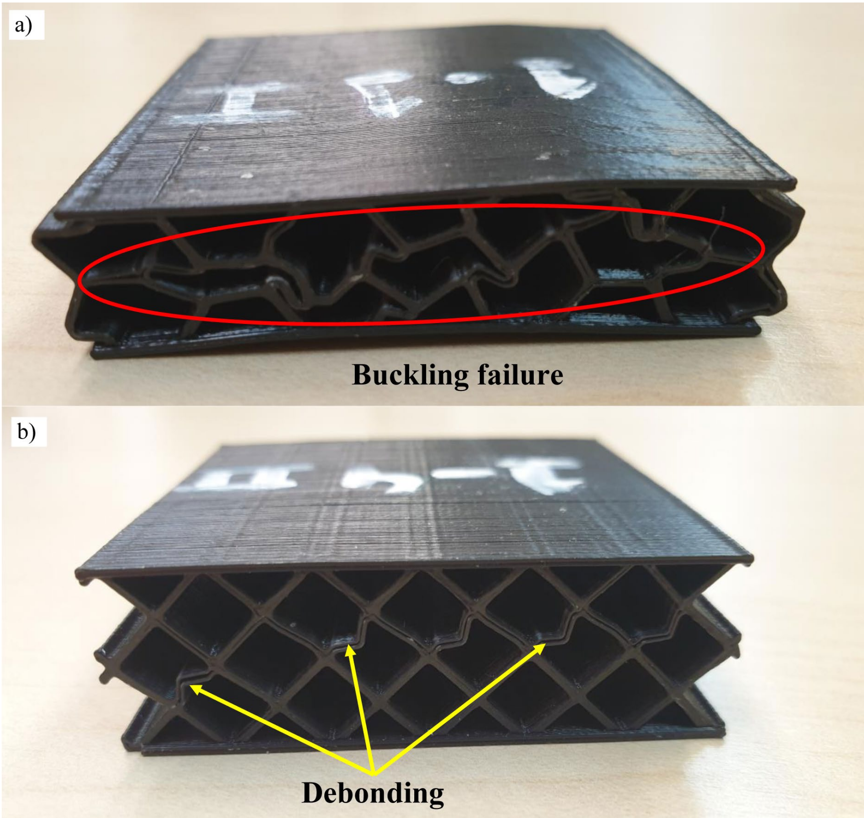


Figure 11

Failures mode of the sandwich specimens after the compression, a) core shearing for the honeycomb sandwich structure, and b) debonding of the extruded filament layers for the rhombus specimens

Conclusions

In the current study, the manufacture and characterisation of 3D-printed two different sandwich structures (honeycomb and rhombus) were executed. At first, using the fused deposition modelling technique, PLA filament was employed to prepare tensile, three-point bending, and compression testing specimens. Then, the characteristics of the sandwich structures fabricated of PLA material were studied based on the tests carried out. It was observed that the rhombus sandwich samples demonstrated the best tensile strength (23.01 MPa), which was 15.3% higher than the honeycomb, due to their reliable core structure. Furthermore, rhombus specimens exhibited the maximum compression strength (7.23 MPa) and flexural strength (11.7 MPa), 35.1% and 39.8%, respectively, higher than the honeycomb. It was noticed that 3D-printed sandwich structural collapse is primarily due to the failure of the core. In the fields of biomedicine and packaging, lightweight sandwich structures are widely utilised to lower the total weight of mechanical components. Therefore, the findings of this work are proposed to provide the industry of these applications with some useful data.

Acknowledgement

This work was supported by the Stipendium Hungaricum Programme and by the Mechanical Engineering Doctoral School, Szent István Campus, MATE University, Gödöllő, Hungary.

References

[1] Erdős F., Németh R., “AMT-based real-time, inter-cognitive communication model”, Acta Polytechnica Hungarica, Vol. 16, No. 6, pp. 115-127, 2019

[2] Rafajłowicz E., “Data Structures for Pattern and Image Recognition and Application to Quality Control”, Acta Polytechnica Hungarica, Vol. 15, No. 4, pp. 233-262, 2018

[3] Hanon M. M., Zsidai L., “Sliding surface structure comparison of 3D printed polymers using FDM and DLP technologies”, IOP Conference Series: Materials Science and Engineering, Vol. 749, pp. 012015, 2020

[4] Antony S., Cherouat A., Montay G., “Fabrication and Characterization of Hemp Fibre Based 3D Printed Honeycomb Sandwich Structure by FDM Process”, Applied Composite Materials, Vol. 27, No. 6, pp. 935-953, 2020

[5] Chaczko Z., Klempous R., Rozenblit J., Adegbija T., Chiu C., Kluwak K., Smutnick C., “Biomimetic Middleware Design Principles for IoT Infrastructures”, Acta Polytechnica Hungarica, Vol. 17, No. 5, pp. 135-150, 2020

[6] Hanon M. M., Ghaly A., Zsidai L., Szakál Z., Szabó I., Kátai L., “Investigations of the Mechanical Properties of DLP 3D Printed Graphene/Resin Composites”, Acta Polytechnica Hungarica, Vol. 18, No. 8, pp. 143-161, 2021

[7] Alzyod H., Ficzere P., "The Influence of the Layer Orientation on Ultimate Tensile Strength of 3D Printed Poly-lactic Acid", JJMIE, Vol. 16, No. 3, 2022

[8] N. Turner B., Strong R., A. Gold S., “A review of melt extrusion additive manufacturing processes: I. Process design and modeling”, Rapid Prototyping Journal, Vol. 20, No. 3, pp. 192-204, 2014

[9] Kun K., "Reconstruction and Development of a 3D Printer Using FDM Technology", Procedia Engineering, Vol. 149, pp. 203-211, 2016

[10] Hanon M. M., Zsidai L., "Comprehending the role of process parameters and filament color on the structure and tribological performance of 3D printed PLA", Journal of Materials Research and Technology, Vol. 15, pp. 647-660, 2021

[11] Liu Z., Wang Y., Wu B., Cui C., Guo Y., Yan C., "A critical review of fused deposition modeling 3D printing technology in manufacturing polylactic acid parts", The International Journal of Advanced Manufacturing Technology, Vol. 102, No. 9-12, pp. 2877-2889, 2019

[12] Ngo T. D., Kashani A., Imbalzano G., Nguyen K. T. Q., Hui D., "Additive manufacturing (3D printing): A review of materials, methods, applications and challenges", Composites Part B: Engineering, Vol. 143, No. December 2017, pp. 172-196, 2018

[13] Dobos J., Hanon M. M., Oldal I., "Effect of infill density and pattern on the specific load capacity of FDM 3D-printed PLA multi-layer sandwich", Journal of Polymer Engineering, Vol. 42, No. 2, pp. 118-128, 2022

[14] Askarinejad S., Rahbar N., "Toughening mechanisms in bioinspired multilayered materials", Journal of The Royal Society Interface, Vol. 12, No. 102, pp. 20140855, 2015

[15] Sugiyama K., Matsuzaki R., Ueda M., Todoroki A., Hirano Y., "3D printing of composite sandwich structures using continuous carbon fiber and fiber tension", Composites Part A: Applied Science and Manufacturing, Vol. 113, pp. 114-121, 2018

[16] Yazdani Sarvestani H., Akbarzadeh A. H., Niknam H., Hermenean K., "3D printed architected polymeric sandwich panels: Energy absorption and structural performance", Composite Structures, Vol. 200, pp. 886-909, 2018

[17] Askarinejad S., Choshali H. A., Flavin C., Rahbar N., "Effects of tablet waviness on the mechanical response of architected multilayered materials: Modeling and experiment", Composite Structures, Vol. 195, pp. 118-125, 2018

[18] Abueidda D. W., Elhebeary M., Shiang C.-S. (Andrew), Pang S., Abu Al-Rub R. K., Jasiuk I. M., "Mechanical properties of 3D printed polymeric Gyroid cellular structures: Experimental and finite element study", Materials & Design, Vol. 165, pp. 107597, 2019

[19] Rajpal R., K. P L., Gangadharan K. V., "Parametric studies on bending stiffness and damping ratio of Sandwich structures", Additive Manufacturing, Vol. 22, pp. 583-591, 2018

[20] Wang X., Jiang M., Zhou Z., Gou J., Hui D., “3D printing of polymer matrix composites: A review and prospective”, Composites Part B: Engineering, Vol. 110, pp. 442-458, 2017

[21] Stepashkin A. A., Chukov D. I., Senatov F. S., Salimon A. I., Korsunsky A. M., Kaloshkin S. D., “3D-printed PEEK-carbon fiber (CF) composites: Structure and thermal properties”, Composites Science and Technology, Vol. 164, pp. 319-326, 2018

[22] Liu W., Wu N., Pochiraju K., “Shape recovery characteristics of SiC/C/PLA composite filaments and 3D printed parts”, Composites Part A: Applied Science and Manufacturing, Vol. 108, pp. 1-11, 2018

[23] Brischetto S., Ferro C. G., Torre R., Maggiore P., “3D FDM production and mechanical behavior of polymeric sandwich specimens embedding classical and honeycomb cores”, Curved and Layered Structures, Vol. 5, No. 1, pp. 80-94, 2018

[24] Saad N. A., Sabah A., “An investigation of new design of light weight structure of (ABS/PLA) by using of three dimensions printing”, Proceedings of the 13th International Conference “Standardization, Prototypes and Quality: A Means of Balkan Countries’ Collaboration”, Brasov, Romania. pp. 3-4 (2016)

[25] Zhang P., Arceneaux D. J., Khattab A., “Mechanical properties of 3D printed polycaprolactone honeycomb structure”, Journal of Applied Polymer Science, Vol. 135, No. 12, pp. 46018, 2018

[26] Arrieta M., Díez García A., López D., Fiori S., Peponi L., "Antioxidant Bilayers Based on PHBV and Plasticized Electrospun PLA-PHB Fibers Encapsulating Catechin", Nanomaterials, Vol. 9, No. 3, pp. 346, 2019

[27] Oladapo B. I., Daniyan I. A., Ikumapayi O. M., Malachi O. B., Malachi I. O., "Microanalysis of hybrid characterization of PLA/cHA polymer scaffolds for bone regeneration", Polymer Testing, Vol. 83, pp. 106341, 2020

[28] Azzouz L., Chen Y., Zarrelli M., Pearce J. M., Mitchell L., Ren G., Grasso M., "Mechanical properties of 3-D printed truss-like lattice biopolymer non-stochastic structures for sandwich panels with natural fibre composite skins", Composite Structures, Vol. 213, pp. 220-230, 2019

[29] Nickels L., "Car with a biodegradable core", Reinforced Plastics, Vol. 61, No. 6, pp. 332-334, 2017

[30] Ficzere P., Borbás L., "Experimental dynamical analysis of specimens’ material properties manufactured by additive technologies", Materials Today: Proceedings, Vol. 12, pp. 352-357, 2019

[31] Lubombo C., Huneault M. A., “Effect of infill patterns on the mechanical performance of lightweight 3D-printed cellular PLA parts”, Materials Today Communications, Vol. 17, pp. 214-228, 2018

[32] Zhao H., Li L., Ding S., Liu C., Ai J., “Effect of porous structure and pore size on mechanical strength of 3D-printed comby scaffolds”, Materials Letters, Vol. 223, pp. 21-24, 2018

[33] Ficzere P., Lukacs N. L., Borbas L., "The Investigation of Interlaminar Failures Caused by Production Parameters in Case of Additive Manufactured Polymers", Polymers, Vol. 13, No. 4, pp. 556, 2021

[34] International Organization for Standardization, ISO 527-1:2012 - Plastics -- Determination of tensile properties -- Part 1: General principles (2012)

[35] International Organization for Standardization, ISO 527-2:2012: Plastics - Determination of tensile properties - Part 2: Test conditions for moulding and extrusion plastics (2012)

[36] Standard A., “ASTM C393-06. Standard test method for core shear properties of sandwich constructions by beam flexure”, ASTM international, pp. 1-7, 2006

[37] Zaharia S. M., Pop M. A., Udroiu R., “Reliability and Lifetime Assessment of Glider Wing’s Composite Spar through Accelerated Fatigue Life Testing”, Materials, Vol. 13, No. 10, pp. 2310, 2020

[38] Zaharia S. M., Enescu L. A., Pop M. A., “Mechanical Performances of Lightweight Sandwich Structures Produced by Material Extrusion-Based Additive Manufacturing”, Polymers, Vol. 12, No. 8, pp. 1740, 2020

[39] Yu J., Wang E., Li J., Zheng Z., “Static and low-velocity impact behavior of sandwich beams with closed-cell aluminum-foam core in three-point bending”, International Journal of Impact Engineering, Vol. 35, No. 8, pp. 885-894, 2008

[40] Samali B., Nemati S., Sharafi P., Tahmoorian F., Sanati F., “Structural Performance of Polyurethane Foam-Filled Building Composite Panels: A State-Of-The-Art”, Journal of Composites Science, Vol. 3, No. 2, pp. 40, 2019

Computer Vision-based Fire Detection using Enhanced Chromatic Segmentation and Optical Flow Model

Daniyar Kaliyev[1], Olga Shvets[1], György Györök[2]

[1]D. Serikbayev East Kazakhstan Technical University
Faculty of Information Technology and Intelligent Systems
A. K. Protazanov Str. 69, 070004 Ust-Kamenogorsk, Kazakhstan
e-mail: {DKaliyev, OShvets}@ektu.kz

[2]Óbuda University, Alba Regia Technical Faculty, Budai út 45, H-8000 Székesfehérvár, Hungary, e-mail: gyorok.gyorgy@amk.uni-obuda.hu

Abstract: Forests are one of the most important natural resources in the world. However, the occurrence of forest fires will burn plants and kill animals. Emergency incidents and events of fires can be dangerous and require quick and accurate decision-making. The use of computer vision for fire detection can provide an efficient solution to deal with these situations. We propose a combined method for detecting fire from a video sequence in monitoring and early fire detection operations. The method is based on motion detection methods, chromatic analysis and image segmentation. To improve the efficiency of the system, image pre-processing algorithms are proposed, and optical flow methods are used to detect the motion in fire video frames. We calculate the growth rate of the fire to reduce false-alarms. The proposed method has been tested on a very large dataset of fire videos captured by drones. It is assumed that the algorithm program is run on a computer that receives data from the camera of the drone that scans the required area. Experimental results demonstrate the effectiveness of our method while keeping their precision compatible with the existing methods.

Keywords: early fire detection; computer vision; image processing

1 Introduction

Forest fires are a powerful natural and anthropogenic factor that significantly changes the functioning and condition of forests. In countries where forests cover a large area, like Kazakhstan, forest fires are a national problem. Although Kazakhstan is a sparsely forested country, the area of the forest fund (the totality of all forests and land for forestry) is 30.4 million hectares, of which 13.3 million hectares are forested (11% of the country's territory). Up to 1000 forest fires are registered annually, covering areas up to 100 thousand hectares. On average, the amount of damage from forest fires per year is about 2.1 billion Kazak Tenge. Obviously, for many countries of the world with significant forest resources, including Kazakhstan, it is important to detect a fire as early as possible, determine its exact location and eliminate it. Therefore, it is very important to create a reliable and intelligent system for monitoring and early fire detection. Traditional fire detection systems include ground towers, aviation, and methods using satellite data. However, many such systems have a number of limitations in terms of cost, time, and space. For example, the use of satellite images makes it possible to detect fires on a continental scale, but they have large time constraints and are not applicable for early fire detection. At the initial stage of ignition, the fire is often not visible behind the barriers, and the smoke clouds often cool down too quickly to be detected. Therefore, it is necessary to use modern image processing tools that can significantly improve the efficiency of many existing systems. To overcome the limitations of traditional methods, unmanned aerial vehicle (UAV) technologies can extend traditional systems. For early forest fire detection, robotic devices such as UAVs can be used as surveillance and exploration devices. For example, one of the representatives of UAVs is quadrocopters. They are multifunctional devices that combine the automatic execution of a flight program from takeoff to landing, the transfer of the necessary information for fire detection, and the use of image data of controlled areas for subsequent computer data processing. They can be used to monitor large hard-to-reach areas such as woodlands, peat bogs, industrial sites, and more.

There are several ways to detect a fire based on information obtained in the form of images and video sequences. For example, an infrared thermal imaging camera, which is based on the detection of a stream of thermal radiation from a fire. To detect and localize wildfires, the system automatically locates hotspots using geo-referenced thermal images. Although the use of an infrared camera gives good results for fire detection, it is not recommended to use them on a drone for large scale use in fire detection systems due to their large size, weight and high cost. In recent years, there has been significant progress in digital data processing technology. At the same time, we have seen significant improvements in image and video processing capabilities, leading to greater use of computer vision systems for fire detection. The visual detection of fire in open spaces is very important for the early detection of fire. In most cases, the appearance and growth of a flame f fire can be considered a sign of an incipient fire, which, as a rule, appears as a bright spot of an indefinite changing shape against a relatively uniform darker image. To detect a fire, it is necessary to have knowledge of fire features (motion parameters, color characteristics, flickering effects) in order to correctly extract data from aerial photographs.

2 Literature Review

Fire detection systems based on machine vision typically extract three different fire data and use it for detection: color, motion, and geometry. The flame is the visible part of the fire and is made up of hot gases. The color of the flame is created by him as a result of burning materials. Depending on the substance and impurities outside, the color of the flame and the intensity of the fire will be different. However, in the case of forest fires, fire has a well-known red-yellow color. Many natural objects have colors similar to fire and can often be misidentified as flames. For this reason, it is very important to distinguish between such false alarms and real fires. Color-based detection methods are the simplest and most used methods for solving detection problems. These methods set the threshold values of the fire pixels after they have been converted to a specific color space. For example, the combination of RGB color space channels with a saturation component from the HSV (Hue, Saturation, Value) color space has been shown to be effective for extracting fire and smoke pixels [1]. The YCbCr color space was used to create a general chromaticity model for flame pixel classification [2]. The YUV color model has also been used for real-time fire detection based on the temporal variation of fire intensity due to the efficient separation of luminance from chrominance compared to the RGB color space [3]. However, the performance of color fire detection methods has been limited by the difficulty of characterizing the smoke, which in most cases strays into clouds. This problem was solved by attentional feature map in capsule networks [4]. Therefore, methods based on a combination of both color and the movement of fire or smoke increase the reliability of fire detection both indoors and outdoors [5]. For example, in [6], a combination of various flame characteristics (color, shape, and motion of the flame) obtained from video obtained by cameras was presented to reduce the number of false alarms caused by a fire. Khan et al. [7] proposed a video-based method using flame dynamics and static flame detection in a room using color, shape, and flame area. Their method may run into problems in early fire detection due to the steps of removing and then applying fire propagation characteristics. Seebamrungsat et al. [8] proposed a method based on the combination of HSV and YCbCr color spaces. Their system has additional color space conversions and is, therefore, better than methods using only one color space. But in their work, only the static characteristics of the fire are used and, therefore, the system is relatively weak and not stable enough. Chen et al. [9] improved the traditional fire detection methods by applying a flame flicker detection algorithm that was included in the algorithm for detecting fires in color video sequences. The test results show the good efficiency and reliability of the proposed algorithms. However, the calculation speed is quite slow, and is suitable for images with a low resolution, and may not be suitable for high-quality images. Kim et al. [10] used an ultra-spectral camera to distinguish flames from common light sources, which is one of the main limitations of ordinary cameras. According to the experimental results, they have achieved good results, but there may be limitations such as the higher cost of the camera. Y. Wang et al. [11] used a new video flame detection based on multi-feature fusion and double-layer XGBoost to increase the accuracy and robustness of flame recognition. The suggested technique enhances the detection rate and may be used in a variety of settings, according to the findings of experiments. Another method commonly used to detect temporal changes in flame boundaries is the wavelet transform [12, 13]. However, to gain satisfactory results in wavelet analysis, the frame rate of the input video sequence must be high, which limits its application. Deep learning has also become an active topic nowadays due to its high recognition accuracy in a wide range of applications. In the studies [14, 15, 16], a deep learning method for fire detection is used and high accuracy is achieved. We also conducted a study on the application of convolutional neural networks for fire detection in [17]. But there are certain limitations when using deep learning technology. For example, deep learning demonstrates better accuracy when working with big data, but it demands collecting a huge amount of actual fire samples taken with the UAV camera [18]. Also, deep learning requires more productive hardware for training, which takes more time.

Among the most recent studies, can be noted the work of Dang Nhu Dinh et al. [19], where they proposed combined fire detection techniques in RGB and YCbCr color spaces based on fire properties. They also propose using the correlation coefficient between consecutive frames to eliminate objects with fire-like color and diminish the vehicle shaking effect to ensure the accuracy of the proposed method in UAVs. P. Huang et al. [20] proposed a combined real-time intelligent fire detection and forecasting approach through cameras based on the computer vision method for practical application in high-fire-risk industries. V. L. Kasyap et al. [21] proposed early detection of forest fire using mixed learning techniques based on UAV composed of YOLOv4 tiny and LiDAR techniques. According to their data, the proposed model outperforms the traditional methods such as Bayesian classifiers, random forest, and support vector machines. But as noted above, there is still not enough data to create a dataset of actual fire samples taken with the UAV camera.

According to the information, which is discussed above, video-based fire detection has been rapidly explored and developed with current technology [22], but most existing fire detection methods need to be improved. In our previous work, we proposed a UAV-based system for monitoring forest fires, so in this research, the goal of our study is to optimize the existing methods and algorithms for analyzing images obtained from the UAV to monitor and detect a forest fire [23]. In this paper, we used the optical flow method for motion detection to isolate potential fire areas from other moving objects in the video frame. Optical flow is an important technique in motion analysis for computer vision-based systems. The proposed method uses fewer parameters than previous fire detection methods, making the process of fine-tuning automatic detection more intuitive.

3 Methods

The general structure of the proposed fire detection algorithm is shown in Fig. 1. The proposed model uses a combination of methods and algorithms for image processing using computer vision. The proposed algorithm can be divided into four steps. In the first stage, the input data is preprocessed for correct detection. Next comes flame region detection using the HSV color space. In the third stage, motion information is modeled by analyzing the temporal and spatial distributions via Optimal Mass Transport optical flow vectors. The last step of the algorithm is measuring the area of the regions extracted in the previous step.

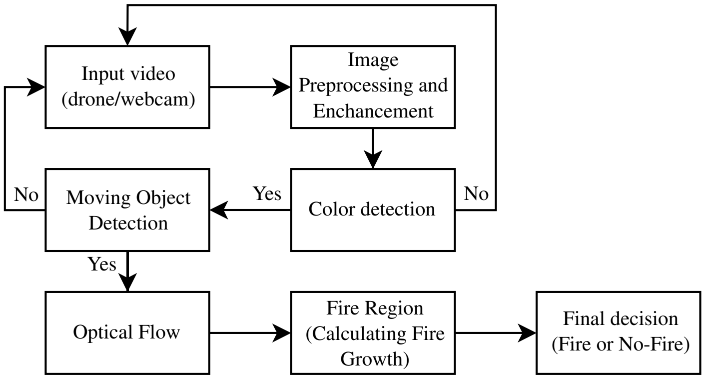


Figure 1
Workflow of the proposed model displaying the different phases

3.1 Image Pre-Processing

Image pre-processing is the stage of image enhancement and adjusting a digital image by changing its brightness, and contrast, removing noise and sharpening the image, and so on. The main goal of this stage is to bring out the details of the image so that the image is more suitable for processing and analysis. Our pre-processing step performs image filtering, contrast enhancement, and color processing. This project used the RGB and HSV color space. To detect a fire, the pixel of the fire area in RGB is explored. Then, for the HSV color space, the RGB image must be converted to HSV before the image can be analyzed. Image preprocessing will increase the reliability of data processing and recognition of the area of interest.

3.1.1 Image Smoothing Filtering

In the process of image formation, transmission, and reception, environmental noises, such as weather and light conditions, inevitably appear on images.

Therefore, before a fire is detected, the image should be smoothed and filtered. To filter and smooth a large number of images, this article applies Gaussian filtering to reduce image noise. Gaussian filtering is a classic way of smoothing images, in which image smoothing occurs by averaging pixels in the neighborhood, in which pixels at different positions are assigned different weights [24].

3.1.2 Contrast Enhancement

One of the main signs of the beginning of a forest fire is the appearance of a smoke cloud with an intensity exceeding the background intensity. Contrast enhancement is a step that is used to remove the effects of contrast variations from images due to varying light conditions. The contrast-enhancing technique is defined in the equations below [25]:

$$g(x,y)=\begin{cases} a_1 f(x,y), & f(x,y)<r_1 \\ a_2\left(f(x,y)-r_1\right)+s_1, & r_1 \geq f(x,y)<r_2 \\ a_3\left(f(x,y)-r_2\right)+s_2, & f(x,y)\geq r_2 \end{cases} \tag{1}$$

$$a_1=\frac{s_1}{r_1},\quad a_2=\frac{\left(s_2-s_1\right)}{\left(r_2-r_1\right)},\quad a_3=\frac{\left(L-s_2\right)}{\left(L-r_2\right)}. \tag{2}$$

where $g(x,y)$ presents the output image, $f(x,y)$ is the input pixel value; s_1, s_2, r_1, and r_2 are the contrast adjusting parameters; a_1, a_2, and a_3 are scaling factors for various grayscale regions and L is the maximum gray level value.

3.1.3 RGB to HSV Conversion

The conversion starts by obtaining red (R), green (G), and blue (B) values on a scale of 0 to 1 inclusive, and the largest and smallest of the R, G, B values with the difference between them. The variable $scale_x$ represents the channel scale, such as 255.

$$R=\frac{R'}{scale_r},\ G=\frac{G'}{scale_g},\ B=\frac{B'}{scale_b} \tag{3}$$

$$m_{\max}=\max\left(R,G,B\right) \tag{4}$$

$$m_{\min}=\min\left(R,G,B\right) \tag{5}$$

$$\Delta=m_{\max}-m_{\min} \tag{6}$$

To get the H (hue) values, we look at the largest of the R, G, B values. The two smallest values are subtracted and divided by the difference between the largest and smallest. Then, we normalize the hue by adding 0, 2, or 4. The resulting H is any real number, and the HSI to RGB conversion algorithm can work with any H value.

$$H = \begin{cases} undefined, & if\ \Delta = 0 \\ \dfrac{G-B}{\Delta}, & if\ m_{\max} = R \\ \dfrac{B-R}{\Delta} + 2, & if\ m_{\max} = G \\ \dfrac{R-G}{\Delta} + 4, & if\ m_{\max} = B \end{cases} \tag{7}$$

$$H' = H \times scale_h \tag{8}$$

The saturation value S is the difference between the largest and smallest values of the color channel divided by the brightness. If V is 0, then the resulting saturation is 0.

$$S = \begin{cases} 0, & if\ V = 0 \\ \dfrac{\Delta}{V}, & otherwise \end{cases} \tag{9}$$

$$S' = S \times scale_s \tag{10}$$

The value V is based on the brightest color channel.

$$V = m_{\max} \tag{11}$$

$$V' = V \times scale_v \tag{12}$$

3.2 Color Segmentation using Multi-Color Space

Color is one of the static characteristics of fire and is the most important phenomenon for determining the flame of fire. To correctly detect a fire, it is necessary to know the color range of the fire. The color of the fire changes relatively little within a particular area and changes greatly when moving from one area to another. Most color-based fire detection methods use the RGB color space because almost all visible cameras have sensors that detect video in RGB format. In the RGB color model, each color appears in its primary spectral components: red, green, and blue. According to the RGB color histogram of the fire image, the value of the red channel of the flame is greater than the value of the green channel, and the value of the green channel is greater than the blue channel. Thus, the RGB model is more suitable for color, but not very suitable for describing the image as interpreted by a human. Therefore, the main disadvantage of the RGB color space is brightness dependence, which means if the brightness of the image changes, the fire pixel segmentation rules do not work properly. Also, it is not possible to separate a pixel value into intensity and chrominance components. Thus, there is a need to convert the RGB color space to another one where the separation between brightness and chrominance is large. The HSV color space is great for this. HSV is a cylindrical color model that converts RGB primaries to a color value that is easier for humans to understand, and those parameters are hue, saturation, and value.

3.2.1 RGB Color Space Segmentation

The RGB additive color model uses light to represent color. In this model, three colors – red, green, and blue – are mixed in various combinations to obtain different colors. Thus, by varying the amount of red, green, and blue, different colors can be created. The intensity of each of the red, green, and blue components are represented on a scale from 0 to 255, with 0 being the least intensity to the maximum intensity of 255. By studying the visual features of fire, it can be observed that fires have unique visual features. These features can be divided into static and dynamic characteristics, which are both used in the classification of fire pixels. In terms of static characteristics, fire displays a distinctive range of red-yellow colors depending on its temperature and this property makes color segmentation one of the most fundamental steps. The later decisions based on shapes and other features depend on how accurately true fire pixels are segmented out. Fire samples with different flame colors, environmental conditions, and backgrounds are collected and initially checked with the following rules:

$$\begin{gathered} R > R_T \\ R \geq G \geq B \\ \left(S \geq \left((255 - R) * \frac{S_T}{R_T} \right) \right) \end{gathered} \tag{13}$$

3.2.2 HSV Color Space Segmentation

Among various color systems, HSV color model is very suitable for human interpretation, since hue, saturation, and value components are intimately related to the way in which human beings perceive color. On the basis of human interpretations of fire features, it is reasonable to assume that the flame color belongs to certain ranges of HSV components. In order to obtain the threshold values to separate flame colors, the histograms shown in Figure 3 were generated for fire pixels of HSV components.

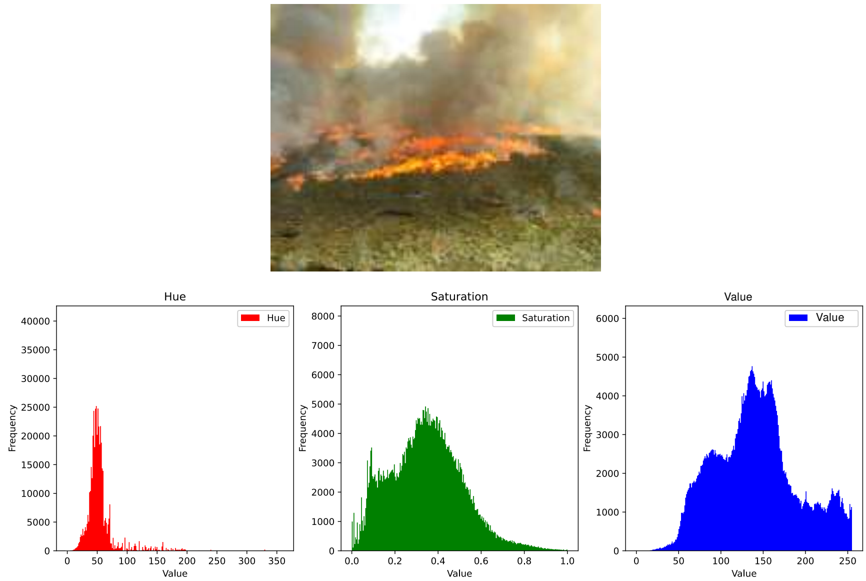


Figure 3
The input image and fire pixel histogram in HSV space

It can be concluded from Figure 3 that the fire pixels usually take much lower H values, but relatively higher S and V values. After a lot of experiments, the following decision rules to segment fire by HSV color space were defined. For flame detection, thresholds are used on each channel of the HSV color space for segmentation. For a brighter environment, the following rules are used:

$$\left.\begin{aligned} 0^\circ \le I_H(x) \le 60^\circ \\ 32 \le I_S(x) \le 100 \\ 127 \le I_V(x) \le 255 \end{aligned}\right\} \tag{14}$$

For a darker environment of image:

$$\left.\begin{aligned} 0^\circ \le I_H(x) \le 60^\circ \\ 20 \le I_S(x) \le 100 \\ 126 \le I_V(x) \le 255 \end{aligned}\right\} \tag{15}$$

3.3 Motion Detection

The proposed color models of flames detect fire regions well. However, various fire-like objects cause false positive alarms in real applications. Fires have dynamic features with a changeable shape as the airflow created by the wind can cause fluctuations and sudden movement of the fire. Some early studies consider fire-like moving objects as fire, but this method causes many false alarms, because fire-colored moving objects may all be falsely detected as fire. Red moving objects are subject to being considered flames. To reduce these false positive fire detection errors, it is reasonable to set the dynamic characteristics of the flame. To determine if the movement is caused by fire or a moving non-fire object, further analysis of the moving areas in the video sequence is needed. In this paper, we propose motion models for flames based on optical flow analysis. Optical flow is described as the two-dimensional distribution of visible motion velocities of brightness patterns in an image plane. This feature can be applied to estimate the local image pixel's movement and specify the velocity of each image pixel between adjacent images. Each pixel in the image corresponds to one velocity vector, and these velocity vectors compose an optical flow field.

3.3.1 Optimal Mass Transport Optical Flow

Classical optical flow models cannot represent the appearance of a fire because they depend on constant brightness. This problem is caused by rapid variation of intensity that exists in the combustion process and the chaotic motion produced by air turbulence. Taking into account the above conditions, the Optical Mass Transport (OMT) optical flow model is a suitable option for fire detection applications. In this work we apply OMT optical flow model due to its advantages in performing motion detection tasks with further dynamic analysis of moving areas, so that non-fire moving objects can be excluded. If the image capture camera is installed on the UAV that is in motion throughout the mission, the fire detection performance will be reduced because all objects in the field of view of the camera are moving. To solve this problem, this study proposes a novel approach to distinguishing the image variations caused by UAV motion and those caused by fire. The main idea of the proposed method is to estimate the discrepancies between the artificial optical flow and the OMT optical flow [26] and the elicitation of the fire pixels from the classified divergences. The structure of the proposed method is shown in Figure 4.

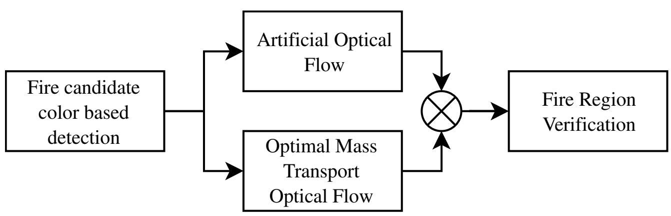


Figure 4
Basic architecture of the motion-based fire detection

3.3.2 Optical Flow Fire Feature Extraction

OMT fire motion pixels are defined as:

$$\Omega_e = \left\{ (x, y) \in \Omega : \left\| \vec{u}(x, y) \right\|_2 > c \cdot \max_{\Omega} : \left\| \vec{u} \right\|_2 \right\} \tag{16}$$

where $0 \le c < 1$ is selected so that an adequate number of pixels can be reserved, $\Omega \subset R^2$ represents an image region.

Two features can be extracted from OMT optical flow depending on dynamic texture of fire: f_1 measures the average magnitude; the directional value f_2 analyzes the motion directionality. They define the two-dimensional feature vector $F = (f_1, f_2)^T$. The given image region Ω and the OMT optical flow field in this region, the magnitude and directional characteristics are chosen through the following procedures [26]. The magnitude and directional characteristics are chosen through the following steps:

$$\text{OMT Transport Energy per pixel: } f_1 = \underset{\Omega_e}{Mean} \left(\frac{I}{2} \left\| \vec{u}_{OMT} \right\|_2^2 \right). \tag{17}$$

OMT Source Matching function is a quantification of how well the ideal source flow pattern matches the intended OMT flow field:

$$\vec{u}_T(x, y) = \begin{bmatrix} u_T(x, y) \\ v_T(x, y) \end{bmatrix} = \exp\left(-\sqrt{x^2 + y^2} \right) \begin{bmatrix} x \\ y \end{bmatrix}. \tag{18}$$

The best matching computed OMT flow field with ideal source flow can be computed by:

$$f_2 = \max_{\Omega} \left\| \left(u_T * \frac{u_{OMT}}{\left\| \vec{u}_{OMT} \right\|_2} + \left(\upsilon_T * \frac{\upsilon_{OMT}}{\left\| \vec{u}_{OMT} \right\|_2} \right) \right) \right\| \tag{19}$$

where ($*$) means convolution.

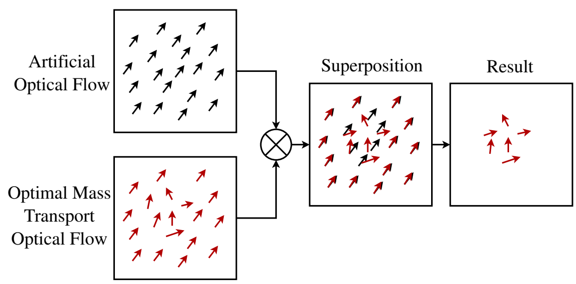


Figure 5
Illustration of the fire pixel identification

3.3.3 Motion Errors Estimation

After calculating two optical flows (an artificial optical flow for calculating camera movements and an OMT optical flow for estimating pixel movement), the next step is a comparison of these two obtained optical flows and calculating their differences. As shown in Figure 5, the artificial optical flow estimates the identical moving direction of the view in the image, while the OMT optical flow evaluates the movement of each pixel in the image. Then, these optical flows are combined into the same image, where if there is a fire in the scene, the OMT optical flow shows differences in motion between pixels.

3.3.4 Fire Region Tracking

On the basis of the estimated moving orientation of each pixel f_{OMT} utilizing the OMT optical flow and the calculated moving orientation of the camera f_α using the artificial optical flow. The difference between them can be acquired by $\Delta f = \left| f_\alpha - f_{OMT} \right|$. To reduce noise and some unexpected errors, f_α is chosen as the average value of the orientations of all pixels in the image. A simple approach to identifying fire pixels is comparing the angular deviations of each pixel with a properly chosen threshold. The decision rule for filtering the background and highlighting fire pixels is designed as follows:

$$P_{FM} = \begin{cases} 1, & if\ \Delta f > \overline{f}, \\ 0, & otherwise. \end{cases} \tag{20}$$

where P_{FM} is the binarized values of pixels obtained by applying the fire pixels motion decision making rule. If Δf outweighs the threshold $\overline{f}$, the pixel is classified as a fire pixel and is set to 1, otherwise the pixel is set to 0. The threshold value can be settled on the basis of the practical condition or by using advanced artificial intelligent methods, such as neural network or fuzzy logic.

3.4 Fire Region Growth Rate

Based on the fact that a hazardous fire grows over time, the area of flame is expected to increase. An example of the growth of the flame area during the burning of wood fuel is illustrated in Figure 6. First, we calculate the fire pixels between two consecutive frames using the following equation, which shows the fire is increasing or decreasing [27]:

$$FP_f = F\left(P(s)\right)_{i+1} - F\left(P(s)\right)_i \tag{21}$$

where F is the frame from the video and $P(s)$ is the number of fire pixels.

To identify the fire growth feature, the flame pixels in the video frames are calculated at regular time intervals and compared. To find the fire pixel difference between two frames at different time intervals, we use the following equation:

$$FP_t = F\left(P(s)\right)_{f(r)+1} - F\left(P(s)\right)_i \tag{22}$$

where $f(r)$ is the frame rate of the video and $P(s)$ is the segmented fire like pixels in the successive images. These results lead to the conclusion that the fire is likely to grow, and this increases the authentication of the real fire.

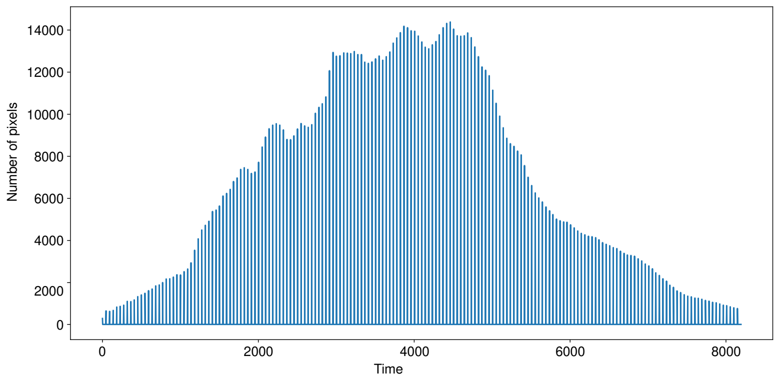


Figure 6
Graph of the amount of fire pixels extracted during the burning of wood fuel

4 Experimental Evaluation and Results

The proposed method has been tested on real cases. Tests have been run on a personal computer (NVidia GeForce RTX 2060 Super (Turing architecture) with 8 GB of graphics memory, AMD Ryzen 5 2600 processor with 3.9 GHz, 16 GB of DDR4 RAM). Our program was implemented using Python language and open-source computer vision libraries. The experimental evaluation was performed using 20 videos collected from multiple sources. Some videos in the dataset contain non-fire frames that are not visible to the eye. The video data frame rate varies from 20 to 30 fps, and the image resolution is a minimum 720 × 480 pixels. Table 1 describes the videos used in the experimental testing. Dataset contains aerial captured forest fires frames. The accuracy is calculated using the formula below:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{23}$$

We also calculated other metrics such as F1-score, precision, and recall. The F1-score is a weighted average of precision and recall, as these scores account for both false positives and false negatives. Precision is the ratio of correctly predicted positive results to the total number of positive solutions. Recall is the ratio of correctly predicted positive observations to all observations of the current class. The detection results were evaluated as follows [28]:

$$\mathrm{Precision} = \frac{TP}{TP + FP} \tag{24}$$

$$\mathrm{Recall} = \frac{TP}{TP + FN} \tag{25}$$

$$F1 = \frac{2 \times \mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \tag{26}$$

where TP - the number of true positives, TN - the number of true negatives, FP - the number of false positives and FN is the number of false negatives.

Table 1

The details of videos for experimental testing and performance evaluation

Input file	Resolution, px	Number of frames	Fire frames	Non-fire frames
Video file 1	640 x 480	278	278	0
Video file 2	640 x 480	325	321	4
Video file 3	1280 x 720	384	375	9
Video file 4	1280 x 720	292	292	0
Video file 5	640 x 480	658	657	1
Video file 6	480 x 480	364	364	0
Video file 7	720 x 480	741	697	44
Video file 8	640 x 480	365	365	0

Table 2

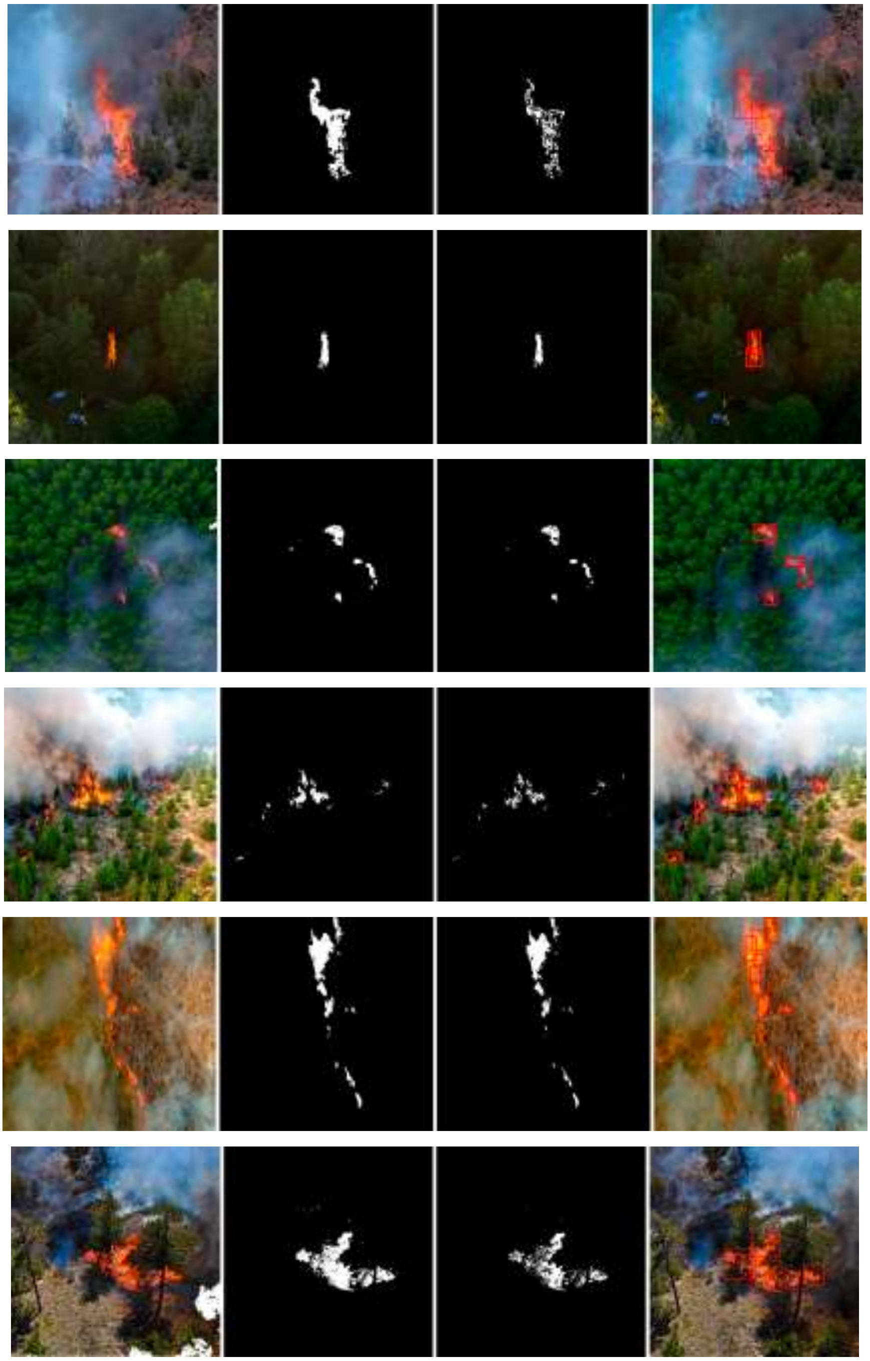
Average indicators of the proposed method

Proposed method	Accuracy, %	Precision, %	Recall, %	F1-score, %
	96.6	94.8	97.6	96.2

Table 3

Comparison of metrics in different techniques with the proposed method

Metric	D.-H. Lee et al. [1]	Khan et al. [7]	D. N. Dinh et al. [19]	**Our proposed method**
Accuracy, %	89.5	95.8	94.7	**96.6**
Precision, %	86.8	93.4	93.1	**94.8**
Recall, %	90.6	91.9	92.4	**97.6**
F1-score, %	88.7	92.6	92.7	**96.2**



(a) (b) (c) (d)

Figure 7

Experimental results of sample videos (a) Original images from video files (1-8), (b) results of color detection (c) optical flow analysis d) final results of fire tracking

Figure 7 shows examples of fire detection. Experimental results show that the accuracy of the method is 96.6%, which proves that our proposed method has higher accuracy and good performance in various scenes. From Table 2 it is clear that the precision and recall of the proposed method show the most satisfactory performance. Additionally, the performance of our proposed method with enhanced color segmentation and an optical flow analysis technique is compared with the model presented by F. Gong et al. [1], Khan et al. [7], and D. N. Dinh et al. [19]. The comparison table is given in Table 3. It can be clearly seen that the proposed method using the OMT optical flow achieved very high performance with 96.6% of accuracy respectively.

Conclusions

In this paper, we proposed a new fire detection method based on a combination of color spaces with optical flow motion detection and candidate region tracking in video captured by UAV. The key difference between the proposed method and existing studies is that our method first performs preprocessing stages followed by RGB and HSV color segmentation to detect fire-like pixels. Then, we provide segmented images to the OMT optical flow model to capture and track only the moving region in the video. Finally, we calculated the fire region growth rate in the video. Experimental results show the proposed method achieves an accuracy of around 96.6%, demonstrating the effectiveness and usefulness of the algorithm. The computational requirements of the proposed method are higher compared with those of the existing algorithms based on motion detection modeling. However, the proposed method is still considered suitable for early fire alarm systems. Future work will focus on performing the optical flow analysis step using a neural network that is trained directly on an aerial captured dataset containing fire signature vectors and creating a fuzzy logic for the detection of smoke generated by a fire in the visible image.

References

[1] F. Gong, C. Li, W. Gong, X. Li, X. Yuan, Y. Ma, T. Song: A Real-Time Fire Detection Method from Video with Multifeature Fusion. Computational Intelligence and Neuroscience, Vol. 2019, Jule 2019

[2] Wahyono, A. Harjoko, A. Dharmawan, F. D. Adhinata, G. Kosala, K.-H. Jo: Real-Time Forest Fire Detection Framework Based on Artificial Intelligence Using Color Probability Model and Motion Feature Analysis. Fire 2022, Vol. 5, No. 23, 2022, 15 p.

[3] C. Emmy Prema, S. S. Vinsley: Multi Feature Analysis of Smoke in YUV Color Space for Early Forest Fire Detection. Fire Technology, Vol. 52, Issue 5, March 2016, pp. 1319-1342

[4] S. Muksimova, S. Umirzakova, S. Mardieva, C. Young-Im: Novel Video Surveillance-Based Fire and Smoke Classification Using Attentional Feature Map in Capsule Networks. Sensors (Basel), Vol. 22, No. 1, January 2022

[5] G. Marbach, L. Markus, B. Thomas: An image processing technique for fire detection in video images. Fire Safety Journal, Vol. 41, No. 4, June 2006, pp. 285-289

[6] N. Dandge, J. Sanghavi: Forest Fire Detection in Forest Video using Image Processing. Journal of Engineering and Applied Sciences, Vol. 12, No. 8, 2017, pp. 8478-8484

[7] R. A. Khan, J. Uddin, S. Corraya, J. Kim: Machine Vision-based Indoor Fire Detection Using Static and Dynamic Features. International Journal of Control and Automation, Vol. 11, No. 6, 2018, pp. 87-98

[8] A. Khalil, S. U. Rahman, F. Alam, I.Ahmad, I. Khalil: Fire Detection Using Multi Color Space and Background Modeling. Fire Technology, Vol. 57, 2021, pp. 1221-1239

[9] J. Chen, Y. He, and J. Wang: Multi-feature fusion based fast video flame detection. Building and Environment, Vol. 45, No. 5, 2010, pp. 1113-1122

[10] H. Kim, C. Song, G.-J. Son, S.-H. Jeong, J.-H. Son, and Y.-D. Kim: Hyperspectral Image-Based Night-Time Fire Detection Using NKNBD. Proceedings of the 2018 7th International Congress on Advanced Applied Informatics, 2018, pp. 974-975

[11] Y. Wang, Y. Li, H. Wu, Y. Duan: Video Flame Detection Based on Multi-feature Fusion and Double-layer XGBoost. Engineering Letters, Vol. 30, No. 2, June 2022

[12] S. Ye, Z. Bai, H. Chen, R. Bohush, S. Ablameyko: An effective algorithm to detect both smoke and flame using color and wavelet analysis. Pattern Recognition and Image Analysis, Vol. 27, No. 1, 2017, pp. 131-138

[13] Z. Jiao, Y. Zhang, J. Xin, Y. Yi, D. Liu, H. Liu: Forest Fire Detection with Color Features and Wavelet Analysis Based on Aerial Imagery. Chinese Automation Congress (CAC), 2018, pp. 2206-2211

[14] V. Kalaivani, P. Chanthiya: A novel custom optimized convolutional neural network for a satellite image by using forest fire detection. Earth Science Informatics, Vol. 15, 2022, pp. 1285-1295

[15] P. Dai, Q. Zhang, G. Lin, M. M. Shafique, Y. Huo, R. Tu, Y. Zhang: Multi-Scale Video Flame Detection for Early Fire Warning Based on Deep Learning. Frontiers in Energy Research, Vol. 10, March 2022

[16] Z. Zhong, M. Wang, Y. Shi, W. Gao: A convolutional neural network-based flame detection method in video sequence. Signal, Image and Video Processing, Vol. 12, No. 1, 2018, pp.1619-1627

[17] D. Kaliyev, O. Shvets: Convolutional neural networks for solving fire detection problems based on aerial photograph. Program Systems: Theory and Applications, Vol. 13, No. 1, 2022 (In Russian)

[18] A. Bouguettayaa, H. Zarzourb, A. M. Taberkita, A. Kechidaa: A review on early wildfire detection from unmanned aerial vehicles using deep learning-based computer vision algorithms. Signal Processing, Vol. 190, January 2022, 14 p.

[19] N. D. Anh, P. V. Thanh, D. T. Lap, N. T. Khai, T. V. An, T. D. Tan, N. H. An, D. N. Dinh: Efficient Forest Fire Detection using Rule-Based Multi-color Space and Correlation Coefficient for Application in Unmanned Aerial Vehicles. KSII Transactions on Internet and Information Systems, Vol. 16, No. 2, February 2022, pp. 381-404

[20] P. Huang, M. Chen, K. Chen, H. Zhang, L. Yu, C. Liu: A combined real-time intelligent fire detection and forecasting approach through cameras based on computer vision method. Process Safety and Environmental Protection, Vol. 164, August 2022, pp. 629-638

[21] V. L. Kasyap, D. Sumathi, K. Alluri, P. Reddy CH, N. Thilakarathne, R. Mahammad Shaf: Early Detection of Forest Fire Using Mixed Learning Techniques and UAV. Computational Intelligence and Neuroscience Volume 2022, Article ID 3170244, 12 pages, 2022

[22] O. Shvets, D. Kaliyev, G. Györök: Automated Satellite And Uav Image Processing System For Emergency Detection In Kazakhstan. Journal of Theoretical and Applied Information Technology, Vol. 100, No. 9, 2022, pp. 2971-2987

[23] D. Kaliyev, O. Shvets, G. Györök: UAV based navigation system for forest fires monitoring. Proceedings of the AIS 2021 – 16th International Symposium on Applied Informatics and Related Areas, November 2021, pp. 47-50

[24] H. K. Xu, Y. Y. Qin, H. R. Chen: An improved edge detection algorithm based on Canny. Infrared Technology, Vol. 36, 2014, pp. 210-214

[25] J. Ahmad, B. Jan, H. Farman, W. Ahmad, A. Ullah: Disease Detection in Plum Using Convolutional Neural Network under True Field Conditions. Sensors, Vol. 20, No. 19, 2020

[26] M. Mueller, P. Karasev, I. Kolesov, A. Tannenbaum: Optical Flow Estimation for Flame Detection in Videos. IEEE Transactions on Image Processing, Vol. 22, No. 7, July 2013, pp. 2786-2797

[27] A. Khalil, S. U. Rahman, F. Alam, I. Ahmad, I. Khalil: Fire Detection Using Multi Color Space and Background Modeling. Fire Technology, Vol. 57, No. 5, 2020, pp. 1221-1239

[28] Y. Valikhujaev, A. Abdusalomov, Y. I. Cho: Automatic Fire and Smoke Detection Method for Surveillance Systems Based on Dilated CNNs. Atmosphere, Vol. 11, No. 11, 2020

Detecting the Absence of Lung Sliding in Ultrasound Videos Using 3D Convolutional Neural Networks

Michal Kolárik, Martin Sarnovský, Ján Paralič

Department of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Letná 9, 04001 Košice, Slovakia; E-mail: michal.kolarik@tuke.sk, martin.sarnovsky@tuke.sk, jan.paralic@tuke.sk

Abstract: During recent years, deep learning models proved to be very effective in multiple tasks in medicine and frequently outperformed other traditional machine learning methods. Especially in tasks where the processing of image or video data is necessary, deep networks present a very popular tool. Especially in medicine, image and video data present a frequent source of data. The work presented in this paper focuses on the use of deep learning models to detect specific phenomena from lung ultrasonography data. We focused on the detection of lung sliding, which can be observed from such data and used by clinicians in the diagnostic process of evaluating of patient's health condition. Previous research in this area mostly focused on processing a sequence of static images obtained from ultrasonography. In our work, we focused on the development of deep learning models able to process short video sequences. We used different architectures of a Resnet model and experimentally evaluated them on a real-world dataset. Then, we compared the results of best-performing architectures with the more traditional approach based on static image processing.

Keywords: deep learning; neural networks; image classification; video classification; medicine; ultrasonography

1 Introduction

Lung diseases are among the most common pathological conditions worldwide. Therefore, accurate diagnosis and monitoring of lung diseases is of general interest. Also, for the reason of the pandemic caused by the SARS-COV-2 virus, there is a great motivation to collect clinical data. This also results in increasing interest in the academic field of artificial intelligence to create intelligent solutions using machine learning [1]. Standard scanning techniques for diagnosing and monitoring lung pathologies include chest X-rays, computed tomography (CT), and magnetic resonance imaging (MRI). However, these diagnostic approaches are based on ionizing radiation. This poses a risk to the patient's health, especially in the case of a high dose or frequent exposure in the context of more frequent monitoring. Ultrasonic technologies can overcome the limitations of the mentioned technologies and thus provide a safer, portable, and cost-effective alternative [2].

Deep learning (DL) is a subfield of machine learning (ML) that attempts to learn abstractions at a higher level using hierarchical neural network architectures [3]. It is currently a very popular field, widely applied in traditional tasks of artificial intelligence (AI), such as text mining and natural language processing [4], [5], image processing, and computer vision [6]-[8], or speech recognition [9]. In medicine, DL proved to be very effective for multiple diagnostic tasks, including medical image analysis [10], [11]. With the increasing availability of sufficient computational resources, DL methods are becoming more applicable for video processing tasks as well. In comparison to images, video can provide important information about the evolution of the studied area over time. Therefore, we can track the movement of objects or the temporal appearance of an object, which can be challenging to obtain from static images. Short video sequences are also a frequent data source in medicine.

One of the sources of video data in medicine is ultrasonography. Our work focuses on the short videos obtained from the lung ultrasound. Our primary motivation is to train the deep learning model to detect the lung sliding phenomenon in the lung ultrasound data. Contrary to other published research, which uses mostly transformed static images (see Section 2), we used a 3D convolutional neural network capable of video processing. We implemented the Resnet model, experimentally evaluated multiple topologies of the model using differently preprocessed data and compared its performance to previous research.

This article is organized as follows. Section 2 provides an overview of the current state of the art in the analysis of ultrasound images using deep neural networks. The following (third) section describes the used data and models, and the fourth section summarizes the experiments and results. The final (fifth) section summarizes the main conclusions and sketches future work.

2 Related Work

When analyzing ultrasound (USG) images, the diagnostician observes certain signs (artifacts) that help to evaluate the patient's state of health. In the case of the lungs, these are mainly artifacts such as *lung sliding*, *A-lines,* and *B-lines*. Movement or *lung sliding* is a natural artifact that occurs in humans. The parietal and visceral pleura tightly adhere in a healthy individual with a minimal amount of fluid between them. This fact allows them to slide over each other during breathing. During the diagnostic process, a clinician may want to examine the patient for lung sliding (or its absence), which can be a symptom of one or more lung diseases. This backward and forward movement of the pleura can be observed on images in *B-mode* or *M-mode*.

M-mode presents images that are a 2D time-slice of a video frame in the lung region. Two patterns can be observed in these images: a *barcode* or a *seashore pattern*. These patterns are an essential feature in the classification of *M-mode* images, whether the sliding movement of the lungs occurs on the given video (*seashore pattern*) or not (*barcode pattern*) [12]. Lichtenstein et al. in work [13] classified the sliding movement of the lungs in patients arriving at the intensive care unit (ICU) with acute respiratory problems while also using the mentioned patterns in *M-mode* images. Subsequently, in combination with other monitored symptoms on ultrasound, they could diagnose lung diseases with high accuracy. In Figure 1, we can see a comparison of these two patterns and an example of the selected slice of the *B-mode* image on the *M-mode* image over time.

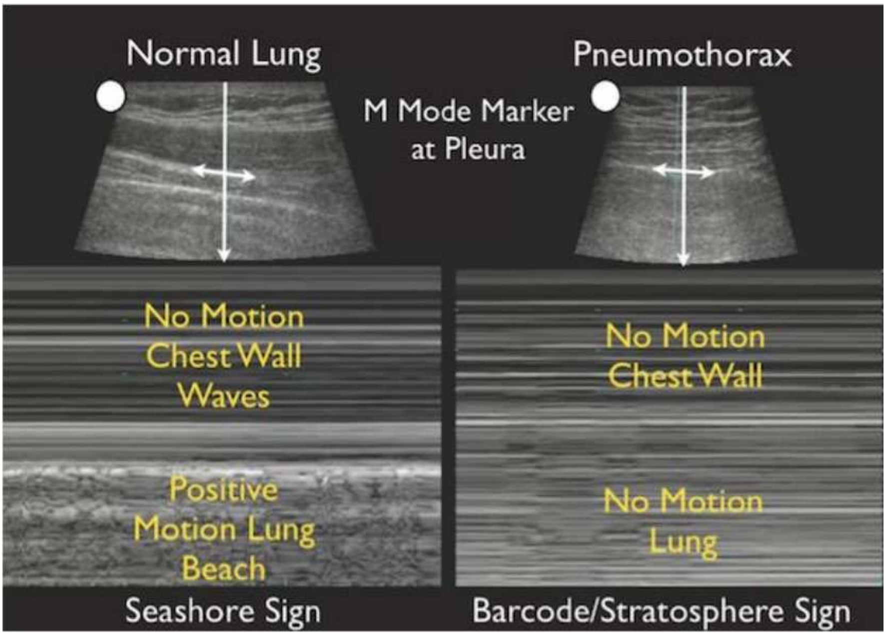


Figure 1
Example of B-mode and M-mode with seashore and barcode sign

Jaščur et al. [14] used deep learning methods for lung sliding classification on *M-mode* images. The result of their work is a system for automatic detection of *lung sliding*, which frees the doctor from the need to perform manual cuts, while with their neural network model with 64-frame architecture, they achieved 89% classification accuracy, 82% sensitivity level and 92% specificity.

Movahedi et al. [15] also devoted themselves to analyzing videos from a USG on the automatic classification of breast lesions based on videos obtained from an ultrasound examination. Their goal was to design an automated method to assist radiologists in breast screening and cancer detection. The proposed algorithm was able to reduce the vibrations that arose during the creation of USG images by the doctor in such a way that the given algorithm evaluates every tenth image, regardless of what results in the previous image were achieved. The results of the proposed algorithm achieved better results than some existing breast lesion detection methods. Following works, even if are from the different field (not specifically USG video analysis). In the paper [16], the authors devoted themselves to the classification of atherosclerotic disease based on the segmentation of the boundaries of the middle part of the vessel wall while using the Kalman snake filter, achieving better results than similar works using dynamic programming. Sonka et al. in [17] performed automatic classification of ultrasound videos with the aim of early detection of cardiovascular disease. Their goal was the analysis of brachial ultrasound images while evaluating their performance.

Ouyang et al. in [18] presented an algorithm called EchoNet-Dynamic based on echocardiography videos that outperforms human experts in the tasks of left ventricular segmentation, ejection fraction estimation and cardiomyopathy assessment. Based on repeated measurements, the paper's authors found that the model achieves comparable or smaller variance than human experts, who need years of practice for reliable and correct decision making. The given model can quickly and accurately diagnose cardiovascular diseases in real-time based on several heart cycles, is more reproducible, and quickly identifies subtle changes in the ejection fraction. This published work also includes an annotated echocardiography dataset containing 10,030 videos.

3 Methods

In the presented work, we analyse *B-mode* videos from the USG examination of the lungs to classify the presence or absence of *lung sliding*. USG examination of the lungs is a relatively new but effective lung examination technique, the main advantage of which is the absence of unnecessary transport of the patient as well as the absence of radiation to the patient, which occurs when using a conventional X-ray examination of the chest. The analysis of such images from an ultrasonographic examination needs a certain level of expertise of the radiologist operating such a system [14].

In Figure 2, we can see what a video image from an ultrasonographic examination of the lungs looks like. The green marked area represents the ribs, blue represents the lung region, and red represents the pleura.

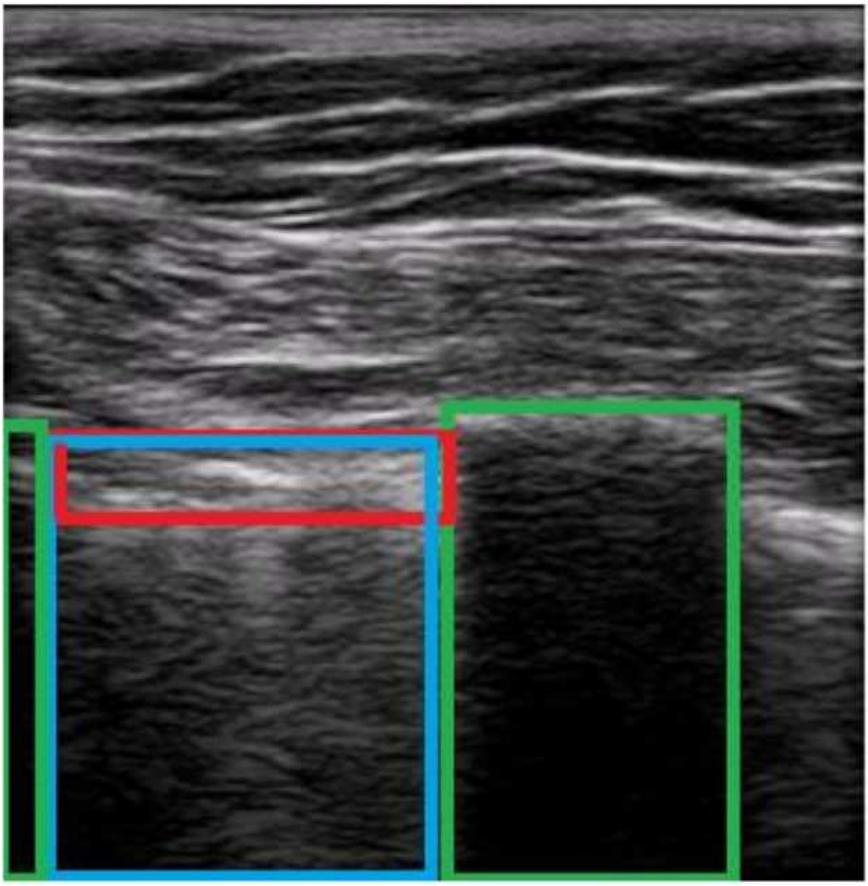


Figure 2
Image from USG video with marked areas of the lung

In this work, we focused on classifying videos containing *lung sliding* and where *lung sliding* is absent. Based on the previous results of works using *M-mode* images, we understood that for the model's decision, an image is needed in time, which creates a *Seashore* or *Barcode* image that is visible in the entire lung area. Therefore, any area of the lung should be suitable for its classification. Moreover, an image of 64 frames in a lung cross-section is already a sufficient sample. We decided to use the information about lung movement directly from the video. In contrast, for classification, we will use only a tiny area of the video in which the lungs are located, which we consider sufficient for the classification of movement or absence of lung movement. The models we will use are 3D convolutional neural networks.

3.1 Dataset Description

In our work, we used the dataset presented in [14], where the given data were collected from an ultrasound examination of the lungs by doctors from the Clinic of Thoracic Surgery of the Jessenius Faculty of Medicine in Martin, Slovakia. The obtained videos are from a pneumothorax examination in which the patient lies on his back, and the doctor scans the front and side of the chest with a linear probe and focuses on three main signs: *sliding lung*, *B-lines,* and *lung point*. All of these videos were recorded in *B-mode*, which captures the impedance of a two-dimensional tissue cross-section. Subsequently, the doctors divided this dataset into two classes, namely the class in which there are videos with the presence of lung sliding movement (28 videos) and the absence of lung sliding movement (20 videos). The dataset contains 48 videos from 48 patients in various stages of postoperative care. The given videos differ from each other in several aspects, specifically in:

- Resolution: 854 x 480 and 640 x 480
- Format: *MP4, AVI*
- Frames per second (FPS): 30 FPS and 54 FPS
- Video length:
    - Lung sliding presence (LSP): 3 – 18 seconds
    - Lung sliding absence (LSA): 2 – 11 seconds

3.1.1 Data Preprocessing

When preprocessing the dataset, we initially focused on unifying the characteristics of the videos. Since the videos contain two different video formats, namely *AVI* and *MP4*, in the first step, we processed these videos to the same type, specifically to the *MP4* format, for easier handling and further work with these videos. We then edited both types of videos to a uniform resolution of 640x480 and 30 FPS, keeping the length of the video in seconds.

In the next step, we cut the videos to a length of 2 seconds (60 frames), which should be sufficient for classifying the presence of lung movement based on the work [5]. In this way, we extended the dataset to 115 videos with the presence of lung movement and 48 videos with the absence of lung movement.

The next step of data preprocessing was to extract from the given videos the areas in which the sliding movement of the lungs occurs. This area is located between the ribs, the entire pleura width (the lung tissue between the two ribs). Since the given videos are shot in different zooms, and there is movement, it is impossible to crop the videos globally.

Based on this need, we created a function that allowed us to manually mark the area of interest in the given videos and extract from this marked area the coordinates we will need to process the dataset further. In Figure 3, we can see how the marking of the pleural areas took place.

Figure 3
The marked pleural area on the image from the USG lung video

In the selected area of interest, we further divided the video into smaller videos with page dimensions of three alternative sizes: 30x30 pixels, 60x60 pixels, and 90x90 pixels with a length of 60 frames. We created two sets. In the first set of videos, we focused on the upper part of the section only, in which the pleura is located, where the movement of the lungs is most visible. We used the first two rows of videos for this.

In the second set, we focused on the entire selected area and created videos of the entire area. In this way, we obtained six different datasets (Figure 4 visualizes the class distributions in the datasets):

- 2_30x30 (60 frames) – containing 1928 videos (1338 pos., 590 neg.)

- all_30x30 (60 frames) – containing 7 774 videos (5435 pos., 2339 neg.)
- 2_60x60 (60 frames) – containing 890 videos (620 pos., 270 neg.)
- all_60x60 (60 frames) – containing 1647 videos (1155 pos., 492 neg.)
- 2_90x90 (60frames) – containing 663 videos (230 pos., 96 neg.)
- all_90x90 (60frames) – containing 663 videos (470 pos., 193 neg.)

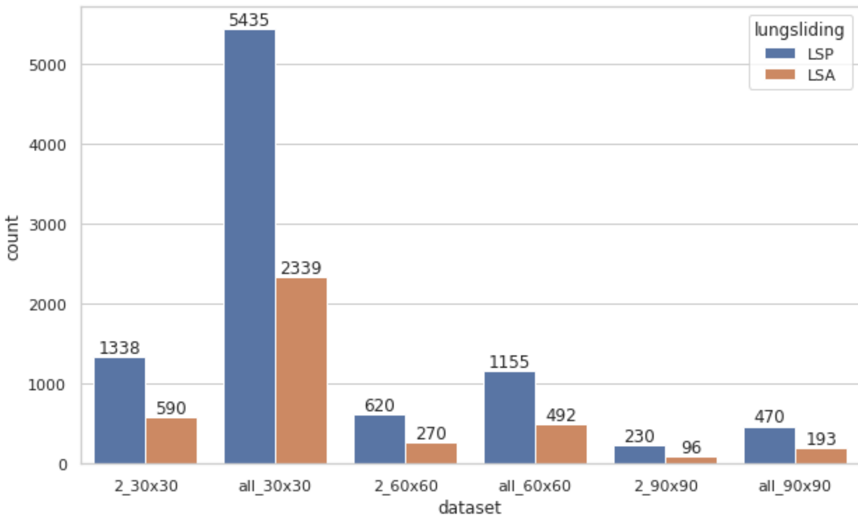


Figure 4
Class distributions in the datasets

These datasets were split into training and test sets in a ratio of 70:30, maintaining a patient-wise split to prevent videos from the same patient from appearing both in the training and test sets.

3.2 Proposed Model

Since lung sliding is a sign that appears over time, it is necessary to work with the time component of the video. One way is to use an *M-mode* image (a slice of one-pixel width and the entire video height in time) or to use a model that can process the whole video as input. In our work, we decided to investigate using the 3D CNN model directly for video processing and motion classification on video.

Considering the success of the Resnet network [19] in work [14], we decided to use Resnet 3D implementation from the freely available library keras-resnet3d[1]. This library allows the processing of 3D image data.

1 https://github.com/JihongJu/keras-resnet3d

The input dimensions of the model were *N * H * W * 3*, where *3* is the number of channels, *H* represents the video height, *W* is the video width, and *N* is the number of video frames. In our experiments, we decided to compare three architectures, *Resnet3D-18*, *Resnet3D-34,* and *Resnet3D-50*, with 18, 34, and 50 layers respectively. The input goes through a 3D convolution layer, batch normalization, and ReLU activation, followed by a max-pooling 3D layer. They are followed by a series of residual blocks with 3D convolutional layers followed by a 3D average pooling layer. The output of the pooling layer is fully connected to the output of the model. Loss is calculated as binary cross-entropy, and the output is a binary prediction of the model.

After the initial tests of the models, we chose the hyperparameters *batch_size = 32* and the *number of epochs = 16*, *optimizer = Adam* for the experiments, and used the *ModelCheckpoint* function to save the model weights while reducing the error function. Subsequently, the model with the smallest error function was used for evaluation.

4 Experimental Results and Evaluation

We used commonly used classification metrics such as accuracy, sensitivity, specificity, and F1 score to evaluate the results.

4.1 Experimental Results on Datasets 30x30

In the first experiment, we worked with a dataset with a resolution of 30x30 pixels for two rows and all rows from the lung region, respectively. We compared three Resnet3D models on each of the datasets. The results of the first experiment can be seen in Table 1. The best results were achieved by the Resnet3D-18 model on the dataset from the whole lung region (all_30x30). This model has the smallest depth among the compared models but achieved the best results. This case resulted in an accuracy of 89.13%, a sensitivity of 93.60%, and F1 score of 92.38%. Interestingly, this model performed the worst for the dataset of two lines selected around the pleura region.

Table 1

Models' performance on the 30x30 videos

Model	Dataset	Accuracy	Sensitivity	Specificity	F1
Resnet3D-18	2_30x30	0.7047	0.5985	0.9344	0.7349
Resnet3D-34	2_30x30	0.7953	0.7083	**0.9836**	0.8256
Resnet3D-50	2_30x30	0.8420	0.8788	0.7624	0.8838
Resnet3D-18	all_30x30	**0.8913**	0.9360	0.7853	**0.9238**
Resnet3D-34	all_30x30	0.8752	**0.9863**	0.6113	0.9175
Resnet3D-50	all_30x30	0.8849	0.9470	0.7376	0.9205

4.2 Experimental Results on Datasets 60x60

For the experiments on 60x60 pixel videos, we followed the same procedure as with the previous datasets. When using datasets with 60x60 pixel videos, we found that the models used on data containing only the pleura surroundings performed better. However, our results deteriorated when using the entire lung height. The *Resnet3D-34* architecture achieved the best classification results with accuracy of 85.96%, specificity of 89.85%, and F1 score of 94.34%. The best-performing architecture that used the full lung height dataset was Resnet3D-50. It achieved accuracy of 79.09%, sensitivity of 87.22%, and F1 score of 85.16%. The results of the experiments can be seen in Table 2.

Table 2

Models' performance on the 60x60 videos

Model	Dataset		Accuracy	Sensitivity	Specificity	F1
Resnet3D-18	2_60x60		0.7640	0.7143	0.8643	0.8019
Resnet3D-34	2_60x60		**0.8596**	0.8403	**0.8985**	**0.8889**
Resnet3D-50	2_60x60		0.7978	0.9076	0.5765	0.8571
Resnet3D-18	all_60x60		0.6848	0.7048	0.6407	0.7547
Resnet3D-34	all_60x60		0.6970	0.6256	0.8544	0.7396
Resnet3D-50	all_60x60		0.7909	**0.8722**	0.6118	0.8516

Compared to the best model for 30x30 pixel videos, the best model for 60x60 videos performed worse across all metrics.

4.3 Experimental Results on Datasets 90x90

This experiment ran on datasets containing 90x90 pixel videos and full lung height. The experiment confirmed that the simpler neural network architecture achieved the best results, and the deeper the architecture, the worse the classification results, as shown in Table 3. Our best model - *Resnet3D-18*, achieved classification accuracy of 88.15%, specificity of 89.40%, and F1 score of 87.68%. Overall, the models performed better than the previous experiment on the entire 60x60 dataset.

Table 3

Models' performance on the 90x90 videos

Model	Dataset	Accuracy	Sensitivity	Specificity	F1
Resnet3D-18	all_90x90	**0.8815**	0.8768	**0.8940**	**0.9150**
Resnet3D-34	all_90x90	0.8419	0.8518	0.8154	0.8870
Resnet3D-50	all_90x90	0.8267	**0.9875**	0.3960	0.8925

4.4 Experimental Results Comparison

Finally, we decided to compare the results of all our experiments. Comparing the experiments in Table 4, we can see that the best results were obtained by the *Resnet3D-18* model on the 30x30 pixel dataset from the whole lung region, by the *Resnet3D-34* on the 60x60 data of the first two rows, and by *Resnet3D-18* on the 90x90 pixel dataset from the entire lung region. Interestingly, according to the doctors, the area around the pleura is important for deciding lung movement, and the movement should be visible just at the pleura.

Table 4
Summary table of all evaluated models

Model	Dataset	Accuracy	Sensitivity	Specificity	F1
Resnet3D-18	2_30x30	0.7047	0.5985	0.9344	0.7349
Resnet3D-34	2_30x30	0.7953	0.7083	0.9836	0.8256
Resnet3D-50	2_30x30	0.8420	0.8788	0.7624	0.8838
Resnet3D-18	all_30x30	0.8913	0.9360	0.7853	0.9238
Resnet3D-34	all_30x30	0.8752	0.9863	0.6113	0.9175
Resnet3D-50	all_30x30	0.8849	0.9470	0.7376	0.9205
Resnet3D-18	2_60x60	0.7640	0.7143	0.8643	0.8019
Resnet3D-34	2_60x60	0.8596	0.8403	0.8985	0.8889
Resnet3D-50	2_60x60	0.7978	0.9076	0.5765	0.8571
Resnet3D-18	all_60x60	0.6848	0.7048	0.6407	0.7547
Resnet3D-34	all_60x60	0.6970	0.6256	0.8544	0.7396
Resnet3D-50	all_60x60	0.7909	0.8722	0.6118	0.8516
Resnet3D-18	all_90x90	0.8815	0.8768	0.8940	0.9150
Resnet3D-34	all_90x90	0.8419	0.8518	0.8154	0.8870
Resnet3D-50	all_90x90	0.8267	0.9875	0.3960	0.8925

Our experiments' results were then compared with the study by Jaščur et al. [14], which used an M-mode approach to time-lapse video images and their classification using Resnet networks. Table 5 shows the results of comparing our best models with the Resnet M-mode model for image classification.

Table 5
Comparison of the best performing model with the Resnet M-mode model (Resnet-18)

Model	Dataset	Accuracy	Sensitivity	Specificity	F1
Resnet3D-18	all_30x30	**0.9193**	**0.9360**	0.7853	0.9238
Resnet3D-34	2_60x60	0.8596	0.8403	0.8985	0.8889
Resnet3D-18	all_90x90	0.8815	0.8768	0.8940	0.9150
Resnet-18	64-frame	0.8900	0.8600	**0.9200**	-

The best-performing model achieved better accuracy and sensitivity but worse specificity. In the case of sensitivity, we achieved more than 7.5% better. Our model learned better to recognize positive examples, i.e., those for which there was a sliding lung movement in the video. The metrics are comparable to those achieved by the authors, while we had a smaller amount of data available for the training phase than the authors had for the 64-frame architecture.

4.5 Explainability

After the model training and evaluation, we wanted to explore the possibility of the model explanation. We decided to use the Vanilla saliency map method [20], which generates a heat map to show which part of the input image (in our case, which pixel in the video) affects the output classification. As a second method, we chose SmoothGrad [21], which is based on the saliency map method. It works on a similar principle, but adds noise at the input of its heat map calculations to reduce the noise at the output of this explainability. Thus, it provides more distinct regions on which to base the model's decisions. In Figure 5, we can see the applied explainability methods for a particular video over time. The given frames are six frames apart to show the given explainability for the entire video. The original video model (input) is in the first row. The second row depicts the frames from the video with Vanilla Saliency Map applied, and in the third row, SmoothGrad result is presented. In both cases, the highlighted areas visualize the pixels most significantly contributing to the classification.

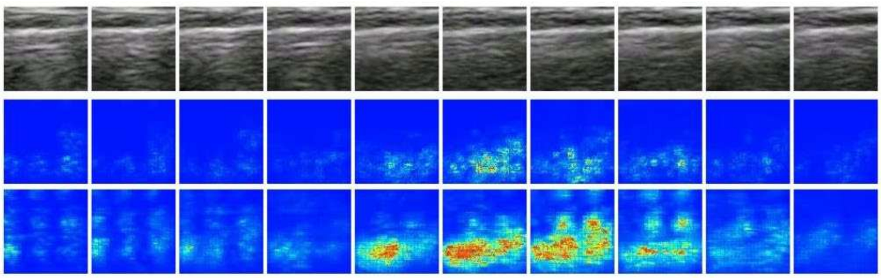


Figure 5
Example of processed frames (first row), and result of XAI methods Vanilla Saliency Map (second row) and SmoothGrad (third row) to each processed frame

Conclusion

This work's goal was to use 3D convolutional neural networks for lung motion classification from USG B-mode videos. We used only a part of the video, thereby obtaining a larger number of samples, reducing the demands on the neural network, and proving that it is possible to replace lung movement in the time domain using M-mode classification directly from the video. We also tried to explain the decisions of the network using XAI methods such as Vanilla Saliency Maps and SmoothGrad. Assuming that additional annotated data from the ultrasonographic examination of the lungs are obtained, further research on the given issue could focus directly on the pleural area and classify the movement of the lungs only from this particular area. This was not possible due to the small amount of input data.

Subsequently, the given experiments presented in the paper could be repeated with additional explainability methods, the use of other neural network architectures, or the use of transfer learning to improve the results of the used network.

Acknowledgment

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-20-0232 and contract No. APVV-17-0550 and by the Slovak VEGA research grant No.1/0685/21.

References

[1] S. B. Desai, A. Pareek, and M. P. Lungren, "Deep learning and its role in COVID-19 medical imaging," *Intell. Med.*, 2020

[2] L. Demi, T. Egan, and M. Muller, "Lung ultrasound imaging, a technical review," *Applied Sciences (Switzerland)* 2020

[3] C. C. Aggarwal and others, *Neural networks and deep learning*. Springer, 2018

[4] V. Maslej-Krešňáková, M. Sarnovský, P. Butka, and K. Machová, "Comparison of Deep Learning Models and Various Text Preprocessing Techniques for the Toxic Comments Classification," *Appl. Sci.*, Vol. 10, No. 23, 2020

[5] V. M. Kresnakova, M. Sarnovsky, and P. Butka, "Deep learning methods for Fake News detection," in *IEEE Joint 19th International Symposium on Computational Intelligence and Informatics and 7th International Conference on Recent Achievements in Mechatronics, Automation, Computer Sciences and Robotics, CINTI-MACRo 2019 - Proceedings*, 2019

[6] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep Learning for Computer Vision: A Brief Review," *Computational Intelligence and Neuroscience*. 2018

[7] V. Maslej-Krešňáková, K. El Bouchefry, and P. Butka, "Morphological classification of compact and extended radio galaxies using convolutional neural networks and data augmentation techniques," *Mon. Not. R. Astron. Soc.*, 2021

[8] Š. Mackovjak, M. Harman, V. Maslej-Krešňáková, and P. Butka, "SCSS-Net: solar corona structures segmentation by deep learning," *Mon. Not. R. Astron. Soc.*, 2021

[9] Y. Dokuz and Z. Tufekci, "Mini-batch sample selection strategies for deep learning based speech recognition," *Appl. Acoust.*, 2021

[10] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep Learning Applications in Medical Image Analysis," *IEEE Access*, 2017

[11] M. I. Razzak, S. Naz, and A. Zaib, "Deep learning for medical image processing: Overview, challenges and the future," in *Lecture Notes in*

Computational Vision and Biomechanics, 2018

[12] S. Singh, H. Kaur, S. Singh, and I. Khawaja, "Basic Insights of Lung Ultrasonography in Critical Care Setting," *Cureus*, 2018

[13] D. A. Lichtenstein and G. A. Mezière, "Relevance of lung ultrasound in the diagnosis of acute respiratory failure the BLUE protocol," *Chest*, 2008

[14] M. Jaščur, M. Bundzel, M. Malík, A. Dzian, N. Ferenčík, and F. Babič, "Detecting the absence of lung sliding in lung ultrasounds using deep learning," *Appl. Sci.*, 2021

[15] M. M. Movahedi, A. Zamani, H. Parsaei, A. T. Golpaygani, and M. R. H. Poya, "Automated analysis of ultrasound videos for detection of breast lesions," *Middle East J. Cancer*, 2020

[16] S. Zhao, G. Li, W. Zhang, and J. Gu, "Automatical Intima-Media Border Segmentation on Ultrasound Image Sequences Using a Kalman Filter Snake," *IEEE Access*, 2018

[17] M. Sonka, W. Liang, and R. M. Lauer, "Automated analysis of brachial ultrasound image sequences: Early detection of cardiovascular disease via surrogates of endothelial function," *IEEE Trans. Med. Imaging*, 2002

[18] D. Ouyang *et al.*, "Video-based AI for beat-to-beat assessment of cardiac function," *Nature*, 2020

[19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016

[20] K. Simonyan, A. Vedaldi, and A. Zisserman, "Deep inside convolutional networks: Visualising image classification models and saliency maps," in *2nd International Conference on Learning Representations, ICLR 2014 - Workshop Track Proceedings*, 2014

[21] D. Omeiza, S. Speakman, C. Cintas, and K. Weldermariam, "Smooth Grad-CAM++: An Enhanced Inference Level Visualization Technique for Deep Convolutional Neural Network Models," Aug. 2019

Optimal Fuzzy Controller, using a Genetic Algorithm for a Ball on Wheel System

Péter Menich, József Kopják

Kandó Kálmán Faculty of Electrical Engineering, Óbuda University, Bécsi út 96/b, H-1034 Budapest, Hungary
email: peter.menich@pentasoft.hu, kopjak.jozsef@kvk.uni-obuda.hu

Abstract: The process of building a fuzzy controller for the Ball on Wheel problem raised an issue with the optimization of the controller parameters. For non-linear systems the parameter values of the controller are based on a subjective estimate and a trial-and-error approach. This is a time-consuming human task, which could be automated. The information obtained from the visualized simulation results gave us the opportunity to increase the controllable area, using a genetic algorithm (GA), by tuning the parameters of the fuzzy membership methods.

Keywords: unstable; fuzzy; genetic algorithm; microcontroller; BLDC; simulink

1 Introduction

The control of an unstable non-linear system is always a challenging task. [1] Fuzzy controllers are well suited for such controlling applications [22] e.g., inverted pendulum [14] control of an unstable bioreactor [23]. Combining the fuzzy controller with various techniques opens up the possibility to increase the performance of the controller without significantly increasing the required computational power. A list of representative examples showing the huge scope for further development of controller design: predictive control [24], data-driven hybrid controller design [25], evolving fuzzy models [26], using lookup tables [29], cascade system structure [30].

Performance can also be improved by optimizing the controller. This can be achieved in several ways, of which the following are some examples: Multitasking genetic algorithm [27], using a Slime mold algorithm [28]

Our goal was to build a very simple fuzzy controller for educational purposes for the “ball on the wheel" problem, where the task is to keep a free-rolling ball on the top of a rotating wheel. Since our goal was to make the controller as simple as possible, we discarded complex solutions and focused on optimizing the controller parameters, i.e., the points of the triangular membership functions.

To measure the quality, we constructed a parameter space that assigned an ITAE (Integral of the time-weighted absolute error) [2] value to the initial speed and position of the ball. It shows the size and shape of the controlled area. We used a genetic algorithm to expand this controlled area as much as possible. Unfortunately, the full parameter space cannot be mapped to evaluate the fitness function, so it was reduced to five points. In these five points, the ITAE values were determined and the sum of these values gave the result of the fitness function. Our approach allowed us to cover a relatively large parameter space in such a way that it can be used to evaluate the fitness function of the genetic algorithm. The visualization allowed us to investigate how the selection of the five points influences the shape of the controllable area and thus the behavior of the controller.

2 Ball on the Wheel

The ball on the wheel problem is the following: a free-rolling ball must be kept on the upper unstable equilibrium of a wheel by controlling its rotation. Ignoring the dissipation, there are an infinite number of solutions when the ball is rolling with constant speed at the upper equilibrium. Calculating with dissipation the ball can be kept also in a constant rolling state where the g acceleration keeps the balance with the dissipation. The controller is planned to keep the ball standing at the upper equilibrium. In order for the controller to be able to do this, it must be able to handle the various disturbances, whether they are the result of an external force or the fact that the ball is not in a stationary position when it is lifted to the upper equilibrium.

It is a good demonstration platform for non-linear, unstable system control [3] [4]. In [5] [6] PID controller, in [7] linear controller with trajectory planning was used to demonstrate a solution for this problem.

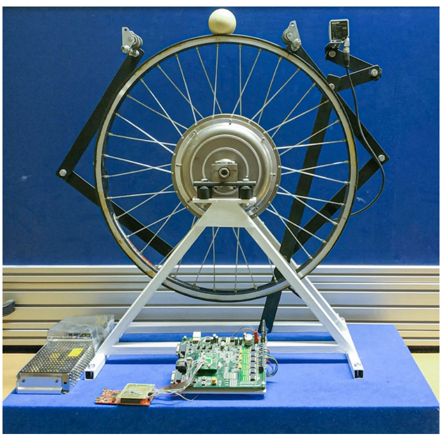


Figure 1
Ball on the wheel system

2.1 The Hardware

The hardware is based on an electric scooter wheel containing a BLDC motor with embedded HALL sensors. The driver is the MCLV development board from Microchip, the controller is an X2C development board using dsPIC33EP256MC502 microcontroller. The software uses a six-step control with PWM. The commutation is based on the inputs of the HALL sensors. The ball's position sensor is a BALLUFF BOD21M-LA04-S92 laser distance sensor. It has a 2.5 ms sample time with 4 samples moving average filter and an analog output.

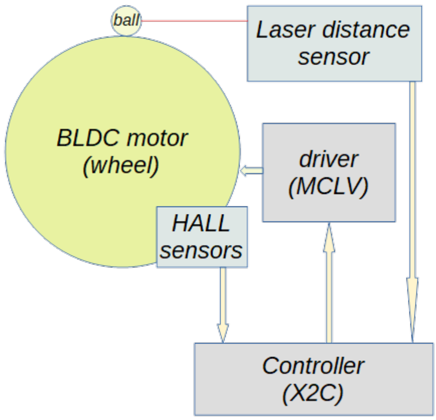


Figure 3
Schematic diagram of the hardware

2.2 The Model

We are using a fuzzy logic controller to balance the ball on the wheel, so the mathematical model building is not necessary, although it gives us the opportunity to examine the system in a virtual environment. From an educational perspective, it provides an opportunity for students to test other methods and controllers before they try them in a real environment.

2.1.1 Mathematical Model

The Lagrange formula: $L = T - V$ [9] (1)

Kinetic energy:

$$T = \frac{1}{2}\Theta_W\dot{\phi}_W^2 + \frac{1}{2}\Theta_B\left(\frac{r_W}{r_B}\left(\dot{\phi}_{pos} - \dot{\phi}_W\right)\right)^2 + \frac{1}{2}m_B\left((r_w + r_B)\dot{\phi}_{pos}\right)^2 \quad (2)$$

where φ_W is the current angle of the wheel and φ_{pos} is the ball position angle. The other parameters can be found in Table 1.

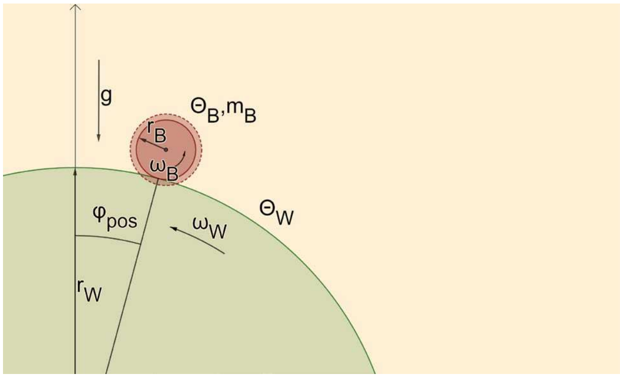


Figure 6
Schematic drawing of the Ball on the wheel

Potential energy (zero level is the center of the wheel)

$$V = -m_B g(r_B + r_w)cos(\phi_{pos}) \tag{3}$$

Generalized forces:

$$Q = \begin{bmatrix} Q_M - Q_W + Q_B \\ -Q_B \end{bmatrix} \tag{4}$$

Where,

Motor: $Q_M = \frac{K_M(U_a - K_U \dot{\phi}_W)}{R_a}$ (5)

where U_a is the armature voltage.

Wheel: $Q_w = K_{W1} \cdot sign(\dot{\phi}_W) + K_{W2} \cdot \dot{\phi}_W$ (6)

Ball:$Q_B = K_{B1} \cdot sign(\dot{\phi}_{pos} - \dot{\phi}_W) + K_{B2} \cdot \dot{\phi}_W$ (7)

The Lagrangian dynamics:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = Q \tag{8}$$

extracting (8) using (1) – (7) gives the following result:

$$\begin{bmatrix} \Theta_W + \Theta_B \frac{r_W^2}{r_B^2} & -\Theta_B \frac{r_W^2}{r_B^2} \\ -\Theta_B \frac{r_W^2}{r_B^2} & \Theta_B \frac{r_W^2}{r_B^2} + m_B(r_w + r_B)^2 \end{bmatrix} \begin{bmatrix} \ddot{\phi}_W \\ \ddot{\phi}_{pos} \end{bmatrix} - \\ - \begin{bmatrix} 0 \\ m_B g(r_B + r_w)sin(\phi_{pos}) \end{bmatrix} = \begin{bmatrix} Q_M - Q_W + Q_B \\ -Q_B \end{bmatrix} \tag{9}$$

The mathematical model contains the following parameters of the physical system:

Table 1
Exact values of the parameters

Θ_W	0.30618	kg · m^2	moment of inertia (wheel)
Θ_B	0.0000731	kg · m^2	moment of inertia (ball)
m_B	0.203	kg	weight (ball)
r_W	0.27	m	radius (wheel)
r_B	0.0254	m	radius (ball)
R_a	0.3	Ω	armature resistance
K_M	2.093	V·s/rad	motor mechanic constant
K_{W1}	0.0023	N·m	Coulomb friction constant (wheel)
K_{W2}	0.0003	N·m·s	viscous friction constant (wheel)
K_U	2.518	N·m/A	motor electric constant
K_{B1}	0.045	N·m	Coulomb friction constant (ball)
K_{B2}	0.005	N·m·s	viscous friction constant (ball)

Using the values from Table 1, in (9) it gives the following differential equations formula:

$$\begin{bmatrix} \ddot{\phi}_{pos} \\ \ddot{\phi}_W \end{bmatrix} = \begin{bmatrix} 22.84 & -17.94 & -0.17 & -0.00235 & 7.115 \\ 0.6 & -56.34 & 0.01 & -0.00738 & 22.375 \end{bmatrix} \begin{bmatrix} sin(\phi_{pos}) \\ \dot{\phi}_W \\ sign(\dot{\phi}_{pos} - \dot{\phi}_W) \\ sign(\dot{\phi}_W) \\ U_a \end{bmatrix}$$

$$A = \begin{bmatrix} 22.84 & -17.94 & -0.17 & -0.00235 & 7.115 \\ 0.6 & -56.34 & 0.01 & -0.00738 & 22.375 \end{bmatrix}$$

$$\begin{bmatrix} \ddot{\phi}_{pos} \\ \ddot{\phi}_W \end{bmatrix} = A \cdot \begin{bmatrix} sin(\phi_{pos}) \\ \dot{\phi}_W \\ sign(\dot{\phi}_{pos} - \dot{\phi}_W) \\ sign(\dot{\phi}_W) \\ U_a \end{bmatrix} \tag{11}$$

2.2.2 Modelling in Matlab/Simulink

The model can be implemented in Simulink from the linear system of differential equations using the Kelvin-Thomson theorem [8]. The highest derivatives are separated on the left side, so the integrators can be used to get the lower derivatives.

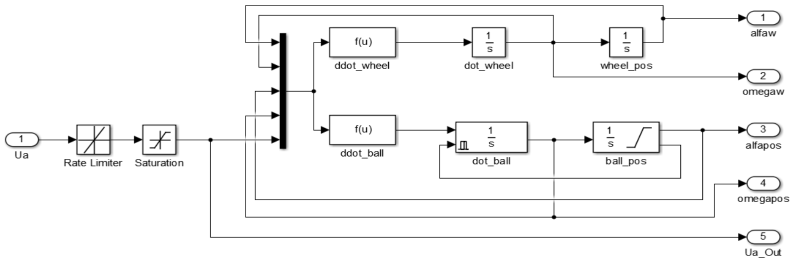


Figure 7
Simulink nonlinear model of the ball on the wheel problem

It is implemented in a subsystem. The input is the armature voltage and the outputs are all of the accessible parameters: the wheel's and the ball's position and angular velocity and the limited armature voltage. The input voltage is saturated to the nominal voltage of the hardware and the change rate is limited because the mathematical model enables fast oscillation of the input voltage, but it is not recommended to get it on the physical hardware. The ball's position is limited by bumpers, which is implemented by saturation. When the ball reaches the end position the velocity of the ball must be set to zero, so the integrator must be reset.

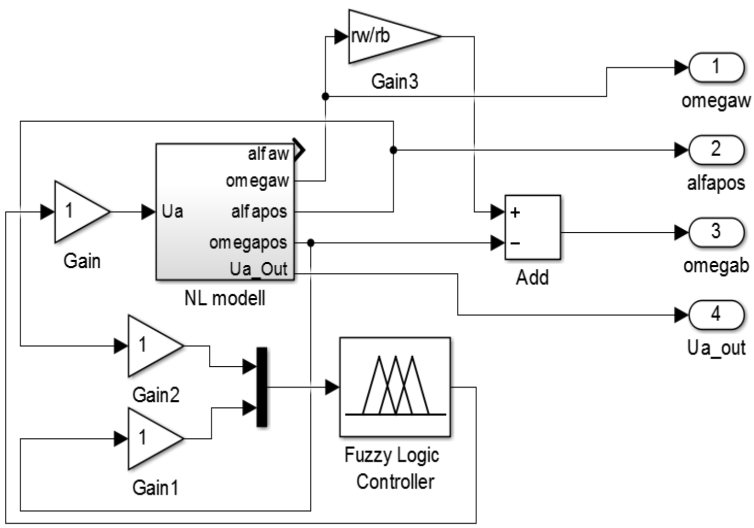


Figure 8
The ball on the wheel fuzzy controller

The current fuzzy controller configuration uses the ball position and velocity values for controlling the movement of the wheel set up the armature voltage.

Evaluating the quality of control

The different parameter setup gives very different behavior. To measure the quality of the control the ITAE (Integral of the time-weighted absolute error) function was implemented.

ITAE: $\int_0^\infty w(t) \cdot |e(t)|\, dt$ (12)

where w(t) is the time weighting and e(t) is the error function.

In our case the $w(t) = t$ (13)

and $e(t) = \alpha_{pos}(t)$ (14)

The ITAE (12) using (13) and (14) is $\int_0^\infty t \cdot \left|\alpha_{pos}(t)\right|\, dt$

The Simulink subsystem:

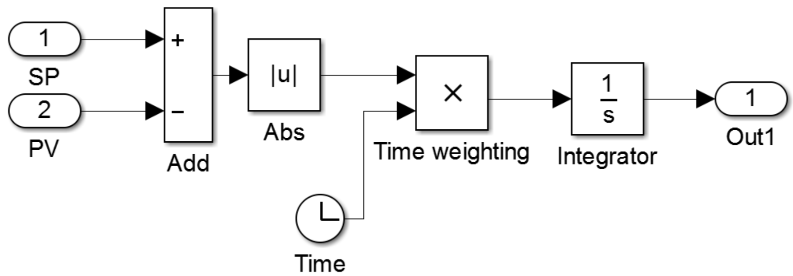


Figure 11

ITAE subsystem

The parameters are: SP – set point, PV – present value

2.2 Fuzzy Controller

The goal was to build a fuzzy controller as simple as possible, so the following restrictions were made:

- Using triangle membership functions
- Related to the symmetric problem, the fuzzyfication of the input values and the output defuzzyfication must be symmetric
- Using Mamdani method
- Using only two inputs
- Using Ruspini partitions [10]

Related to these restrictions the fuzzy controller can be defined by using 6 parameters. For each membership function (two inputs and one output) it needs two parameters. In Figure 7 the dotted lines mark the positions. The first parameter is the distance of line 'a' from center and the second is the distance line 'b' from the line 'a'.

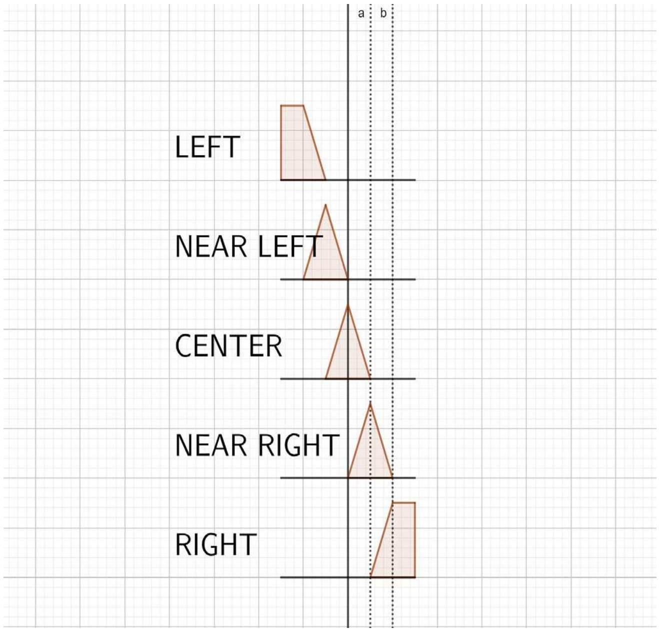


Figure 13
Fuzzy membership functions

The input 1 – ball position membership functions:

center: $\{[-a_{pos},0],[0,1],[a_{pos},0]\}$

near left, near right: $\{[-b_{pos},0],[-a_{pos},1],[0,0]\}$, $\{[0,0],[a_{pos},1],[b_{pos},0]\}$

far left, far right: $\{[min_{pos},1],[-b_{pos},1],[-a_{pos},0]\}$, $\{[a_{pos},0],[b_{pos},1],[max_{pos},1]\}$

The input 2 – ball angular velocity membership functions:

stop: $\{[-a_{vel},0],[0,1],[a_{vel},0]\}$

slow left, slow right: $\{[-b_{vel},0],[-a_{vel},1],[0,0]\}$, $\{[0,0],[a_{vel},1],[b_{vel},0]\}$

fast left, fast right: $\{[min_{vel},1],[-b_{vel},1],[-a_{vel},0]\}$, $\{[a_{vel},0],[b_{vel},1],[max_{vel},1]\}$

The output – armature voltage membership functions:

stop: $\{[-a_{out},0],[0,1],[a_{outl},0]\}$

low left, low right: $\{[-b_{out},0],[-a_{out},1],[0,0]\}$, $\{[0,0],[a_{out},1],[b_{out},0]\}$

high left, high right: $\{[min_{out},1],[-b_{out},1],[-a_{out},0]\}$, $\{[a_{out},0],[b_{out},1],[max_{out},1]\}$

The 6 parameters are: a_{pos}, b_{pos}, a_{vel}, b_{vel}, a_{out}, b_{out}

These parameters are used for both input membership functions (ball position, angular velocity of the position) and for the output membership function. [11-17]

3 Optimization

The model described in the previous chapter can be used to study the behavior of the system in a virtual environment with the given fuzzy controller parameters. We can study how the system behaves under different initial conditions. With given fuzzy parameters, we can plot the three-dimensional ITAE function image to study the controller behavior around the equilibrium.

Our goal is to design a controller that is less sensitive to disturbances, and to do this we want to tune the parameters of the fuzzy controller. The model also gives us the possibility to use it in the calculation of the fitness function of the genetic algorithm. To do this, we found that evaluating parameter space only at 5 points has a suitable effect on the shape of the controllable region. The sum of the ITAE function evaluated at the five points is used to calculate the value of the fitness function.

3.1 Genetic Algorithm

Genetic algorithms were inspired by natural evolution. The DNAs of the individuals are the current parameter values. The probability of survival and the chances of reproduction are based on the fitness function which determines the quality of the individual. [18-20]

DNA of each individual: $x = \begin{bmatrix} a_{pos} \\ b_{pos} \\ a_{vel} \\ b_{vel} \\ a_{out} \\ b_{out} \end{bmatrix}$

The genetic representation is the six parameters described above, and the fitness function is derived from the ITAE function.

The genetic algorithm feature of Matlab can handle restrictions in the following format: $A \cdot x \leq b$, where A and b can be defined, and x is the DNA's vector representation.

To have better performance the parameter space should be decreased as much as possible. The following restriction was made for the x parameter vector:

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} \cdot x \leq \begin{bmatrix} 0.5 \\ 15 \\ 24 \\ -0.01 \\ -0.01 \\ -0.01 \\ -0.01 \\ -1 \\ -1 \end{bmatrix}$$

This means that position must be in [-0.5,0.5] the angular velocity in [-15,15] range and the output must be in [-24,24] interval. The minimum distance between points must be 0,01 for input and 1 for output.

3.2 Fitness Function

The previously defined ITAE function measures the quality of the controller's response for the initial state, but our goal is to define the controller parameters to have a reasonable control near the area of the equilibrium. The following three parameters can describe the actual state of the system:

- Position of the ball
- Angular velocity of the wheel
- Angular velocity of the ball

Certainly, other linear combinations of these three parameters could be used also.

Since the goal is to have the ball standing at the equilibrium, only the first two parameters are used for the quality measurement. So, the first coordinate of the parameter space is the ball's position the second one is the speed of the wheel. The simulation starts with these parameters where the ball lateral speed is zero. In this case the initial angular velocity of the ball: $\omega_b = \frac{r_w \cdot \omega_w}{r_b}$

Unfortunately to evaluate the result of the fitness function for the whole parameter space is not possible, so only some predefined points were examined to check the size of the controllable area near the equilibrium. The redesigned simulation for the optimization has five parallel simulations and the fitness function is the summary of the ITEA values. The minimization criteria will cause the algorithm to hold the ball in the equilibrium at as many points out of five as it can. The best case is when all of them are controllable.

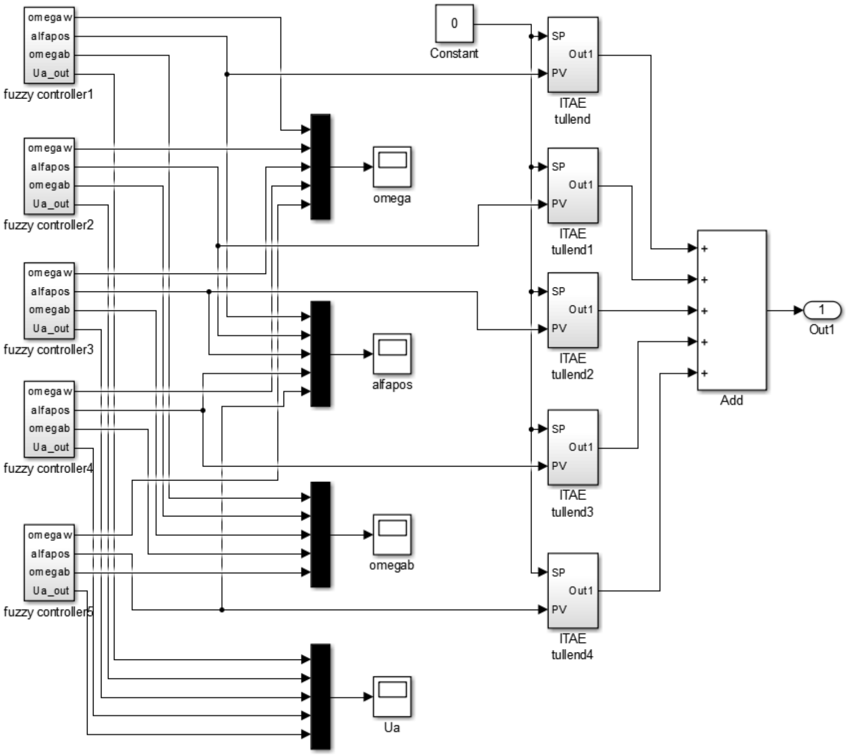


Figure 15
Five parallel simulation

3.3 Visualization

The visualization of the parameter space gives the opportunity to examine the behavior of the controller. To visualize we used the function$f: \mathbb{R}^2 \rightarrow \mathbb{R}_0^+$, where the domain is the initial position of the ball and the initial angular velocity of the wheel$\left(\alpha_{pos} \times \omega_w\right)$, the co-domain is the ITAE result.

Figure 9 shows the controllable area for a basic fuzzy membership function which was set similarly to Figure 7. The controllable area is where the value of the ITAE function is low (blue). It can be seen that the narrow path can be controlled mainly where the signed summary of the kinetic and potential energy of the system is near zero.

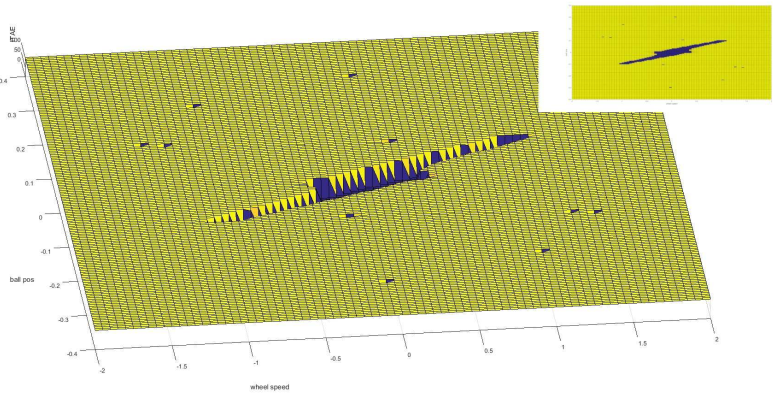


Figure 17

Controllable area visualization for simple equidistant fuzzy membership function

3.4 Expanding the Controllable Area

Unfortunately, the evaluation of the ITAE for the whole parameter space, is not possible. To get an acceptable runtime result we had to reduce the number of points to evaluate. The evaluation of five different points gives a big enough impact on the desired shape of the controllable area and provides acceptable calculation time. The summary of the ITAE values of these points gave the result of the fitness function. The implementation can be seen in Figure 8. Expanding the area on the ball position axis reduced the control on the other (Figure 10).

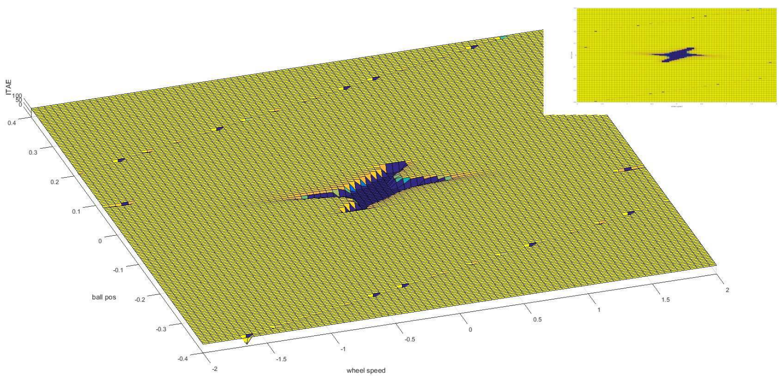


Figure 19

Expanded ball position area near to zero wheel speed

The result of the wheel speed insensitivity gives a narrow field for the ball's position (Figure 11).

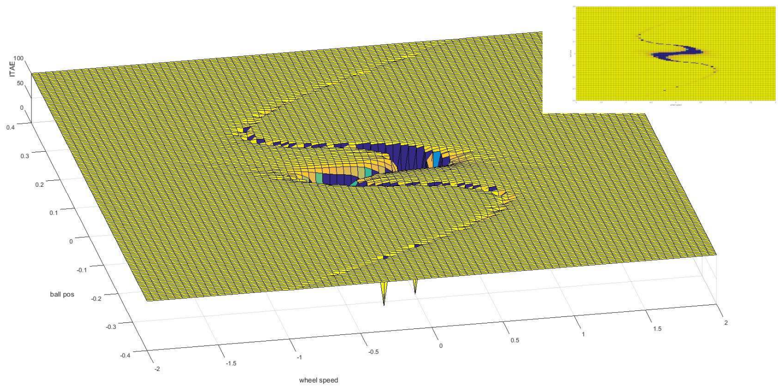


Figure 21
Narrow field

Putting the ball in a controllable position is not part of the controlling task. It can be done by a simple algorithm. Roll the ball on the bumper to get enough kinetic energy to “climb” up, turning the kinetic energy into potential energy. Near the equilibrium, the controller can be turned on. The path to the central area in Figure 11 gives us the possibility of putting the ball near to the equilibrium by the controller from the given state. We implemented this method on the physical system. Spinning the ball on the bumper using the right wheel speed for less than a second is enough for us to be able to turn on the controller and then it can put the ball to the right position and hold it there.

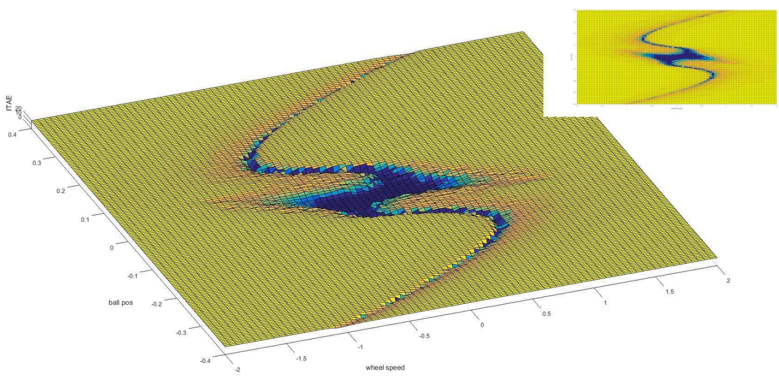


Figure 23
Big enough area with path to the center

4 Implementation

The original hardware was built by Artúr Kasovitz and Imre Dobany. The wheel's suspension was rigid so it had to be modified. The controller board was extended by the X2C module because it is our standard educational platform. [21]

For the software implementation, we used the C language on the MPLABX development environment by Microchip. With the given restrictions a very simple fuzzy controller was enough to handle the problem.

Since the primary goal of the project is to use the fuzzy controller for educational purposes, we have always tried to keep the controller as simple as possible, so we focused on the controller parameters. The mathematical model and the computer simulation allowed us to visualize the parameter space, i.e., how the controller behaves for a given initial ball position and ball speed. The visualization made it clear in which directions the controllable area needed to be increased. For this purpose, we constructed a simplified fitness function using only 5 points in the parameter space. Thus, using the mathematical model and the genetic algorithm to set up the parameters, we were able to maximize the performance of our controller and keep the implementation easy.

The dataflow of the software can be seen in Figure 13.

The software uses the laser distance sensor and the HAL sensors as input signals. The laser distance sensor is analog, so it is transformed to a digital value by the 12-bit A/D converter of the microcontroller. The position change gives the angular speed information. The output of the fuzzy controller is a signed armature voltage value. The sign is the direction information, and the absolute value gives the PWM duty cycle. The commutation event of the 6 steps controller is based on the signals of the HAL sensors and the direction information from the fuzzy controller. The output signal is controlled by the PWM duty cycle.

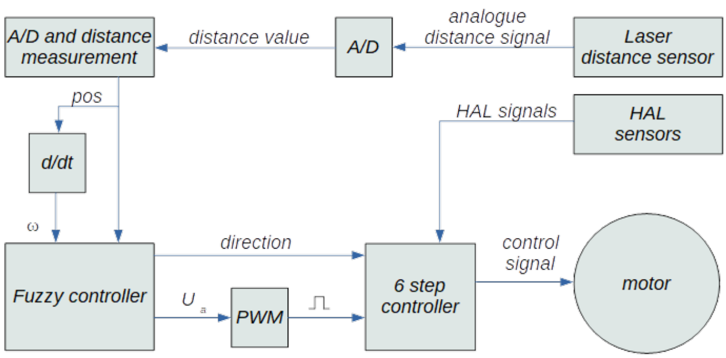


Figure 25
Software dataflow

Running the test can be seen on YouTube: https://youtu.be/hBOBb6hBWXw

Conclusions

Our goal was to build a relatively simple, fuzzy controller, for the implementation of the "ball on the wheel" system. To achieve this, we had to find a good parameter setup for the membership functions. We used a mathematical model and Matlab and Simulink tools to build up a virtual environment. In this system, we defined a relatively simple parameter space and fitness function, which was drawn as a $\mathbb{R}^2 \rightarrow \mathbb{R}$function. In this system we used a genetic algorithm to find a good setup for our controller parameters. An additional advantage of the visualization was that we found an alternative way for the initial lifting of the ball.

Acknowledgement

This work was supported by Óbuda University. We would like to thank to Artúr Kasovitz, Imre Dobány, Máté Gyimesi for their work on the hardware.

References

[1] Gunter Stein: Respect the unstable. Respect the unstable. IEEE Control Systems, (2003) 23(4), pp. 12-25

[2] Fernando G. Martins: Tuning PID controllers using the ITAE criterion, International Journal of Engineering Education, (2005) Vol. 21, No. 5, pp. 867-873

[3] H. K. Khalil: Nonlinear Systems. Upper Saddle River, (1996) NJ 07458: Prentice Hall, second ed.

[4] J. E. Slotine, W. Li: Applied Nonlinear Control, (1991) Englewood Cliffs, NJ: Prentice - Hall, Inc.

[5] M. T. Ho, J. M. Lu_ "H_∞ PID controller design for Lur'e systems and its application to a ball and wheel apparatus, International Journal of Control, (2005) Vol. 78, No. 1, pp. 53-64

[6] M. Ho, H. Lin: Balance Control of Ball and Wheel Systems via Feedback Linearization, (2006) Proceedings of the 45th IEEE Conference on Decision and Control, San Diego, CA, pp. 3926-3931

[7] S.-O. Lindert, C. Höfler, K. Schlacher: Nonlinear controlling using Raspberry PI (Nichtlineare Regelung mit einem Raspberry PI.) e & i Elektrotechnik Und Informationstechnik (2018) 135(3) pp. 286-293

[8] R. Herman: Solving differential equatins using Simulink, (2019) Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License

[9] J. H. Ginsberg: Advanced Engineering Dynamics second edition, (1998) Cambridge University Press

[10] Enrique H. Ruspini: A new approach to clustering, (1969) Information and Control Volume 15, Issue 1, pp. 22-32

[11] L. Kóczi T., D. Tikk: Fuzzy systems (Fuzzy rendszerek), (2000) Typotex Kiadó, Budapest

[12] M. Takács, Modified Distanece Based Operators in Fuzzy Approximate Reasoning, (2006) IEEE International Cenference on Mechatronics, Budapest pp. 297-299

[13] Wang, L.-X.: A Course in Fuzzy Systems and Control. (1997) Englewood Cliffs, NJ: Prentice-Hall

[14] Wang, L.-X.: Stable Adaptive Fuzzy Controllers with Application to Inverted Pendulum Tracking, (1996) IEEE Trans. on Syst., Man and Cyberneticspart B, Vol. 26, pp. 677-691

[15] Mathworks-MATLAB: Fuzzy Logic Toolbox, https://nl.mathworks.com/products/fuzzy-logic.html [Online] Available: https://nl.mathworks.com/products/fuzzy-logic.html

[16] B. Kosko: Fuzzy engineering, (1997) Prentice-Hall

[17] T. Takagi, M. Sugeno: Fuzzy Identification of Systems and its Applications to Modeling and Control, (1985) IEEE Trans. on Systems, Man, and Cybern., SMC-15(1) pp. 116-132

[18] A. Álmos, S. Győri, G. Horváth, A. Várkonyiné Kóczy: Genetic algorithms (Genetikus algoritmusok), (2002) Typotex Kiadó, Budapest

[19] William M. Spears: Evolutionary Algorithms, (2000), The Role of Mutation and Recombination. Springer, Berlin, Heidelberg, New York, (Natural Computing Series)

[20] Mathworks-MATLAB: ga Find minimum of function using genetic algorithm ,https://nl.mathworks.com/help/gads/ga.html [Online] Available: https://nl.mathworks.com/help/gads/ga.html

[21] X2C+ Development board, http://x2c.microstickplus.com/ [Online] Available: http://x2c.microstickplus.com/

[22] M. Olivares, P. Albertos: Fuzzy PD control of an unstable system, IFAC Proceedings Volumes Vol. 33, Issue 4, (2000) pp. 217-222

[23] M. Galluzzo, C. Cirino: Sliding Mode Fuzzy Logic Control of an Unstable Bioreactor (2013), Chemical engineering transactions Vol. 32, (2013) pp. 1213-1218

[24] A. Reda, A. Bouzid, J. Vásárhelyi: Model Predictive Control for Automated Vehicle Steering, Acta Polytechnica Hungarica Vol. 17, No. 7, (2020) pp. 163-182

[25] R.-C. Roman, R.-E. Precup, E. M. Petriu: Hybrid data-driven fuzzy active disturbance rejection control for tower crane systems, European Journal of Control, Vol. 58, (2021) pp. 373-387

[26] R. Precup, T. Teban, A. Albu, A. Borlea, I. A. Zamfirache and E. M. Petriu: Evolving Fuzzy Models for Prosthetic Hand Myoelectric-Based Control, IEEE Transactions on Instrumentation and Measurement, Vol. 69, No. 7, (2020) pp. 4625-4636

[27] D.-R. Wu, X.-F. Tan: Multitasking genetic algorithm (MTGA) for fuzzy system optimization, IEEE Transactions on Fuzzy Systems, 28(6), (2020) pp. 1050-1061

[28] R.-E. Precup, R.-C. David, R.-C. Roman, A.-I. Szedlak-Stinean, E. M. Petriu: Optimal tuning of interval type-2 fuzzy controllers for nonlinear servo systems using Slime Mould Algorithm, International Journal of Systems Science (2021)

[29] A. Turnip, J. H. Panggabean: Hybrid Controller Design based Magneto-rheological Damper Lookup Table for Quarter Car Suspension, International Journal of Artificial Intelligence, vol. 18, no. 1, (2020) pp. 193-206

[30] T. Haidegger, L. Kovács, R.-E. Precup, B. Benyó, Z. Benyó, S. Preitl: Simulation and control for telerobots in space medicine, Acta Astron., Vol. 181, No. 1, (2012) pp. 390-402

Petri Net-based S^3PR Models of Automated Manufacturing Systems with Resources and Their Deadlock Prevention

František Čapkovič

Institute of Informatics, Slovak Academy of Sciences, Dúbravská cesta 9, 845 07 Bratislava, Slovakia, e-mail: Frantisek.Capkovic@savba.sk

Abstract: Correct allocation of resources in Automated Manufacturing Systems (AMS) is very important, especially in order to avoid deadlocks and their consequences. Petri Nets (PN) are frequently used for modeling AMS. S^3PR (Systems of Simple Sequential Processes with Resources) model of Resource Allocation Systems (RAS) based on PN are defined, analyzed and controlled here. S^3PR are modeled by Ordinary PN (OPN). After defining and creation of such models the deadlock prevention will be performed by two deadlock prevention methods, namely (i) the method based on elementary siphons, and (ii) the method based on preventing strict minimal siphons from being emptied in another way (by means of circuits, holders of resources and complementary siphons). For illustration, two practical examples will be introduced. Both approaches are very useful not only for reliable deadlock-free control of existing AMS, but also at design of new AMS of such kind.

Keywords: automated manufacturing systems; deadlock prevention; Petri nets; resource allocation systems; siphons; supervisor; traps

1 Introduction

Petri Nets (PN) in general are defined as follows. A Petri net is a four-tuple $N = (P, T, F, W)$, where P and T are finite nonempty sets. $P = \{p_1, p_2, \ldots, p_n\}$ is a set of places ($|P| = n$) and $T = \{t_1, t_2, \ldots, t_m\}$ is a set of transitions ($|T| = m$). It is valid that $P \cup T \neq \emptyset$ and $P \cap T = \emptyset$. $F = (P \times T) \cup (T \times \mathrm{P})$ is called a flow relation of the net. It is represented by directed arcs from places to transitions and from transitions to places. $W : (P \times T) \cup (T \times P) \longrightarrow \mathrm{N}$ is a mapping that assigns a weight to an arc: $W(f) > 0$ if $f \in F$ and $W(f) = 0$ otherwise. Here, $\mathrm{N} = \{0, 1, 2, \ldots\}$, containing natural numbers plus zero.

$N = (P, T, F, W)$ is *ordinary net* (OPN) denoted as $N = (P, T, F)$, if $\forall f \in F$, $W(f) = 1$. N is said to be a *generalized net* (GPN) if $\exists f \in F$, $W(f) > 1$. Consequently, PN marking M being an $(n \times 1)$ vector (frequently named also as the state vector) is evolved in the matrix/vector form as $M_{k+1} = M_k + [N].\sigma_\mathrm{k}$, $k \in \mathrm{N}$, with M_0 being initial marking, where $[N] = [Post]^T - [Pre]$ is the $(n \times m)$ incidence matrix based on the set F, and σ_k is a $(m \times 1)$ firing vector of transitions (frequently named also as control vector). More details about PN can be found in [1]-[4]. While the foundations of PN were laid by C. A. Petri in his PhD Thesis [1], many other authors have developed the PN theory into its present form. Among them we should mention at least [2]-[4]. Some particulars about PN were mentioned also in [5]-[7].

A PN marking M is usually understand as a vector. $M(p)$ denotes the number of tokens in place p. For economy of space $\sum_{p \in P} M(p)p$ is used to denote the vector M. Place p is marked at M if $M(p) > 0$. A subset $S \subseteq P$ is marked (unmarked) at M if $M(S) > 0$ ($M(S) = 0$). If M_0 is an initial marking of a net N, (N, M_0) is called a *marked net*. A *state machine* is PN, where each transition has only one input and only one output place. PN where a place p is both an input and output place of a transition t is called *self-loop* PN. Here, in this paper, only PN without self-loops will be used.

2 Preliminaries

Resource allocation systems (RAS) represent a special class of automated manufacturing systems (AMS), where the attention is focused on resources. Resources are understood as a finite set of devices like robots, machine tools, automatically guided vehicles, transport belts, input/output devices, etc. Finite set of different processes of AMS (e.g. production lines) are competing each other for access to such resources. The competition may induce the existence of deadlocks. There exist several standard paradigms of RAS - see e.g. [8]-[11], where particulars about some of them are introduced, and [7], where a summary of most frequently used paradigms as well as their relation to PN in general are mentioned. In this paper, only one of them, namely S^{3}PR, modeled by ordinary Petri nets, will be presented and investigated. More complicated paradigms of RAS, e.g. extended S^{3}PR (ES^{3}PR) or S^{4}PR (Systems of Sequential Systems with Shared Process Resources), are modeled by means of generalized Petri nets. They will be investigated , the future.

2.1 S^{3}PR Model of AMS

A simple sequential process (S^{2}P) [7], where a review of definitions published in [12]-[14] are introduced, is a Petri net $N = (P_A \cup \{p^0\}, T, F)$, where

- $P_A \neq \emptyset$ is called the set of *activity places;* $\emptyset$ is an empty set
- $p^0 \notin P_A$ is called the idle *process place*
- N is a strongly connected *state machine*
- every circuit $\boldsymbol{C}$ of N contains place p^0.

A simple sequential process with resources (S^2PR) is a Petri net $N = (\{p^0\} \cup P_A \cup P_R, T, F)$ with P_R being a set of *resource places*, such that

- the subnet generated by $X = P_A \cup \{p^0\} \cup T$ is a S^2P
- $P_R \neq \emptyset$ and $(P_A \cup \{p^0\}) \cap P_R = \emptyset$
- $\forall t \in {}^\bullet p$, $\forall t' \in p^\bullet$, $\exists r_p \in P_R$, ${}^\bullet t \cap P_R = t'^\bullet \cap P_R = \{r_p\}$, where ${}^\bullet p$ expresses all input transitions of the place p, $p^\bullet$ represents all output transitions of p; ${}^\bullet t$ expresses all input places of the transition t, $t'^\bullet$ represents all output places of the transition t'
- the following statements are verified
  1. $\forall r \in P_R$, ${}^{\bullet\bullet}r \cap P_A = r^{\bullet\bullet} \cap P_A \neq \emptyset$
  2. $\forall r \in P_R$, ${}^\bullet r \cap r^\bullet = \emptyset$
- ${}^{\bullet\bullet}(p^0) \cap P_R = (p^0)^{\bullet\bullet} \cap P_R = \emptyset$

Here, ${}^\bullet r$ represent all input transitions of the resource place r, ${}^{\bullet\bullet}r = \bigcup_{t \in {}^\bullet r} {}^\bullet t$ is the set of all input places of all input transitions of the place r, $r^{\bullet\bullet} = \bigcup_{t \in r^\bullet} t^\bullet$ represents the set of all output places of all output transitions of the resource place r; ${}^{\bullet\bullet}(p^0)$ expresses all input places of all input transitions of the place p^0, $(p^0)^{\bullet\bullet}$ represents the set of all output places of all output transitions of the place p^0.

S^3PR N is composed of n S^2PR $N_1, N_2, \ldots, N_n$, i.e. $O_{i=1}^{n} N_i$.

2.1.1 Composition of Two S^2PR Into S^3PR

To illustrate the composition of two S^2PR N_1, N_2 into S^3PR N let us introduce the following.

An initial marking M_0 of S^2PR N is called an acceptable initial marking for N if

- $M_0(p^0) \geq 1$
- $M_0(p) = 0, \forall P_A$
- $M_0(r) \geq 1, \forall P_R$

S^2PR N with such a marking is said to be an *acceptable marked*.

In Simple Sequential Processes with Resources S^2PR N

- $P_A = P_{A1} \cup P_{A2}$
- $P^0 = \{p_1^0\} \cup \{p_2^0\}$
- $P_R = P_{R1} \cup P_{R2}$
- $T = T_1 \cup T_2$
- $F = F_1 \cup F_2$ is also S^3PR

S^{3}PR $N = N_1 \circ N_2$ (where $\circ$ symbolizes the composition of nets) is acceptably marked when

- $\forall i \in \{1,2\}$, $\forall p \in P_A \cup \{p_i^0\}$, $M_0(p) = M_{0i}(p)$
- $\forall i \in \{1,2\}$, $\forall r \in P_{Ri} \setminus P_C$, $M_0(r) = M_{0i}(r)$, where $P_C = P_{R1} \cap P_{R2} \neq \emptyset$
- $\forall r \in P_C$, $M_0(r) = \max\{M_{01}(r), M_{02}(r)\}$
- Places from P_A symbolize activities - a token in a place $p \in P_A$ models an active process - e.g. a part being processed. Places from P_R represent resources - e.g. a buffering capacity of resources, shared devices like robots and machines, etc. Tokens in a place $r \in P_R$ model the available buffering capacity of resource r. Markings represent states with a physical meaning. In this sense, only acceptable initial markings are considered. If the system is well defined and its initial marking is correct, all the markings that are reachable from it will represent possible states of the system and have physical meanings.

In Figure 1 we can see the S^{2}P N and two S^{2}PR N_1, N_2 nets, while their composition S^{3}PR net consisting of two S^{2}PR nets is displayed in Figure 2.

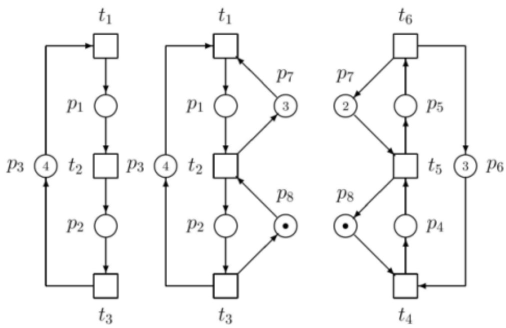


Figure 1
S^{2}P net (left) and two S^{2}PR nets (middle and right)

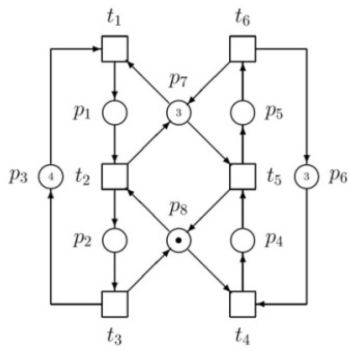


Figure 2
S^{3}PR net composed from two S^{2}PR nets N_1, N_2, i.e. $N = N_1 \circ N_2$

S^3PR in Figure 2 has the set of places $P^0 = \{p_1^0\} \cup \{p_2^0\} = \{p_3, p_6\}$, $P_A = P_{A1} \cup P_{A2} = \{p_1, p_2, p_4, p_5\}$, $P_R = P_{R1} \cup P_{R2} = \{p_7, p_8\}$. From Figure 2 is clear that ${}^\bullet p_7 = \{t_2, t_6\}$ and ${}^{\bullet\bullet}p_7 = {}^\bullet t_2 \cup {}^\bullet t_6 = \{p_1, p_5, p_8\}$, $p_7^\bullet = \{t_1, t_5\}$ and $p_7^{\bullet\bullet} = \{t_1^\bullet \cup t_5^\bullet\} = \{p_1, p_5, p_8\}$. Clearly, ${}^{\bullet\bullet}p_7 = p_7^{\bullet\bullet}$. In S^3PR, only one shared resource is allowed to be used at each stage in a job.

2.2 Deadlocks, Petri Net Siphons, Traps and *P*-Invariants

A *deadlock* in general is a state in which two or more processes are each waiting for the other one to execute, but neither can continue, , ,. Hence, deadlock is undesirable and rather bad phenomenon in PN models of real production processes.

There are four conditions for a deadlock occurring known as Coffman conditions [15]. A deadlock will never occur if one of these conditions is not satisfied. These conditions are the following:

1. There is a resource that cannot be used by more than one process at the same time (i.e. the mutual exclusion condition)
2. There are processes already holding resources are waiting for additional resources or may request new resources held by other processes (i.e. the hold and wait condition)
3. No resource can be forcibly removed from a process holding it. Resources can be released only by the explicit action of the process (i.e. not using a preemption condition)
4. Two or more processes form a circular chain where each process waits for a resource that the next process in the chain holds (the circular wait condition)

Petri net siphons, traps and invariants are structural PN parameters. They are intensively used at the deadlocks prevention. A nonempty subset $S \subset P$ is called a *siphon* if every transition having an output place in S has an input place in S. A nonempty subset $Q \subset P$ is called a *trap* if every transition having an input place in Q has an output place in Q.

If S has no token in a marking of N it remains without any token in each successor marking of N. If Q has at least one token in a marking of N it remains marked under each successor marking on N. If every non-empty siphon includes a marked trap, no dead marking is reachable. S is called an *empty siphon* at M_0 if $M_0(S) = \sum_{p \in S} M_0(p) = 0$. Such siphon is inclined to evocate deadlocks. The main aim of the deadlock prevention is the effort to prevent emptying siphons.

An $(n \times 1)$ vector I is the *P-invariant* (place invariant) if and only if $I \neq 0$ and $I^T \cdot [N] = 0^T$, where $[N]$ is the incidence matrix of N, 0 is zero vector. $\|I\| = \{p|I(p) \neq 0\}$ is called the support of I. $\|I\|^+ = \{p|I(p) > 0\}$ is the positive support of I.

2.2.1 Illustrative Example 1

The simple illustration of the siphon and trap is introduced in Figure 3. There, the siphon $S = \{p_2, p_4, p_5, p_6\}$ and the trap $Q = \{p_1, p_3, p_5, p_6\}$. Alike, in Figure 2 the siphon $S = \{p_2, p_5, p_7, p_8\}$ and the trap $Q = \{p_1, p_4, p_7, p_8\}$.

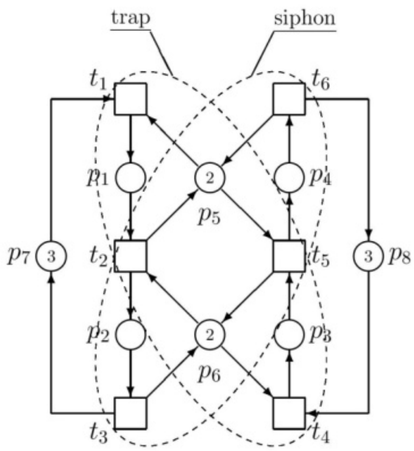


Figure 3
The siphon and trap in the S³PR N

2.2.2 Minimal, Strict Minimal and Elementary Siphons

If a siphon does not properly contain another siphon, it is called a *minimal siphon*. The set of minimal siphons is denoted by Π. A minimal siphon S is called a *strict minimal siphon* (SMS) if there is no siphon contained in it as a proper subset. A strict minimal siphon is a siphon containing neither another siphon nor trap except itself.

Having a matrix $[\lambda]$ consisting of rows being strict minimal siphons, then the linearly independent rows of the matrix $[\eta] = [\lambda]. [N]$ point out on *elementary siphons* in $[\lambda]$. Denote Π_E as a set of elementary siphons. In general - see [13], $|\Pi_E| \leq \min\{|P|, |T|\}$.

Other rows of $[\eta]$ point out on *dependent siphons* in $[\lambda]$. The dependent siphon may be *strict (strongly) dependent* or *slack (weakly) dependent.* It depends on whether the linear combination coefficients are all positive or not.

A siphon $S \notin \Pi_E$ is called *strongly dependent* siphon with respect to (w.r.t.) elementary siphons if $\eta_S = \sum_{Si \in \Pi_E} a_i.\eta_{Si}$, where $a_i \geq 0$. A siphon $S \notin \Pi_E$ is called *weakly dependent* siphon w.r.t. elementary siphons if $\exists A, B, A \neq \emptyset, B \neq \emptyset, A \cap B = \emptyset$ and $\eta_S = \sum_{Si \in A} a_i.\eta_{Si} - \sum_{Sj \in B} b_j.\eta_{Sj}$, where $a_i, b_j > 0$.

3 Deadlock Prevention in S3PR Models of RAS

There are several approaches to deadlock prevention [12]-[14], [16]-[23]. While in [12] elementary siphons are defined and their usage in the deadlock prevention is described, [13] is devoted to methods of the deadlock resolution in AMS. In [14] details of the supervisor synthesis for AMS are introduced. A survey of siphons is performed in [16], while a method of deadlock prevention without the need to enumerate complete set of siphons is presented in [17]. Different kinds of siphons, namely compound and complementary ones are analyzed in [18], while the control of elementary and dependent siphons is presented in [19]. The deadlock avoidance policy for AMS with assembly operations is described in [20]. Deadlock prevention methods depending on size of AMS are compared in [21]. Controllability for dependent siphons in S^3PR based on elementary siphons are tested in [22]. A practical usage of modeling and supervisory control in railway systems is presented in [23].

Two principled methods of deadlock prevention are presented and applied here on the S^3PR PN model of a real system and illustrated by examples: (i) the approach based on elementary siphons; and (ii) the approach based on preventing SMS from being emptied by means of analyzing circuits of PN models and using complementary siphons and downstream and upstream siphons.

3.1 Deadlock Prevention Method Based on Elementary Siphons

This approach is based on elementary siphons and siphons dependent on them. Both kinds of siphons were defined above in the part 2.2.2.

3.1.1 Illustrative Example

Consider an AMS schematically displayed in Figure 4.

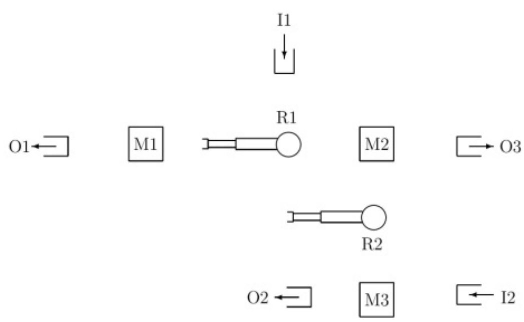


Figure 4
The scheme of the AMS structure

This AMS contains three machines M1 - M3, two robots R1, R2, two input devices I1, I2 and three output devices O1 - O3.

The scheme of the technological process being under way in it is displayed in Figure 5. Two types of parts P1 and P2 are processed as it is denoted by routing.

P1 is taken from I1 by R1 and put either into M1 or into M2. After processing P1

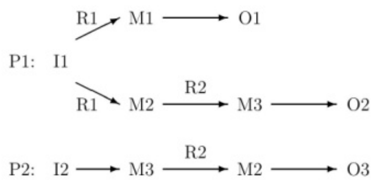


Figure 5
The scheme of the technological process of the AMS

by M1, P1 is moved to O1 by M1. After being loaded to M2, P1 is processed by M2 and then moved from M2 to M3 by R2. After being processed by M3, P1 is finally moved to O2 by M3. In the production of P1, R1 and M1, or R1, M2, R2, and M3 are used. Similarly, P2 is taken from I2 by M3, and after being processed by M3 it is moved from M3 to M2 by R2. Finally, after being processed by M2, P2 is moved to O3 by M2. To produce part type P2, M3, R2, and M2 are used. The S^{3}PR N model of the AMS is in Figure 6. Places p_7 and p_2 represent the operations of R1 and M1, respectively, i.e. one sequence at producing of the part type P1. Similarly, p_7 and p_3, p_5, p_6 represent the operations of R1, M2, R2, and M3, respectively, i.e. another sequence at producing of the part type P1. For production sequence of part type P2 p_8, p_4, p_9 represent the operations of M3, R2, and M2, respectively. The number of tokens in p_1, i.e., $M(p_1) = 5$, represents the number of concurrent activities that can take place for P1. The number of tokens in p_{10}, i.e., $M(p_{10}) = 3$, represents the number of concurrent activities that can take place for P2. Places p_{11} and p_{15} denote the resources R1 and M1, respectively. Places p_{12} - p_{15} denote shared resources M2, R2, and M3, respectively.

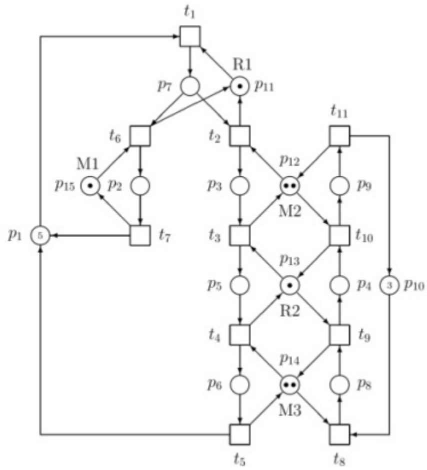


Figure 6

The S[3]PR PN model of the AMS

Initial markings of p_{11}, p_{13}, and p_{15}, are all one as robots can hold one part and M1 can process one part at a time. Initial markings of p_{12} and p_{14} are two as either of M2 and M3 can process two parts at a time.

The net has 10 minimal siphons. However, seven of them are equal to (i.e. contain) traps and ergo they are prevented from emptying. As it was pointed out in [6], [7], and also mentioned above, such siphons cannot be emptied. Namely, only siphons being inclined to be emptied are dangerous, because they may lead to deadlocks. It means that there are such 3 strict minimal siphons (SMS) here. Namely,

$S_1 = \{p_4, p_6, p_{13}, p_{14}\}$; $S_2 = \{p_5, p_9, p_{12}, p_{13}\}$; $S_3 = \{p_6, p_9, p_{12}, p_{13}, p_{14}\}$

or in the matrix form

$$[SMS] = [\lambda]^T = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

The incidence matrix of N displayed in Figure 6 is $[N] = [Post]^T - [Pre]$, where

$$[Pre] = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$[Post]^T = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$[\eta]^T = [\lambda]^T.[N] = \begin{pmatrix} 0 & 0 & -1 & 1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \end{pmatrix}$$

The *rank* $([\eta]) = 2$, because only two of its rows are linearly independent. Namely, the third row is the sum of first and second one: $\eta_{S3} = \eta_{S1} + \eta_{S2}$. It means that there are two elementary siphons S_1, S_2 and one dependent siphon S_3. To ensure the deadlock prevention in our S3PR model, we have to add two control places V_{S1} and V_{S2} in order to control S_1 and S_2, respectively. The strongly dependent siphon S_3 can never be emptied in our case. Based on η_{S1} and η_{S2} we have the supervisor with the structure $[N]_S = [Post]_S^T - [Pre]_S$ resulting from $[\eta]$, where negative entries yield $[Pre]_S$, while positive entries yield $[Post]_S^T$. Namely,

$$[Pre]_S = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

$$[Post]_S^T = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

The first row of $[Pre]_S$ symbolizes directed arcs from V_{S1} to transitions of uncontrolled N displayed in Figure 6, while the first row of $[Post]_S^T$ symbolizes directed arcs from transitions of N to V_{S1}. Analogically, the second row of $[Pre]_S$ symbolizes directed arcs from V_{S2} to transitions of N, while the second row of $[Post]_S{}^T$ symbolizes directed arcs from transitions of N to V_{S2}. Hence, we obtain the new net N_1 displayed in Figure 7. It is composed of both the original uncontrolled net N displayed in Figure 6 and the supervisor N_S. Its incidence matrix $[N_1] = \begin{bmatrix} [N] \\ [N]_S \end{bmatrix}$.

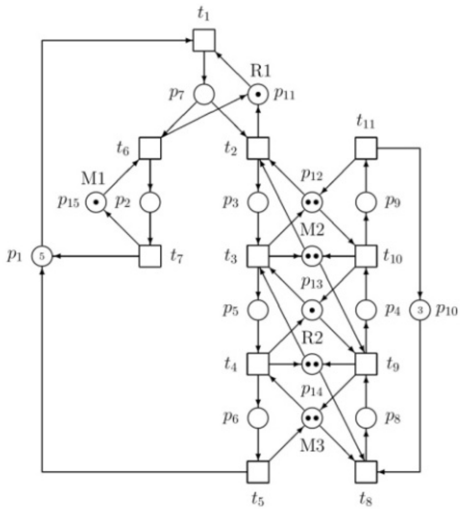


Figure 7
The controlled S³PR PN model of the AMS

3.1.2 Setting the Marking of Monitors

Now it is important to find a suitable marking of the control places (monitors) V_{S1} and V_{S2}. Namely, an inadequate setting of marking of these monitors may cause other deadlocks in the controlled plant.

For setting the marking of monitors V_{S_i}, i =1, n_m (n_m is a number of monitors) are valid the following general rules. Let $S = \{p_i, p_j, ..., p_k\}$ be SMS of a net system (N_0, M_0), where $N_0 = (P_0, T_0, F_0)$. Add a control place V_S to N_0 to make P-vector $I = (0, \ldots, 1_i, \ldots, 1_j, \ldots, 1_k, \ldots, 0, -1)^T$ be a P-invariant of a new net system (N_1, M_1), where $\forall p \in P_0 \backslash S$, $I(p) = 0$, $I(V_S) = -1$, $\forall p \in P_0$, $M_1(p) = M_0(p)$, and $[N_1]=[[N_0]^T \,|L_{VS}^T]^T$, where L_{V_S} is a row vector due to the addition of the place V_S. Let $M_1(V_S) = M_0(S) - \xi_S$, where $1 \leq \xi_S \leq M_0(S)$. Then, S is an invariant-controlled SMS and hence always marked at any reachable marking of the net system (N_1, M_1). Namely, I is a P-invariant and $\forall p \in (P_0 \cup \{V_S\}|\backslash S$, $I(p) < 0$. Note than $I^T.M_1 = I^T.M_0 = M_0(S) - M_1(V_S) = \xi_S > 0$. Thus, S is an *invariant-controlled* siphon.

To make a siphon S be always marked in a net system, we have to keep at least one token staying at S at any reachable marking of the net system. Suppose someway is found which controls S never to be emptied and the least number of tokens staying at S is denoted as, say, ξ. As mentioned above, ξ is called the *siphon control depth* variable. It is obvious the larger ξ is, the more behavior of the modeled system will be restricted, which, in Petri net formalism, means more reachable states will be forbidden. Therefore, let the siphon control depth variable be as small as possible, i.e. 1, whenever possible.

After N_0 is extended by V_S, the incidence matrix $[N_0]$ is extended by one row, denoted by L_{V_S}. Note that $I^T = (\lambda_S{}^T\text{-}1)$ and I is a P-invariant of N_1. Consequently, we have $I^T.[N_1] = 0^T$ and $\lambda_S{}^T.[N_0] - L_{V_S} = 0^T$. It means that $\lambda_S{}^T.[N_0] = L_{V_S}$ and $[N_1] = =[[N_0]^T \mid (\lambda_S{}^T.[N_0])^T]^T = [[N_0]^T \mid \eta_S]^T$. We can see that $L_{V_S} = \eta_S{}^T$.

It is easy to check from Figure 7 that $I_1 = p_4 + p_6 + p_{13} + p_{14} - V_{S_1}$ and $I_2 = p_5 + p_9 + p_{12} + p_{13} - V_{S2}$ are P-invariants of N_1. Clearly, $S_1 = \{p_4, p_6, p_{13}, p_{14}\}$ is invariant-controlled by I_1, since $\| I_1 \|^+ = S_1$ and $I_1{}^T.M_1 = M_1(S_1) - M_1(V_{S_1}) = 3 - 2 > 0$, and $S_2 = \{p_5, p_9, p_{12}, p_{13}\}$ is invariant-controlled by I_2, since $\| I_2 \|^+ = S_2$ and $I_2{}^T.M_2 = M_2(S_2) - M_2(V_{S_2}) = 3 - 2 > 0$. Here, $\| I_i \|^+$, $i = 1, 2$, are the positive supports of I_i.

S_3 is a redundant siphon. We can see that in uncontrolled net N the summary marking of S_3 (the sum of marking of its places) is $M_0(S_3) = 5$ while the summary marking of S_1 and S_2 are $M_0(S_1) = M_0(S_2) = 3$. Here, $M_0(S_i)$ means the marking of S_i.

Let $\xi_{S1} = \xi_{S2} = 1$, we have $M_0(S_3) > M_0(S_1) + M_0(S_2) - \xi_{S1} - \xi_{S2}$. Thus, S_3 can never be emptied after S_1 and S_2 are controlled via adding two control places (monitors) V_{S_1}, V_{S_2}, as it is shown in Figure 7.

It is easy to check that S_1 and S_2 can never be emptied. In such a way deadlocks in our S^{3}PR net are prevented and controlled system can operate safely (without deadlocks) and reliably.

3.2 Method Preventing Strict Minimal Siphons from Emptying

This method, sometimes called as *classical method*, consists of the work with circuits, the set of holders of resources, and with complementary siphons.

3.2.1 Circuits and Complementary Siphons in S^{3}PR

Let $\boldsymbol{C}$ be a circuit of N and x and y be two nodes of $\boldsymbol{C}$. Node x is said to be *previous* to y *iff* (if and only if) there exists a path in $\boldsymbol{C}$ from x to y, the length of which is greater than one and does not pass over the idle process place p^0. This fact is denoted by $x \prec_C y$. In general, the symbol $\prec$ means a *generic strict order relation*, while $\nprec$ symbolizes the assertion *'does not precede'*.

Let x and y be two nodes in N. Node x is said to be previous to y in N *iff* there exists a circuit $\boldsymbol{C}$ such that $x \prec_C y$. This fact is denoted by $x \prec_N y$.

Let x be a node and $A \subseteq P \cup T$ be a set of nodes in N. Then $x \prec_N A$ *iff* there exists a node $y \in A$ such that $x \prec_N y$ and $A \prec_N x$ *iff* there exists a node $y \in A$ such that $y \prec_N x$.

For $r \in P_R$, $H(r) = {}^{\bullet\bullet}r \cap P_A$ (the operation places P_A that use r), is called the set of holders of r.

$[S] = (\bigcup_{r \in S_R} H(r)) \setminus S$ is called the *complementary set* of the siphon S.

In the net N in Figure 6, $\boldsymbol{C} = p_1t_1p_7t_2p_3t_3p_5t_4p_6t_5p_1$ is a circuit and the *elementary path* $EP(p_7, p_6) = p_7t_2p_3t_3p_5t_4p_6$ is a path in $\boldsymbol{C}$. The support of $EP(p_7, p_6)$ is $\{p_7, t_2, p_3, t_3, p_5, t_4, p_6\}$ and the support of $\boldsymbol{C}$ is $\{p_1, t_1, p_7, t_2, p_3, t_3, p_5, t_4, p_6, t_5\}$. Clearly, we have $p_7 \prec_C p_6$ and $p_7 \prec_N p_6$.

Consider the same AMS with the same structure like that in Figure 4 - Figure 6, however, with another initial marking of resources, displayed in Figure 8 (left).

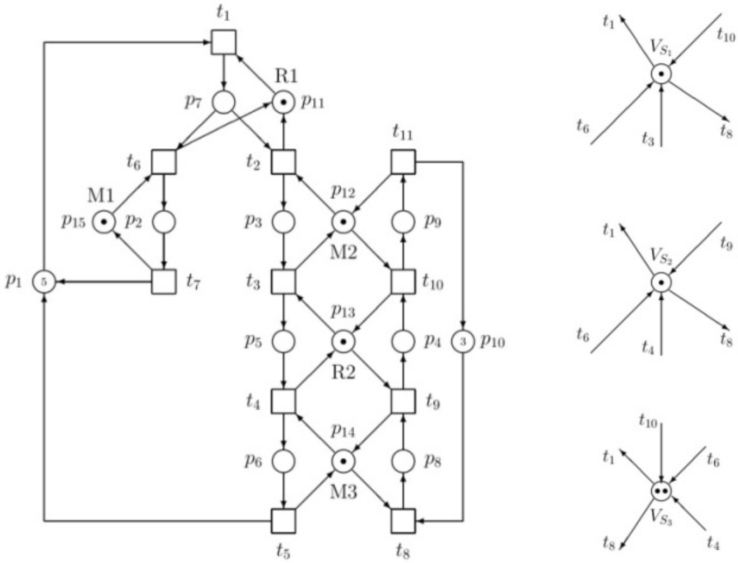


Figure 8

The uncontrolled S^{3}PR PN model of the AMS (left) and thee monitors creating the controller (right)

Circuits are structural parameters. Therefore, $\boldsymbol{C}$ introduced above in connection with Figure 6 it is valid here too. There are 3 minimal siphons in the net introduced in Figure 8, namely, $S_1 = \{p_5, p_9, p_{12}, p_{13}\}$, $S_2 = \{p_4, p_6, p_{13}, p_{14}\}$, and $S_3 = \{p_6, p_9, p_{12}, p_{13}, p_{14}\}$. They are the same as those in connection with Figure 6. Their complementary siphons are $[S_1] = \{p_3, p_4\}$, $[S_2] = \{p_5, p_8\}$, and $[S_3] = \{p_3, p_4, p_5, p_8\}$.

3.2.2 Downstream and Upstream Siphons in S^{3}PR

Let $\Delta^+(t)$ ($\Delta^-(t)$) denote the set of downstream (upstream) siphons of a transition t and $\boldsymbol{P}_S$ denote the adjoint set of a siphon S in an S^{3}PR $N = \bigcirc_{i=1}^{n} N_i = (P^0 \cup P_A \cup P_R, T, F)$. Then

1. $\Delta^+ : T \longrightarrow 2^{\Pi}$ (2^{Π} is the power set of the set Π being the set of minimal siphons) is a mapping defined as follows: If $t \in T_i$, then $\Delta^+(t) = \{S \in \Pi \mid t \prec_{\overline{N_i}} [S]^i\}$. If $S \in \Delta^+(t)$ then the set $[S]^i$ is *reachable* from t, i.e., there exists a path in $\overline{N_i}$ leading from t to an operation place $p \in P_{Ai}$ that is *not included* in S but uses a resource of S, where $[S]=\bigcup_{i=1}^{n} [S]^i$, $P_A =\bigcup_{i=1}^{n} P_{Ai}$, and $[S]^i =[S] \cap P_{Ai}$.
2. 2. $\Delta^- : T \longrightarrow 2^{\Pi}$ is a mapping defined as follows: If $t \in T_i$, then $\Delta^-(t) = \{S \in \Pi \mid [S]^i \prec_{\overline{N_i}} t\}$.
3. $\forall i \in N_n, \forall S \in \Pi$, $\boldsymbol{P}_S^i = [S]^i \cup \{p \in P_{Ai} \mid p \prec_{\overline{N_i}} [S]^i\}$, and $\boldsymbol{P}_S = \bigcup_{i=1}^{n} \boldsymbol{P}_S^i$, where $N_n = \{1, 2, \ldots, n\}$.

The downstream siphons are $\Delta^+(t_1) = \Delta^+(t_2) = \Delta^+(t_8) = \{S_1, S_2, S_3\}$, $\Delta^+(t_3) = \{S_2, S_3\}$, and $\Delta^+(t_4) = \Delta^+(t_{10}) = \emptyset$. Analogically, upstream siphons are $\Delta^-(t_1) = \Delta^-(t_2) = \Delta^-(t_6) = \Delta^-(t_7) = \emptyset$, $\Delta^-(t_3) = \{S_1\}$, $\Delta^-(t_4) = \Delta^-(t_5) = \{S_1, S_2, S_3\}$. The adjoint sets are $\boldsymbol{P}_{S1} = \boldsymbol{P}_{S1}^1 \cup \boldsymbol{P}_{S1}^2 = (\{p_3\} \cup \{p_7\} \cup \{p_4\} \cup \{p_8\} = \{p_3, p_4, p_7, p_8\}$, $\boldsymbol{P}_{S2} = \boldsymbol{P}_{S2}^1 \cup \boldsymbol{P}_{S2}^2 = (\{p_5\} \cup \{p_7, p_3\} \cup \{p_8\} = \{p_3, p_5, p_7, p_8\}$, and $\boldsymbol{P}_{S3} = \boldsymbol{P}_{S3}^1 \cup \boldsymbol{P}_{S3}^2 = (\{p_3, p_5\} \cup \{p_7\} \cup \{p_4, p_8\} = \{p_3, p_4, p_5, p_7, p_8\}$.

3.2.3 Implementation of Monitors and Setting their Markings

The net $(N_V, M_{0V}) = (P_A \cup P^0 \cup P_R \cup P_V, T, F \cup F_V, M_{0V})$ is the controlled system of the net (N, M_0) *iff*:

1. $P_V = \{V_S \mid S \in \Pi\}$ is a set of monitors and there is a bijective mapping between Π and P_V (i.e. one-to-one and onto mapping; it can be inverted).
2. $F_V = F_V^1 \cup F_V^2 \cup F_V^3$ with
   $F_V^1 = \{(V_S, t) \mid S \in \Delta^+(t), t \in P^{0\bullet}\}$
   $F_V^2 = \{(t, V_S) \mid t \in [S]^\bullet, S \notin \Delta^+(t)\}$
   $F_V^3 = \bigcup_{i=1}^{n}\{(t, V_S) \mid t \in T_i \setminus P^{0\bullet}, S \notin \Delta^-(t), {}^\bullet t \cap P_{Ai} \subseteq \boldsymbol{P}_S^i, t \nprec [S]^i\}$
3. M_{0V} is defined as follows:
   - $\forall p \in P_A \cup P^0 \cup P_R, M_{0V}(p)$; and
   - $\forall V_S \in P_V, M_{0V}(V_S) = M_0(S) - 1$.

In our example three monitors are needed to prevent three SMS from being emptied.

Take first the siphon $S_1 = \{p_5, p_9, p_{12}, p_{13}\}$ as an example. We can see that $P^0 = \{p_1, p_{10}\}$. Thus, $P^{0\bullet} = \{t_1, t_8\}$. As a result, we have $\{(V_{S1}, t_1), (V_{S1}, t_8)\} \subseteq F_V^1$.

Due to $[S_1] = \{p_3, p_4\}$, $[S_1]^\bullet = \{t_3, t_{10}\}$. We can see that $S_1 \notin \Delta^+(t_3)$ and $S_1 \notin \Delta^+(t_{10})$. Consequently, $\{(t_3, V_{S1}), (t_{10}, V_{S1})\} \subseteq F_V^2$. Let us find the arcs related to V_{S1} in F_V^3. Put $T_\alpha = (T_1 \backslash P^{0\bullet}) \cup (T_2 \backslash P^{0\bullet})$; $T_\beta = \{t|\ S_1 \notin \Delta^-(t), t \in T\}$; $T_\gamma = \{t|\ {}^\bullet t \cap P_{A1} \subseteq \boldsymbol{P}_S{}^1\} \cup \{t|\ {}^\bullet t \cap P_{A2} \subseteq \boldsymbol{P}_S{}^2\}$; $T_\delta = \{t|\ t \nprec [S_1]^1\} \cup \{t|\ t \nprec [S_1]^2\}$. Hence, $T_\alpha = \{t_2,t_3,t_4,t_5,t_6,t_7,t_9,t_{10},t_{11}\}$; $T_\beta = \{t_1,t_2,t_6,t_7,t_8,t_9\}$; $T_\gamma = \{t_2,t_3,t_6,t_9,t_{10}\}$; $T_\delta = \{t_3,t_4,t_5,t_6,t_7,t_{10},t_{11}\}$. It can be seen that $T_\alpha \cap T_\beta \cap T_\gamma \cap T_\delta = \{t_6\}$. Thus, $(t_6, V_{S1}) \in F_V^3$. For siphons S_2, S_3, monitors V_{S2}, V_{S3} can be added where $\{(V_{S2}, t_1), (V_{S2}, t_8), (V_{S3}, t_1), (V_{S3}, t_8)\} \subseteq F_V^1$, $\{(t_4, V_{S2}), (t_9, V_{S2}), (t_4, V_{S3}), (t_{10}, V_{S3})\} \subseteq F_V^2$, and $\{t_6, V_{S2}), (t_6, V_{S3})\} \subseteq F_V^3$. The supervised system is displayed in Figure 8.

As to marking of monitors $M_0(V_{S1}) = M_0(p_4) + M_0(p_9) + M_0(p_{12}) + M_0(p_{13}) - 1 = 0 + 0 + 1 + 1 - 1 = 2 - 1 = 1$; $M_0(V_{S2}) = M_0(p_4) + M_0(p_9) + M_0(p_{12}) + M_0(p_{13}) - 1 = 0 + 0 + 1 + 1 - 1 = 2 - 1 = 1$; $M_0(V_{S3}) = M_0(p_6) + M_0(p_9) + M_0(p_{12}) + M_0(p_{13}) + M_0(p_{14}) - 1 = 0 + 0 + 1 + 1 + 1 - 1 = 3 - 1 = 2$.

Three monitors in Figure 8 (right) controlling the plant with the PN model (left) are drawn separately in order to avoid confusing at drawing crisscross mutual interconnections between the PN model of uncontrolled plant and monitors. In spite of the separate drawing, it is clear which monitors are connected with which transitions.

Conclusions

Process of dealing with the allocation of resources may complicate prevention of S^3PR net systems from deadlocks. Two approaches to deadlock prevention of S^3PR net systems modeling automated manufactory systems (AMS) containing common resources (e.g. competitively utilized several manufacturing devices or other kinds of resources, like buffers of parts, etc.) were presented in this paper.

First of approaches is based on elementary siphons of the PN models of S^3PR net systems, while the second one is based on preventing strict minimal siphons of such PN models from being emptied, in another way. The former approach is more analytical (better expressed in analytical terms) and more friendly for processing by computer because it uses linear algebra. The latter approach, based on preventing strict minimal siphons from being emptied, utilizes the analysis of circuits, computing the holders of resources and complementary siphons, what needs some preprocessing or more complicated algorithm able to handle operations from the set theory. On the other hand, also in the former approach the situation may be a little hindered by the needfulness to deal with strongly or weakly dependent siphons if it is necessary (i.e. if they are not automatically prevented before emptying already within the framework of preventing the elementary siphons).

In any case, both approaches are very useful in the deadlock prevention in AMS with the requirement of correct resource allocation. Moreover, they are very useful not only for reliable control of existing AMS, but also at the design of new deadlock-free AMS like those.

For computing S^3PR siphons and traps themselves, the MATLAB based tool GPenSIM [24] was used. MATLAB itself (or at least GNU Octave) is suitable also for computer application both of the deadlock prevention methods.

Benefits following from the application of such deadlock prevention methods yield deadlock free RAS designed off-line (still before their actual deployment in practice). It means that such methods intensively help us at the AMS design. This at the least rapidly decreases a risk of defects in operation of real AMS as well as prevents their shutdowns. In such a way, it is possible to avoid significant economic losses. It is main advantage of the approaches presented in this paper. However, other external disturbances unrelated to deadlocks, cannot be prevented in such a way.

Acknowledgement

This work was partially supported by the Slovak Grant Agency for Science VEGA under Grant No. 2/0020/21.

References

[1] C. A. Petri: Communication with Automata, Ph. D. Thesis. Technical University od Darmstadt, Germany, 1962, 128 pages (in German)

[2] W. Reisig: Petri Nets, Springer, Berlin Heidelberg, 1985

[3] T. Murata: Petri Nets: Properties, Analysis and Applications, Proceedings of the IEEE, Vol. 77, No. 4, 1989, pp. 541-580

[4] J. Desel, W. Reisig: Place/Transition Petri Nets. In: W. Reisig, G. Rozenberg (Eds.): Advances of Petri Nets, Lecture Notes in Computer Science, Vol. 1941, Springer, Heidelberg, 1998, pp. 122-173

[5] F. Čapkovič: Modeling and Control of Discrete-Event Systems with Partial Non-Determinism Using Petri Nets. Acta Polytechnica Hungarica, Vol. 17, No. 4, 2020, pp. 47-66

[6] F. Čapkovič: Control of Deadlocked Discrete-Event Systems Using Petri Nets. Acta Polytechnica Hungarica, Vol. 19, No. 2, 2022, pp. 213-233

[7] F. Čapkovič: Modeling and Control of Resource Allocation Systems within Discrete-Event Systems by Means of Petri Nets - Part 1: Invariants, Siphons and Traps in Deadlock Avoidance. Computing and Informatics, Vol. 40, No. 3, 2021, pp. 648-689

[8] X. Guan, Y. Li, J. Xu, C. Wang, S. Wang: A Literature Review of Deadlock Prevention Policy Based on Petri Nets forAutomated Manifacturing Systems, International Journal of Digital Content Technology and its Applications (JDTCA), Vol. 6, 2012, pp. 426-433

[9] H. Yue, K. Y. Xing, H. S. Hu,W. Wu, H. Y. Su: Petri-Net-Based Robust Supervisory Control of Automated Manufacturing Systems, Control Engineering Practice, Vol. 54, 2016, pp. 176-189

[10] A. Farooq, H. Huang, X. L. Wang: Petri Net Modeling and Deadlock Analysis of Parallel Manufacturing Processes with Shared-Resources, The Journal of Systems and Software, Vol. 83, 2010, pp. 675-688

[11] H. Hu, Y. Liu, L. Yuan: Supervisor Simplification in FMs: Comparative Studies and New Results Using Petri Nets, IEEE Transactions on Control System Technology, Vol. 24, 2016, pp. 81-95

[12] Z. W. Li and M. C. Zhou: Elementary Siphons of Petri Nets and their Application to Deadlock Prevention in Flexible Manufacturing Systems. IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans, Vol. 34, No. 1, 2004, pp. 38-51, doi:10.1109/TSMCA.2003.820576

[13] Z. W. Li and M. C. Zhou: Deadlock Resolution in Automated Manufacturing Systems. A Novel Petri Net Approach, London, Springer Press, 2009

[14] G. Y. Liu: Supervisor Synthesis for Automated Manufacturing Systems Based on Structure Theory of Petri Nets. Doctoral Thesis, Paris Graduate School for Informatics, Telecommunications and Electronics, Paris, France, 2014

[15] E. G. Coffman, M. J. Elphick, A. Shoshani: Systems Deadlocks, ACM Computing Surveys, Vol. 3, No. 2, 1971, pp. 66-78

[16] G. Y. Liu and K. Barkaoui: A Survey of Siphons in Petri Nets. Information Sciences, Vol. 363, 2016, pp. 198-220. Available at: https://doi.org/10.1016/j.ins.2015.08.037

[17] C. F. Zhong and Z. W. Li: A deadlock prevention approach for flexible manufacturing systems without complete siphon enumeration of their Petri net models. Engineering with Computers, Vol. 25, 2009, pp. 269-278

[18] D. Y. Chao and Y. L. Pan: Uniform Formulas for Compound Siphons, Complementary Siphons and Characteristic Vectors in Deadlock Prevention of Flexible Manufacturing Systems. Journal of Intelligent Manufacturing, Vol. 26, No. 1, 2015, pp. 13-23, doi: 10.1007/s10845-013-0757-7

[19] Z. W. Li and M. C. Zhou: Control of Elementary and Dependent Siphons in Petri Nets and Their Application. IEEE Transactions on Systems, Man, Cybernetics, Part A, Systems and Humans, Vol. 38, No. 1, 2008, pp. 133-148, doi: 10.1109/TSMCA.2007.909548

[20] J. C. Luo, Z. Q. Liu and M. C. Zhou: A Petri Net Based Deadlock Avoidance Policy for Flexible Manufacturing Systems with Assembly Operations and Multiple Resource Acquisition. IEEE Transactions on Industrial Informatics, Vol. 15, No. 6, 2019, pp. 3379-3387

[21] E. A. Nasr, A. M. El-Tamimi, A. Al-Ahmari, H. Kaid: Comparison and Evaluation of Deadlock Prevention Methods for Different Size Automated Manufacturing Systems. Hindawi Publishing Corporation, Mathematical Problems in Engineering, Vol. 2015, Article ID 537893, 19 pages. http://dx.doi.org/10.1155/2015/53789

[22] D. Y. Chao: Improved Controllability Test for Dependent Siphons in S^3PR Based on Elementary Siphons. Asian Journal of Control, Vol. 12, No. 3, 2010, pp. 377-391

[23] M. P. Cabasino, A. Giua and C. Seatzu: Modeling and supervisory control of railway networks using Petri nets, IEEE Transactions on Automation Science and Engineering, Vol. 5, No. 3, 2008, pp. 431-445

[24] R. Davidjaruh: GPenSIM, General Purpose Petri Net Sumulator for MATLAB Platform. Available at: http://www.davidrajuh.net/gpensim/

Method and System for Measuring the Cutting Torque and Feed Force used for Hole Processing

Sándor Ravai-Nagy

Technical University of Cluj-Napoca, Faculty of Engineering, Baia Mare, str. Dr.V.Babes, 62A, 430083, Baia Mare, Romania, sandor.ravai@imtech.utcluj.ro

Abstract: The introduction of new materials in the cutting processes generates the need for supplementing existing knowledge with new data specific to these materials. By monitoring the machining parameters, information through which these can be optimized is obtained. The aim of the paper is highlighted by the author's concern for the development of a system for measuring the feed forces and cutting torques for research and educational activities. The proposed measurement system is developed for drilling processes. In this paper, the components of the equipment and their func-tions are presented. The research method is the experimental one. The final part of this research presents some demonstrative tests meant to exemplify the capability of the system, as well as the conclusions drawn and the personal point of view which highlights new fields of research.

Keywords: feed force; cutting torque; drilling; data acquisition; measurement; equipment development

1 Introduction

The machining of plastics has specific features [1, 2]. Within the framework of experiments carried out by several researchers in the field of hole machining [3-5], the influence of the machining parameter on the quality of surfaces and dimensional accuracy was studied [6-8]. For in-depth knowledge of the process, the monitoring of cutting torque [9-11] and feed force [12-14] would be useful. By monitoring the cutting torque and feed force we will be able to get a more comprehensive picture of the phenomena that occur in the area of chip removal [15-17].

The aim of the work is to develop low-cost research equipment with high measurement accuracy. The equipment is developed for the study of the cutting process (cutting torque and feed force) with twist drills in industrial plastics. This study is needed because the behaviors of plastics is very different from that of metals when they are being cut. Small changes in the drill – cutting parameters assembly have major effects on the quality of the machined hole [18].

By creating certain equipment or a device for measuring the cutting torque and feed force, research and publication opportunities will be created, such as:

- The influence of low temperatures on the accuracy of machining and on the quality of the processed surface in the case of industrial plastics;
- Dimensional accuracy of holes processed in industrial plastics;
- The influence of cutting speed when machining the thread with the tap on the resistance of the thread;
- Study of the process of the internal thread forming with thread forming taps;
- Study of threaded metal insertions mounted in parts made of industrial plastic and aluminum.

During the development of the equipment, the main criterion was its execution incurring minimum expenses. There exists equipment for monitoring cutting parameters, and an interest on the researchers' part in the development of such equipment can be seen. Recently, various simple or complex systems have been developed for measuring feed forces and cutting torque.

In educational systems, cost related aspects when ensuring the logistical aspects required for the experiments to be carried out with the involvement of the students at bachelor, master and doctoral level are fundamental. When developing a method and a specific piece of equipment, the directions of experimental research in which the local academic environment is involved and wants to develop were analyzed. Said development is hampered or blocked by the high costs incurred by the purchasing of measuring equipment.

By implementing the method and the equipment proposed in this research, the quality of the experiments, carried out with the students will increase. It becomes possible to present and repeat experiments on the basis of which students become familiar with various aspects related to the machining of materials, the influences of the machining parameters and processing conditions on the accuracy and quality of the surfaces resulting from the processing.

The purpose of developing the measurement system and method leads to:

- The possibility of carrying out demonstrations within the framework of teaching activities, and
- The building of the applied research facilities.

2 State of Art on the Characteristics of Existing Dynamometers

Studying the technical data of the equipment offered for sale or developed by the research teams, I made a synthesis of the measurable cutting moments, Table 1.

In the paper I present an equipment developed for measuring the cutting torque in the range 0-0.5 Nm respectively 0-40 Nm to study the processing of holes in plastics by cutting. With this equipment I focus my studies on holes with diameters in the range 0-24 mm.

Table 1

Synthesis of existing dynamometers

No.	Equipment	Cutting torque [Nm]	References
1	Kistler 9272 (plate dynamometer)	–200 ... 200	[19]
2	Kistler 9170A (rotating dynamometer)	–150 ... 150	[19]
3	Kistler 9171A (rotating dynamometer)	–1 000 ... 1 000	[19]
4	Kistler 9109AA - MicroDin (plate dynamometer)	-50 … 50	[22]
5	Research device 1 (rotating dynamometer)	0 ... 160	[16]
6	Research device 2 (rotating dynamometer)	0 ... 30	[20]
8	Research device 3 (rotating dynamometer)	0 ... 40	[21]
9	Research device 4 (rotating dynamometer)	0 ... 10	[15]
10	Developed device v1 (plate device)	0 ... 0.5 (0 ... 1)	Present paper
11	Developed device v2 (plate device)	0 ... 20 (0 ... 40)	Present paper

The identified measuring devices and equipment partially cover the measuring range 0-10 Nm and the hole diameter 0-24 mm, the field in which the cutting torque we want to study.

Through the designed equipment, I complete this gap and create the facility to research with a high accuracy the moments that appear when processing the holes with the diameter in the range 0-24 mm.

Through these measurements (with a high accuracy) we propose the analysis the cutting torque and his influence on processing precision and the quality of the obtained surface.

3 System and Method for Measuring Cutting Torque and Feed Force when Processing Holes

The proposed method and working principle aim is to highlight the influence of machining parameters on the feed force and the cutting torque on drilling.

The requirements imposed on the drafting and design of the measurement system and method were:

- real time measurement of the feed force and the cutting torque when machining holes and internal threads;
- graphical visualization of the variation of the feed force and of the torque during processing;
- the acquisition of data regarding feed force and cutting torque during processing in order to further analyze the experimentally collected data.

The objectives pursued by achieving the measurement system and method:

- real time measurement and visualization of the feed force and/or cutting torque when processing hole type internal surfaces of rotation and internal threads;
- acquisition of data regarding feed force and cutting torque;
- processing and reading the experimental data resulting from measuring the feed force and cutting torque;
- visualization of phenomena during machining through their influence on the cutting torque and feed force;
- studying the feed force, cutting torque and their effects on the following machining processes:
    - Drilling;
    - Thread tapping;
    - Thread forming;
    - Reaming.
- the behavior of different plastics when processing them through machining;
- studying the influence of cutting tool geometry on the feed force and cutting torque;
- studying the surface quality depending on machining parameters and the phenomena identified during processing;
- determining the influence of low temperature on the feed force and the surface quality resulting from processing.

4 Operation Principle of the Proposed Measuring System

During the machining, the twist drill type cutting tool is stressed by a cutting torque and a feed force. Some of the measuring solutions of the torque and force values are oriented to their determination with equipment fixed on the cutting tool. The force and torque, which stress the cutting tool, will also stress the semi-finished product, except it will do so in the opposite direction.

In the case of the proposed solution (Figure 1), the force and the torque will be measured with the help of certain measuring equipment having also the function of fixing the semi-finished product to the machine tool. The cutting tool is fixed in the machine tool spindle with dedicated tool holders.

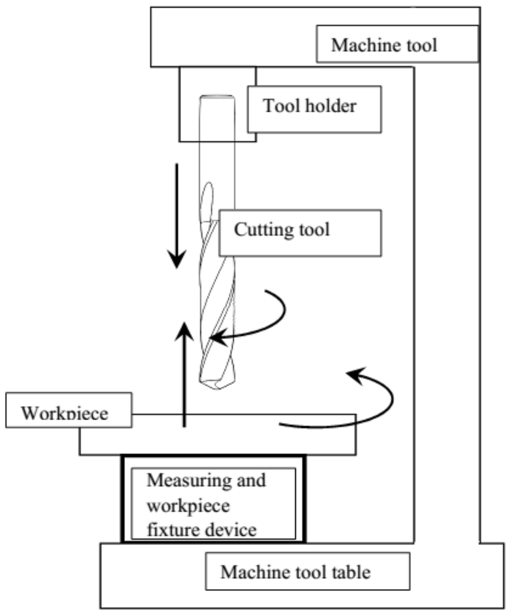


Figure 1
Essential connections

5 The Structure of the Measuring Device

The measuring device has been developed in a modular structure so that it is easy to use in research.

As a component it consists of a "3" base plate on which the other subassemblies are mounted, depending on the application for which the system is used.

Figures 2, 3 and 4 and Table 2 show the component parts of the system depending on the measurement objectives.

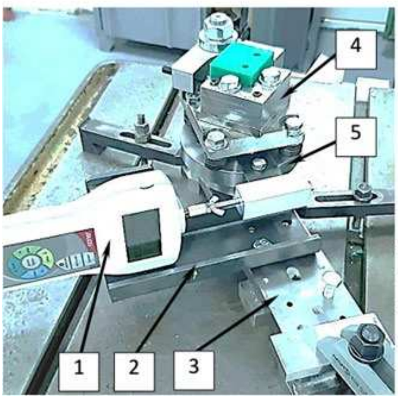
Table 2
Components of the measuring system

No.	Designation	Comments
1	Dynamometer A	(used to measure the force component of the cutting torque)
2	Support of dynamometer A	
3	Base plate	
4	Test item fixture device	
5	Torque module	
6	Spacer for dynamometer support A	
7	Dynamometer B (external sensor)	(used for measuring the feed force)
8	Assembly of thermal insulation elements	(used in case of cryo-cutting type experiments)

5.1 How to Assemble the Measuring System Depending on the Measurement Objectives

Case 1. At measuring the cutting torque (Figure 2), the equipment will be configured as follows:

On the "base plate" 3 using the "support" 2 the "dynamometer A" marked 1 will be fixed. Also on the "base plate" 3 the ensemble 5 "torque module" is fixed, on which the element 4 "test item fixing device" is mounted.



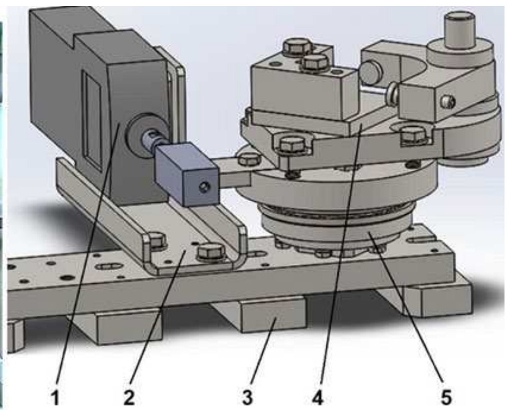


Figure 2
Assembled measuring device for torque measurements

After this, the measuring rod of the dynamometer is coupled with the arm of the torque module.

The ensemble thus made is placed, centered and fixed on the table of the drilling or milling machine tool.

Case 2. In case one desires to measure both the torque and the feed force (Fig. 3), ele-ments 6 "spacer" and 7 "dynamometer B" will also appear in the structure of the said equipment.

When assembling the measuring system, between the base plate "3" and the torque module "5", the element 7 "External sensor of Dynamometer B" is mounted.

By adding the "External sensor of Dynamometer B" 7 the feed force can be measured. The position of the torque module 5 towards the base plate 3 being modified in its turn, the dynamometer A 1 will also be repositioned by mounting a spacer 6 between the base plate 3 and the support 2. The height of the spacer is equal to the height of the external sensor of the dynamometer B.

Case 3. Research at low temperatures

In this case, between the torque module 5 and the test item fixture device 4, an en-semble of parts with the role of thermal insulation is introduced. It is also on the thermal insulation elements that the tray with the cryogenic agent for the cooling of test item is mounted. For these experiments, the test item fixture device will also be cooled.

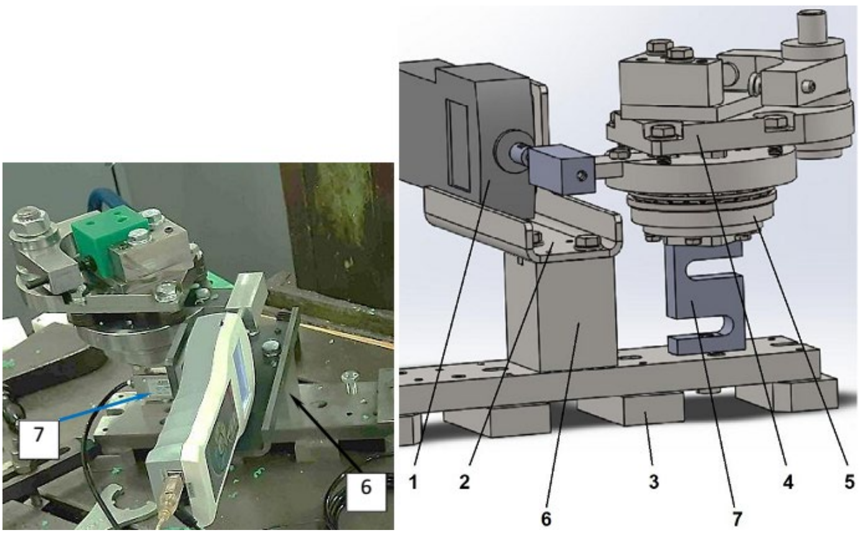


Figure 3

Assembled measuring device for torque and feed force measurements

5.2 Test Specimen

Within the framework of the experiments performed using the proposed and implemented measuring system and method, rectangular parallelepiped test specimens shall be used.

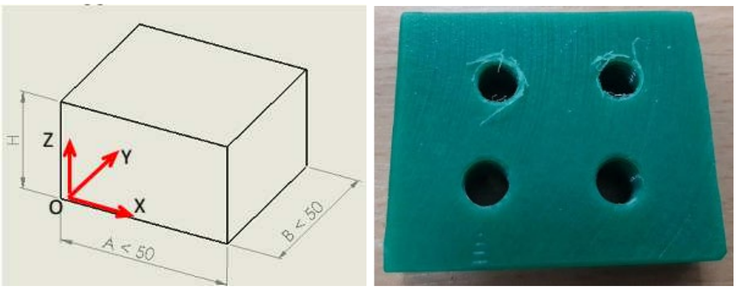


Figure 4
Test specimen

The test specimen fixture device was made so that it could fix the test specimen with the side of the base (A and B) measuring less than or being equal to 50 mm. The height of the test item (H) will be chosen depending on the depth of the hole that is being processed.

5.3 Test Specimen Fixture Device

The device, by its construction, will ensure the same positioning origin for all the pieces from the experimental lots. The origin and axes of the device are the same (Figure 5) as the origin and axes of the test specimen (Figure 4). By rotating the rectangular parallelepiped test specimen, each corner of it can be an origin for each machined hole. This aspect is important, especially when processing the threads, because first the holes can be processed with a twist drill in each test specimen, and then one changes the twist drill and the tool holder of the twist drill with the tool holder with tap collet in which the tap for the experiment is fixed.

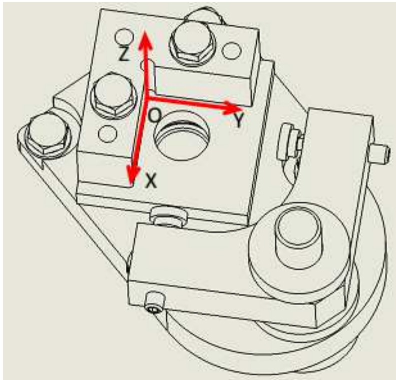


Figure 5
Test specimen fixture device

5.4 Torque Module

The torque module 5 is a set of parts consisting of circular plates of special configurations (5.1, 5.2, 5.5), centered together by means of two bearings (one radial 5.3 and one axial 5.4) so that they rotate smoothly in relation to each other.

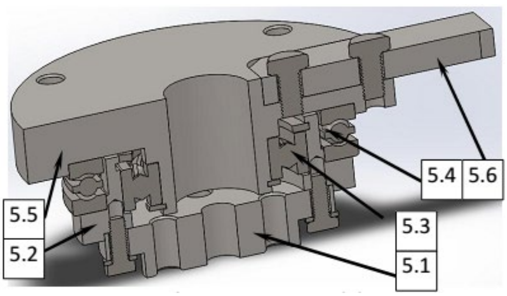


Figure 6
Torque module

The arm 5.6 of the torque module rests on the rod of the dynamometer A (marked with 1, in Figure 2) in order to measure the force component of the cutting torque. Knowing the length of the arm and the force at the end of the arm, the torque can be determined.

The measuring device was designed with two arms 5.6 of different lengths in order to increase the measuring field of the cutting torques. This is why two sets of holes for fixing the support 2, of the dynamometer A, are machined on the base plate 3.

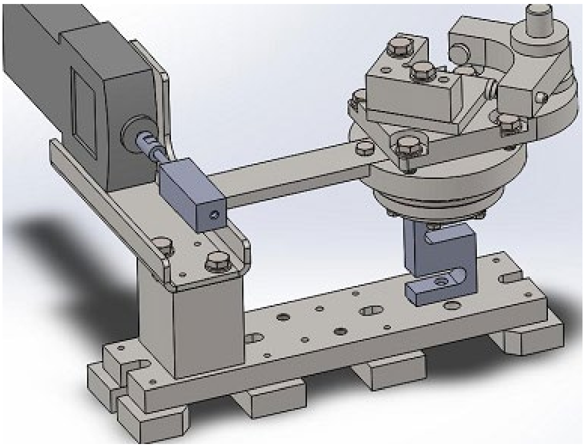


Figure 7
Measuring device in the version with arm of 200 mm

5.5 Measuring Accuracy

Considering that dynamometer A can measure forces up to 200 N, depending on the size of the cutting torque predicted for measurement, the arm of 100 mm or the arm of 200 mm will be used.

With the device equipped with arm of L1=100 mm, it will be possible to measure the cutting torque in the range 0-20 Nm. According to Figure 8 and knowing that the maxi-mum force which can be measured with the dynamometer A (FB200) is 200 N, it is possible to determine the maximum torque that can be measured according to relation 1, as well as the accuracy of this measurement (relation 2).

Measured value: 200 N x 0.1 m = 20 Nm (1)

Measuring accuracy: 0.05 N x 0.1 m = 0.005 Nm (2)

Where:

200 N - the maximum force that can be measured with dynamometer A (FB200)

0.05 N – measuring accuracy of the dynamometer A (FB200).

0.1 m – the length of the arm of the torque module 5

In the case of a 5 N dynamometer (FB5 dynamometer) with a measuring accuracy of 0.001 N according to those presented above, we will be able to measure the cutting torque in the range 0-0.5 Nm with a measuring accuracy of 0.0001 Nm.

With the device equipped with an arm of L2=200 mm and 200 N dynamometer, it will be possible to measure the cutting torque in the range 0-40 Nm.

The calculation of the maximum measuring torque and its measuring accuracy are determined by relations 3 and 4. In case of 5 N dynamometer the measuring range are 0-1 Nm.

Measured value: 200 N x 0.2 m = 40 Nm (3)

Measuring accuracy: 0.05 N x 0.2 m = 0.01 Nm (4)

Where:

200 N - the maximum force that can be measured with dynamometer A (FB200)

0.05 N - measuring accuracy of the dynamometer A (FB200).

0.2 m – the length of the arm of the torque module 5

Using torque arms below 1 m (the length of the arm of the force resulting from the cutting torque being 0.1 and 0.2 m respectively) the measuring accuracy of the cutting torque will be:

- 5 times higher for the torque module with arm L=0.2 m and

- 10 times higher at the torque module with arm L=0.1 m than the measuring accuracy of the dynamometer used.

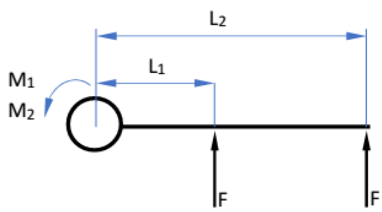


Figure 8
Scaling of measuring accuracy

5.6 The Need of a Correction Coefficient of the Measuring System

During the design phase the frictional forces that appear in the bearings were considered zero because, depending on the bearing manufacturer and its mounting conditions, they will have different values.

After making the equipment, the chosen solution was to experimentally determine the correction coefficient applied to the cutting torque calculated with the formula: the force measured with the dynamometer A x the length of the arm of the torque module.

Thus, I recommend checking the correction coefficient through measurements before each mounting of the equipment and its positioning on the machine tool on which the experiments will be carried out. At the same time, I can conclude that each piece of equipment must be calibrated with its own correction coefficient.

5.7 Determination of the Correction Coefficient of the Measuring System

To determine the correction coefficient a torque screwdriver is used. In this case, we used a Proxxon MicroClick MC5 torque screwdriver with a measuring range of 1-5 Nm and according to the ”Certification of Conformity” with a measurement error of 0,03 Nm, with uncertainty of 0,13

The technological ensemble used to determine the correction coefficient is shown in Figure 9.

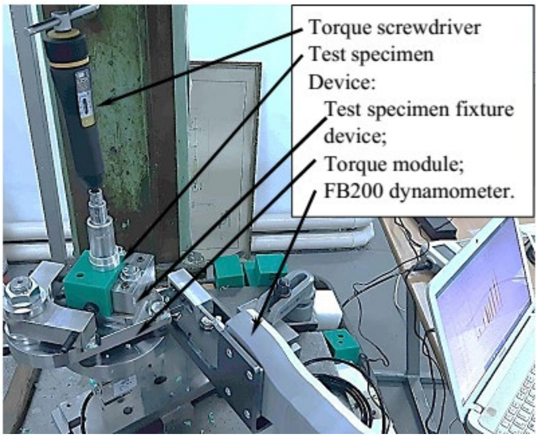


Figure 9
Determination of the correction coefficient

On the torque screwdriver each value of the known torque indicated in the "Certification of Conformity" was adjusted successively, and the device was actuated by the test item element so that the dynamometer record the value of the force component of the torque. The correction coefficient k is defined (formula 5) by the ratio between the test torque (torque set on the torque screwdriver) and the torque determined with the developed equipment (calculated from the force measured by the dynamometer and the arm length).

$$k = \frac{a}{b} \tag{5}$$

Where:

- torque adjusted on the torque screwdriver
- torque determined with de developed equipment (force x arm).

In Table 3, I centralized the measured data and the calculated coefficients. Figure 10 shows the graph of variation of the coefficient calculated according to the torque set on the torque screwdriver.

Table 3
Measured data and the calculated coefficient

Torque screwdriver MC5	Developed equipment			Correction coefficient	
Adjusted torque	Measured force	Arm of force (construc ive)	Determined torque		Average
a			b	ki	k
[1]	[2]	[3]	[4] = [2x3]	[5] = [1 / 4]	[6]
Nm	N	m	Nm		
1.01	9.65	0.1	0.965	1.046	1,05
3.1	29.45	0.1	2.945	1.052	
5.28	50.25	0.1	5.025	1.050	

Obs.

- The values of the torque in the column [1] are values certified by the metrological bulletin.

- The coefficient k is obtained as the arithmetic average of the coefficients determined in column [5]

Based on the experimental data presented in Figure 9 and Table 3 in this case a correction coefficient of 1.05 will be used within the framework of experiments performed with the proposed system.

When using the arm L2=200 mm of the torque module, the determination of the correction coefficient is determined analogously to the method presented.

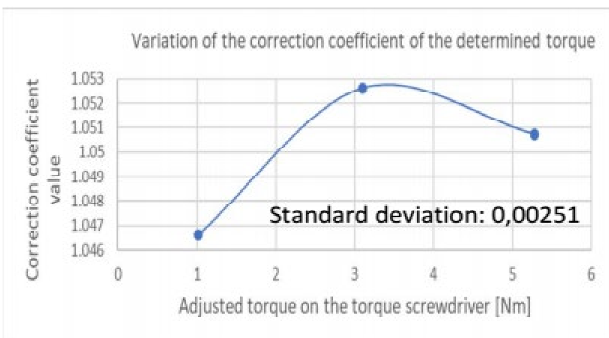


Figure 10
Variation of the torque correction coefficient

6 Applications with the Method and System Proposed for Measuring the Cutting Torque and the Feed Forces used in Hole Processing

To demonstrate the capability of the method and the system proposed and subsequently performed, several experiments are presented, as well as the working technique, the obtaining of the experimental data and some fields of experimental research for which the system will be used.

The presented experiment is the measurement of cutting torque and the feed force when machining holes with a twist drill.

It is our wish to highlight:

- the need for research in the field of new materials for which there is no research data (plastics, composite materials, etc.)
- the aim to reduce research costs
- the aim to support experimental research and education.

6.1 Experiment: Measuring the Cutting Torque and the Feed Force in Hole Processing with the Twist Drill

Within the framework of the experiment, I present the mode of operation for measuring the feed force and cutting torque for processing a hole with a diameter of Ø 6.9 mm, with a twist drill of general use HSS-R ISO 235 (DIN 338), on a universal drilling machine, at a spindle speed of 550 rot/min and a feed of 0.10 mm/rot. Processed material PA6.

6.1.1 Phase 1. Aligning the Equipment with the Machine Tool Spindle

The measuring device is placed on the table of the drilling machine, and the spindle of the torque module is aligned with the axis of the main spindle of the drilling machine.

The alignment of the axis of the measuring system and of the drilling machine (implicitly also of the cutting tool in the current situation of the drill) is important because the coaxially errors will negatively influence the accuracy of the measured values of the cutting torque.

The alignment is achieved by fixing a dedicated centering device in the drilling machine spindle, preferably with a meter with dial, the feeler of which will come into contact with the bore in the test item fixture device. (From the constructive viewpoint), since the assembly phase the concentricity and coaxiality between the bore in the tool holder and the bore of the torque module is ensured.)

The device is fixed on the table of the drilling machine with some clamps.

The dynamometers are connected to the computer on which the experimental data obtained will be recorded and saved. AxisFM data acquisition software is launched for each dynamometer.

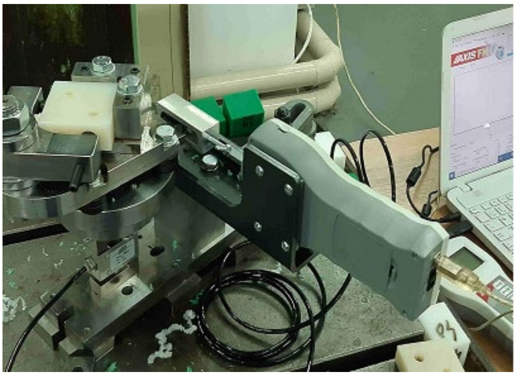


Figure 11
Phase 1. Alignment of equipment according to the axis of the machine tool

6.1.2 Phase 2. Equipment Calibration

At this stage, the proper operation of the dynamometers is checked and the correction coefficient of the measured cutting torque is established.

The dynamometer is calibrated for measuring the feed force:

- reset the dynamometer in this way the mass of the torque module, the mass of the test item fixture device, and the mass of the test item are eliminated.
- place weights with the known mass on the test item fixture device and compare the mass of the weight with the value indicated by the dynamometer.
- determine the correction coefficient of the measured torque.

Follow the steps already described in the chapter "Establishing the correction coefficient of the cutting torque".

Practically, after centering the axes of the torque module and the cutting tool, a test item is fixed in the tool holder in which an internal thread is machined in which a screw is inserted. Using this screw, the known torques will be applied, adjusted on the torque screwdriver, and the values of the forces at the end of the arm of the torque module will be recorded.

Based on the data recorded by dynamometer A, the correction coefficient of the cutting torque will be determined.

At each mounting of the measuring system on a machine tool it is necessary to go through phase 1 “Aligning the equipment with the machine tool spindle” and phase 2 “Equipment calibration” in order to ensure the measurement accuracy and obtain correct data.

6.1.3 Phase 3. Carrying out the Actual Experiment

Fix the studied cutting tool in the drilling machine spindle, with the specific tool holder devices;

Adjust the drilling machine feed travel limiter switch to protect the measuring device from possible accidental processing;

Fix the test item made from the studied material in the test item fixture device;

Adjust the spindle speed and the feed;

Set the ”sampling time” on the dynamometers;

Start the data acquisitions for the dynamometers;

Start the drilling machine and with the mechanical feed and the spindle speed already adjusted, process a hole. Extract the cutting tool from the hole and stop the drilling machine;

Stop the data acquisition of the dynamometers;

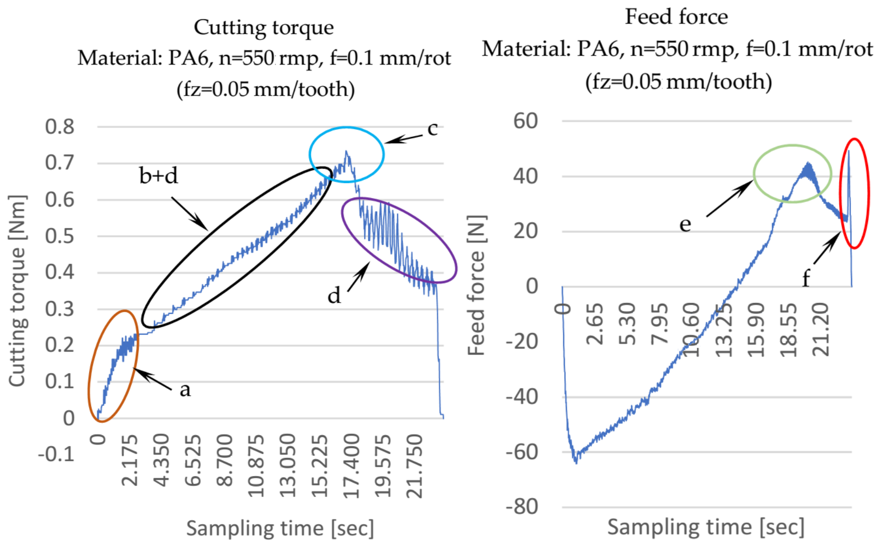
Open an Excel file and every time the data obtained from each dynamometer is exported on the first ”sheet”, resulting in two ”sheets” with data referring to feed force and force component of the cutting torque. Save the file.

The experiment was repeated 3 more times to complete the holes in the test specimen.

6.1.4 Phase 4. Data Processing

The Excel file with the experimental data is opened and the cutting torque is calculated with the following data: the measured force, the arm of the torque module and the correction coefficient (paragraphs 5.6, 5.7).

In Figure 12 present the graph of the variation of the feed force and the cutting torque obtained experimentally.



a) b)

Figure 12
Variation of the feed force and the cutting torque: (a) cutting torque; (b) feed force

To highlight the repeatability of the measurements in Figure 13 present in the same graph the cutting torques for the 4 holes processed in the test specimen.

According to the research topic, processing and interpreting the data will continue.

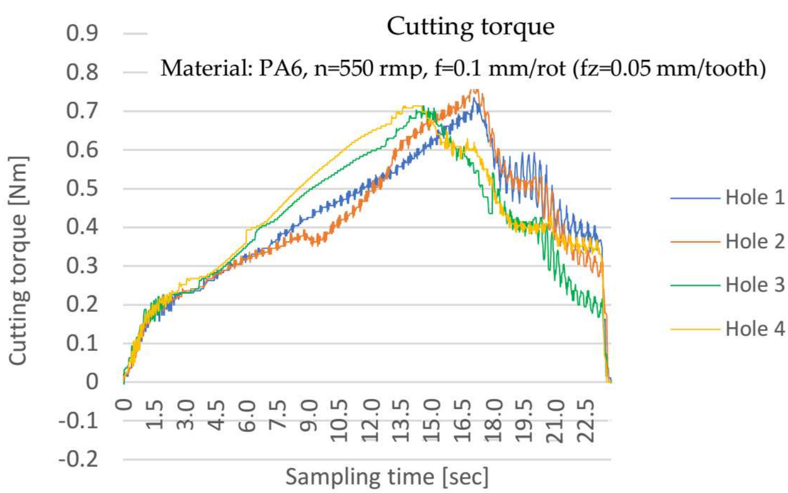


Figure 13
Repeatability of the measurements

6.1.5 Phase 5. Result Analysis

The graph in (Figure 12a) shows the evolution of the force component of the cutting torque from the beginning of the bore processing to the withdrawal of the drill from the hole.

The graph in (Figure 12b) shows the evolution of the feed force from the beginning of the bore processing to the withdrawal of the drill from the hole.

On the mentioned graphs (Figure 12) different areas can be identified:

a) the entry area of the point of the twist drill in the material;

b) continuous cutting area;

c) the drill exit area from the material;

d) the friction area between the drill and the wall of the processed hole. After the drill exits on the opposite side of the test item, both the feed force and the cutting torque do not decrease to zero (0) because the material tightens on the drill.

e) on the diagram of the feed force one can observe its evolution in time. Due to the helical blade, the material tends to climb along the tool like a nut. The fixing force of the test item fixture device is required so that the test item is not removed from the device.

f) the exit moment of the drill from the bore using the fast feed can be seen on the OX axis of time on the graph. There is an increase in the negative feed force due to the tendency to lift the semi-finished product from the device.

7 Analysis of the Measuring Method and System from an Economic Perspective

In this research an affordable system with which to perform exploratory tests and allow the development of research in the field of hole processing was designed and developed.

A comparative cost analysis is presented below.

The costs on the market of a ”turnkey” equipment amount to the values of 55000-65000 euros.

I value the conceived, designed and executed device to 3820 euros.

– Purchase of supplies (two dynamometers, an axial bearing, a radial bearing and assembly components) in the amount of 820 euro

- Self-managed activities valued at 3000 euro
- These activities can be split into design activities (10 days x 8 hours x 20 euro/hour = 1600 euro)
- Processing by cutting and mounting (50 hours x 25 eur/hour = 1250 euro)
- Other linked activities.

Therefore, the carried out device is 1/80 if it is related to the actual expenses from the value of an equipment on the market (Or 1/17 from the value of a piece of equipment from the market if it is related to the total expenses with the manufacturing of the device.)

A comparison of advantages and disadvantages is presented in Table 4.

Table 4

Advantages versus disadvantages of the measuring system

DEVICE		COMMERCIAL EQUIPMENT	
Low costs	+	-	High costs
Big size	-	+	Low size
Poor measuring range	-	+	High measuring range
High measuring accuracy	+	-	Low measuring accuracy
Usable only for hole processing (drill, tap, reamer, etc.)	-	+	Can also be used for milling processing
Each dynamometer has separate software and the purchased data must be merged into an Excel file.	-	+	Purchased data automatically ends up in a single file.
The cutting torque results from the multiplication of the force with the arm of the force	-	+	The cutting torque results directly.
Simple control and calibration.	+	-	Hard control and calibration.

Conclusions

Based on the information presented in this paper, the following conclusions can be highlighted:

- A method and a system have been developed for measuring the cutting torque and the feed force through which cutting experiments can be performed at a “cost efficient” level;
- A usable system has been developed to study the phenomena that occur during the processing of holes, threads and data acquisition for further analysis;
- The system can be used for drilling, tapping and reaming processes;
- The presented system and method can reduce the research costs in the field of drilling, tapping and reaming;

- It can be used routinely in the teaching process;
- The system, through its construction, increases the measurement accuracy of the incorporated equipment;
- The system allows the extension of research for several fields.

References

[1] Dutta H, Debnath K, Sarma DK (2020) A study of wire electrical discharge machining of carbon fibre reinforced plastic. In: Shunmugam M, Kanthababu M (eds) Advances in unconventional machining and composites. Lecture notes on multidisci-plinary industrial engineering, Springer, Singapore, pp. 451-460 (Proceedings of the AIMTDR 2018)

[2] Zhang, C.; Zhang, X.; Duan, Y.; Xia, Y.; Ming, Y.; Zhu, Y. Deformation Resistance Performance of Carbon Fiber-Reinforced Plastic Machined by Controlling Drilling Area Temperature below the Glass Transition Temperature. Materials 2021, 14, 1394, https://doi.org/10.3390/ma14061394

[3] Jamkamon, K.; Janmanee, P. Improving Machining Performance for Deep Hole Drilling in the Electrical Discharge Machining Process Using a Step Cylindrical Electrode. Appl. Sci. 2021, 11, 2084

[4] Aamir, M.; Tolouei-Rad, M.; Vafadar, A.; Raja, M. N. A.; Giasin, K. Performance Analysis of Multi-Spindle Drilling of Al2024 with TiN and TiCN Coated Drills Using Experimental and Artificial Neural Networks Technique. Appl. Sci. 2020, 10, 8633

[5] Berzosa, F.; Rubio, E. M.; de Agustina, B.; Davim, J. P. Geometric Optimization of Drills Used to Repair Holes in Magnesium Aeronautical Components. Metals 2020, 10, 1534

[6] Machno, M. Impact of Process Parameters on the Quality of Deep Holes Drilled in Inconel 718 Using EDD. Materials 2019, 12, 2298

[7] Rebora, A. U.; Vernassa, G. Transverse Circular Holes in Cylindrical Tubes Loaded in Traction and in Flexion: A New Ana-lytical Approximation of the Stress Concentration Factor. Materials 2020, 13, 1331

[8] Machno, M. Investigation of the Machinability of the Inconel 718 Superalloy during the Electrical Discharge Drilling Process. Materials 2020, 13, 3392

[9] Dragoi, M.-V.; Rosca, D. M.; Folea, M. F.; Oancea, G. A Fully Symmetrical High Performance Modular Milling Cutter. Symmetry 2021, 13, 496

[10] Luo, H.; Fu, J.; Wu, T.; Chen, N.; Li, H. Numerical Simulation and Experimental Study on the Drilling Process of 7075-t6 Aerospace Aluminum Alloy. Materials 2021, 14, 553

[11] Benedicto, E.; Rubio, E. M.; Carou, D.; Santacruz, C. The Role of Surfactant Structure on the Development of a Sustainable and Effective Cutting Fluid for Machining Titanium Alloys. Metals 2020, 10, 1388

[12] Parasuraman, S.; Elamvazuthi, I.; Kanagaraj, G.; Natarajan, E.; Pugazhenthi, A. Assessments of Process Parameters on Cutting Force and Surface Roughness during Drilling of AA7075/TiB2 In Situ Composite. Materials 2021, 14, 1726

[13] Nur, R.; Yusof, N. M.; Sudin, I.; Nor, F. M.; Kurniawan, D. Determination of Energy Consumption during Turning of Hardened Stainless Steel Using Resultant Cutting Force. Metals 2021, 11, 565

[14] Prakash, C.; Pramanik, A.; Basak, A. K.; Dong, Y.; Debnath, S.; Shankar, S.; Singh, S.; Wu, L. Y.; Zheng, H. Y. Investigating the Efficacy of Adhesive Tape for Drilling Carbon Fibre Reinforced Polymers. Materials 2021, 14, 1699

[15] Liang, Q.; Zhang, D.; Coppola, G.; Mao, J.; Sun, W.; Wang, Y.; Ge, Y. Design and Analysis of a Sensor System for Cutting Force Measurement in Machining Processes. Sensors 2016, 16, 70

[16] Yaldız S., Unsacar F., Saglam H., Isık H., Design, development and testing of a four-component milling dynamometer for the measurement of cutting force and torque Mechanical Systems and Signal Processing. Vol. 21, 2007, pp. 1499-1511, ISSN 0888-3270

[17] Qin, Y., Wang, D. & Yang, Y. Integrated cutting force measurement system based on MEMS sensor for monitoring milling process. Microsyst Technol 26, 2095-2104 (2020)

[18] Ravai-Nagy, S.; Pop, A. B.; Titu, A. M. Determination of Processing Precision of Hole in Industrial Plastic Materials. Polymers 2023, 15, 347

[19] Kistler Group, Brochure: Cutting force measurements in research and development. (2021) https://www.kistler.com/files/download/960-002e.pdf

[20] Totis, G., Adams, O., Sortino, M., Veselovac, D. Klocke, F., Development of an innovative plate dynamometer for advanced milling and drilling applications. Measurement, Vol. 49, 2014, pp 164-181

[21] Qin, Y.; Zhao, Y.; Li, Y.; Zhao, Y.; Wang, P. A., High Performance Torque Sensor for Milling Based on a Piezoresistive MEMS Strain Gauge. Sensors 2016, 16, 513, https://doi.org/10.3390/s16040513

[22] Kistler Group, Brochure: MicroDyn. Multicomponent dynamometer. (9109AA_003-346e-08.18)

An Approximate Model for Determining the Resistance of a Hemispherical Ground Electrode Placed on a Non-homogeneous Truncated Cone

Dragan Vuckovic[1], Dejan Jovanovic[1], Nenad Cvetkovic[1], Dragan Tasic[1], Karolina Kasas-Lazetic[2]

[1] University of Nis, Faculty of electronic engineering, A. Medvedeva 14, 18000 Nis, Serbia; dragan.vuckovic@elfak.ni.ac.rs; dejan.jovanovic@elfak.ni.ac.rs; nenad.cvetkovic@elfak.ni.ac.rs; dragan.tasic@elfak.ni.ac.rs

[2] University of Novi Sad, Faculty of technical sciences, Trg Dositeja Obradovica 6, 21 000 Novi Sad, Serbia; kkasas@uns.ac.rs

Abstract: In this paper we present an approximate analytically oriented approach for determining the resistance of a hemi-spherically-shaped ground electrode placed on the top of a mountain. The mountain is modelled as a non-homogeneous truncated cone consisting of two homogeneous domains with different specific conductivities values. The given procedure extends the existing, previously proposed procedure in which the mountain was modelled as a homogeneous truncated cone. The procedure, proposed in this paper, includes application of the Estimation method, based on idea of finding arithmetic mean of maximal and minimal resistance value. The obtained results are validated and compared with those obtained using the COMSOL program package, based on the finite element method.

Keywords: Estimation method; grounding system; non-homogeneous soil; resistance; hemispherical electrode

1 Introduction

The resistance of the grounding electrode is its most important characteristic and therefore it is of the utmost importance to develop methods for its estimation. This value is influenced by electrode geometry, conductor characteristics, soil structure related to the specific conductivity structure and physical shape of the surrounding ground. The approach of modelling ground as homogeneous half space of flat ground has been used a decades ago [1-2]. The similar is with the procedures which include non-homogeneous ground modeled as multi-layered [3-4], sectoral

[5], semi-spherically [6-7] or semi-cillindrically shaped domain [8]. All these approaches include using of various appropriate numerical methods [9-14]. Very often, facilities having grounding systems as a necessary part are placed at mountains or hills (e.g. antenna towers, wind turbines, etc.) [15-19]. Usually, those terrains are of low specific conductivity. Research dealing with the problem of characterization of grounding systems installed in such places is not so common. An interesting procedure for analysis of a hemispherical ground electrode placed at the top of a hill is proposed in [20]. It is based on the idea of approximating a mountain with a homogeneous truncated cone, while the hemispherical electrode surface is modelled as calotte. These assumptions allow generating approximate analytical expression for the resistance value. In [21], the above-mentioned procedure is improved by assuming current density distribution in two different forms, depending on the observed domain. In this paper, an extension of the approach from [21] is proposed. It offers the possibility of modelling a hill as a non-homogeneous domain consisting of two homogeneous areas, each having different electrical characteristics. The Estimation method application [8, 22-24] is part of the procedure described in this paper. This method is based on the idea of determining the desired approximate value as an arithmetic mean of the upper and lower limits of the interval where the corresponding solution is expected to be. The COMSOL program package, based on the finite element method, is used to validate the results. The obtained results and the data analysis performed during validation suggest that the proposed, relatively simple approach is satisfactory accurate, especially for practical engineering purposes. There is no need for any integration involved in the procedure which reduces the resistance determination to an arithmetic equation. The proposed approach can also be extended to a hill modelled as a truncated cone consisting of three or more domains of different specific conductivity values. Also, the procedure can be used for both, flat or semi-spherically-shaped boundary surface between domains of different specific conductivity values.

2 Problem Description and Solution Procedure

2.1 Problem Description

The hemispherical ground electrode placed at the top of a hill approximated with a truncated cone is observed, as depicted in Figure 1. The cone consists of two homogeneous domains having specific conductivities σ_1 and σ_2. The radius of the electrode is r_0 and cone base radius is r_t. Other geometry parameters from Figure 1 are self-explanatory. In this paper, proposed solution for the structure from Figure 1 is based, as it has been already emphasized, on approaches from

[20] and [21]. However, the chosen model of non-homogeneous truncated cone is more complex and realistic structure related to those from [20] and [21]. One could expect that in general case ground structure is non-homogeneous and using the model from Figure 1 is a good way to take ground non-homogeneity into account.

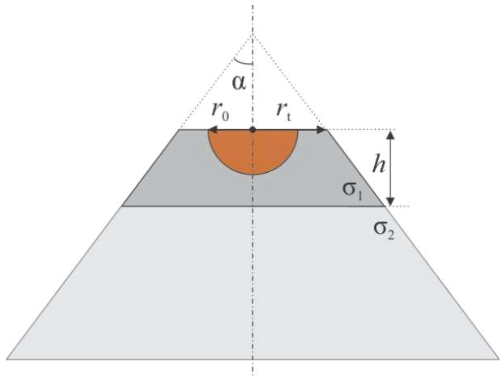


Figure 1
The hemispherical ground electrode at the top of a hill

2.2 Solution Procedure

Since approaches from [20] and [21] are the bases of solving of the problem illustrated in Figure 1, they will be briefly described in this chapter, before presentation of extended procedure applied on model from Figure 1.

2.2.1 Basic Procedure

In [20], the problem of hemispherical grounding electrode having radius r_0, placed at the top of the truncated homogeneous cone of a specific conductivity σ (Figure 2) is analysed and corresponding analytical solution for the low-frequency resistance is derived. A brief description of this procedure is given in this chapter.

The essence of the procedure given in [20] is approximation of the hemispherical electrode by a spherical sector of the radius R_1 (the center of spherical sector is at the fictitious peak of the cone), Figure 2. The parameters d and α are marked in the Figure 2. It is assumed that electrode is fed by the quasi-stationary current I.

From Figure 2 follows:

$$R_1 = r_0 + d \,,\ d = r_0 \cot\alpha \ \text{ and } \ R_1 = r_0\left(1+\cot\alpha\right). \qquad (1)$$

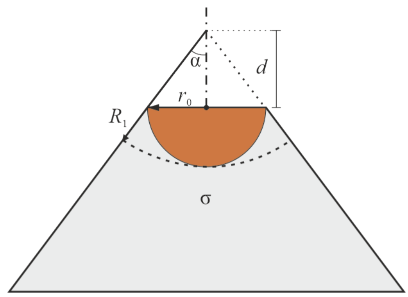


Figure 2
The hemispherical ground electrode at the top of a homogeneous truncated cone

A surface area of a spherical sector of radius R (coordinate orgin for this coordinate coincide with the fictitious pick of the cone) is $S = \Omega R^2$, where $\Omega = 2\pi(1-\cos\alpha)$ is solid angle. Hence, surface area is

$$S = R^2 2\pi(1-\cos\alpha). \tag{2}$$

Symmetry of the approximate geometry has as a consequence that current density depends only on radial coordinate r, i.e

$$\vec{J} = \frac{I}{S}\hat{R} = \frac{I}{R^2 2\pi(1-\cos\alpha)}\hat{r},\ R_1 < R < \infty. \tag{3}$$

Now, using the local form of Ohm's law, $\vec{J} = \sigma\vec{E}$, where $\vec{E}$ is electric field vector, the approximate potential of the electrode surface is

$$\varphi_s = \int_{R_1}^{\infty} \frac{J}{\sigma}\,\mathrm{d}R = \int_{R_1}^{\infty} \frac{I}{2\pi\sigma(1-\cos\alpha)R^2}\,\mathrm{d}R = \frac{I}{2\pi\sigma(1-\cos\alpha)R_1}. \tag{4}$$

Considering that $R_1 = r_0(1+\cot\alpha)$, the electrode surface potential is

$$\varphi_s = \frac{I}{2\pi\sigma(1-\cos\alpha)(1+\cot\alpha)r_0}. \tag{5}$$

Now, resistance of the hemispherical grounding electrode is

$$R_e = \frac{\varphi_s}{I} = \frac{I}{2\pi\sigma(1-\cos\alpha)(1+\cot\alpha)r_0}. \tag{6}$$

2.2.2 Improved Basic Procedure

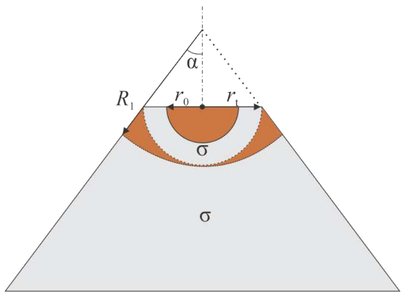


Figure 3
The hemispherical ground electrode at the top of a homogeneous truncated cone

The basic procedure from [20] is extended in [21] and applied to the problem of the hemispherical ground electrode having radius smaller than the radius of the cone basis, Figure 3. The electrode is fed in the center by quasi-stationary current I. The truncated cone has specific conductivity σ, while the radii of the hemispherical electrode and upper cone base are r_0 and r_t respectively ($r_0 < r_t$). This approach includes approximation of the current field with two different expressions.

Firstly, in area defined by $r_0 < r < r_t$, where r is radial coordinate having origin at the center of the hemispherical electrode, the current density vector is assumed as

$$\vec{J}_1 = \frac{I}{2\pi r^2}\hat{r},\ r_0 < r < r_t . \tag{7}$$

As in [20], a spherical sector of the radius R_1 (Figure 3) is introduced into the model. Below this sector, defined by $R_1 < R < \infty$, where R is radial coordinate having origin at the fictitious pick of the cone, for the current density vector the following expression is used

$$\vec{J}_2 = \frac{I}{2\pi\left(1-\cos\alpha\right)R^2}\hat{R},\ R_1 < R < \infty \tag{8}$$

Now, the potential of the electrode can be determined as

$$\varphi_s = \int_{r_0}^{r_t} \frac{J_1}{\sigma}\,\mathrm{d}r + \int_{R_1}^{\infty} \frac{J_2}{\sigma}\,\mathrm{d}R = \int_{r_0}^{r_t} \frac{I}{2\pi\sigma r^2}\,\mathrm{d}r + \int_{R_1}^{\infty} \frac{I}{2\pi\sigma\left(1-\cos\alpha\right)R^2}\,\mathrm{d}R . \tag{9}$$

From expression (9) follows

$$\varphi_s = \frac{1}{2\pi\sigma}\left[\frac{1}{r_0} - \frac{1}{r_t} + \frac{1}{(1-\cos\alpha)R_1}\right]. \tag{10}$$

Consenquently, the resistance of the hemispherical electrode from Figure 3 is

$$R_e = \frac{\varphi_s}{I} = \frac{1}{2\pi\sigma}\left[\frac{1}{r_0} - \frac{1}{r_t} + \frac{1}{(1-\cos\alpha)R_1}\right]. \tag{11}$$

2.2.3 Solution Procedure for the system from Figure 1

In order to approximately determine resistance of the hemispheric electrode from Figure 1, the model depicted in Figure 4 will be analysed. The truncated cone consists of two homogeneous domains having specific conductivities σ_1, i.e. σ_2. The electrode is fed by quasi-stationary current *I.* The radius of the electrode is r_0 and cone basis radius is r_t. The boundary surface between two domains is the spherical sector of the radius R_2.

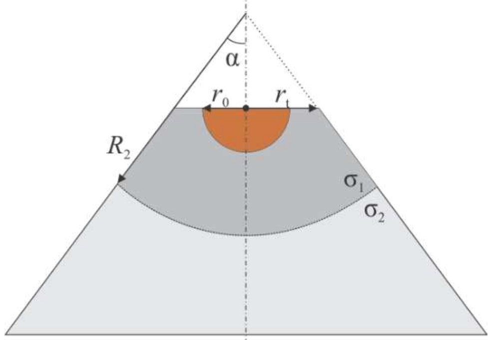


Figure 4
Hill approximated with two-domain truncated cone

A part of the system from Figure 4, consisting of hemispherical electrode and domain having specific conductivity value σ_1, is approximated as in [21] (and explained in 2.2.2), as it is shown in Figure 5. As proposed in [21], described part has been replaced with a hemispherical electrode placed in a shell with boundary of radius R_1. As already written, r_0 and r_t are the distances from the centre of the hemisphere, while R_1 and R_2 are distances from the origin positioned at the fictitious top of the cone. As in [21], the current field in a hemispherical shell around the electrode is assumed as radial, having a current density

$$\vec{J}_1 = \frac{I}{2\pi r^2}\hat{r},\ r_0 < r < r_t. \tag{12}$$

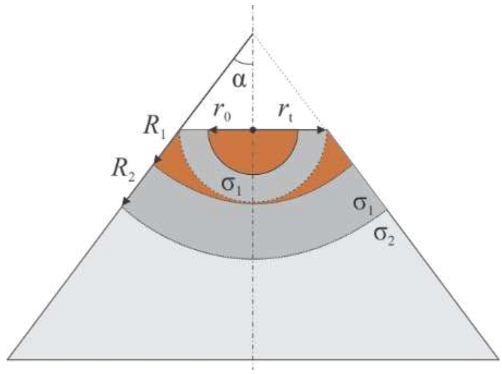


Figure 5
Illustration of the procedure from [21] for part of the system of specific conductivity σ_1

The current field in the rest of the domain of specific conductivity σ_1 is assumed as radial (following the approach from [21]), related to the fictitious top of the cone. It can be characterized by current density vector [21],

$$\vec{J}_2 = \frac{I}{2\pi\left(1-\cos\alpha\right)R^2}\hat{R},\ R_1 < R < R_2\,. \tag{13}$$

In expression (13), R corresponds to the radial distance from the top of the cone, while $\hat{R}$ corresponds to radial ort. The same expression can also be applied for current density vector in the hill domain of specific conductivity σ_2 (defined with $R_2 < R < \infty$), based on the boundary condition for normal component of quasi-stationary current density vector, for $R = R_2$.

Now, the potential of the electrode surface related to the referent point placed on a large distance from the electrode, based on previous assumptions, can be determined as

$$\varphi_s = \int_{r_0}^{r_t}\frac{J_1}{\sigma_1}\mathrm{d}r + \int_{R_1}^{R_2}\frac{J_2}{\sigma_1}\mathrm{d}R + \int_{R_2}^{\infty}\frac{J_2}{\sigma_2}\mathrm{d}R\,. \tag{14}$$

In the previous expressions, dr and dR are differentials of the radial coordinates defined in the text above.

Now, using Ohm's law and equations (12)-(14), the following approximate expression for the electrode potential is obtained.

$$\varphi_s = \int_{r_0}^{r_t}\frac{I}{2\pi\sigma_1 r^2}\mathrm{d}r + \int_{R_1}^{R_2}\frac{I}{2\pi\sigma_1\left(1-\cos\alpha\right)R^2}\mathrm{d}R + \int_{R_2}^{\infty}\frac{I}{2\pi\sigma_1\left(1-\cos\alpha\right)R^2}\mathrm{d}R\,. \tag{15}$$

From (15), the resistance of the hemispherical electrode from Figure 3 is,

$$R_\mathrm{e} = \frac{\varphi_\mathrm{s}}{I} = \frac{1}{2\pi\sigma_1}\left\{\frac{1}{r_0} - \frac{1}{r_\mathrm{t}} + \frac{1}{(1-\cos\alpha)}\left[\frac{1}{R_1} - \frac{1}{R_2}\right]\right\} + \frac{1}{2\pi\sigma_2}\frac{1}{(1-\cos\alpha)R_2}. \quad (16)$$

where, from Figure 5 follows that

$$R_1 = (1+\cot\alpha)r_\mathrm{t}. \quad (17)$$

Finally, using equation (16) and the Estimation method [8, 22-24], it is possible to form an approximate expression for determining the ground electrode's resistance from Figure 1. In Figure 6 are labelled upper (R_{2e}) and lower (R_{2i}) values of boundary surface radii. The approximate resistance of the system is determined as arithmetic mean of the resistance values obtained for Figure 6,

$$R_{2i} = (1+\cot\alpha)r_\mathrm{t} \text{ and } R_{2e} = \frac{h + r_\mathrm{t}\cot\alpha}{\cos\alpha}. \quad (18)$$

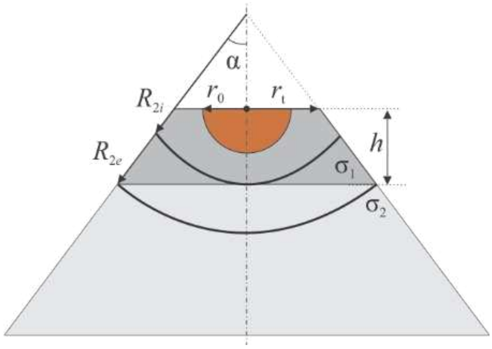


Figure 6
Estimation method application

Now, for $R_2 = R_{2i}$ the value of the electrode resistance $R_\mathrm{e} = R_{\mathrm{e}i}$ is

$$R_{\mathrm{e}i} = \frac{1}{2\pi\sigma_1}\left\{\frac{1}{r_0} - \frac{1}{r_\mathrm{t}} + \frac{1}{(1-\cos\alpha)}\left[\frac{1}{R_1} - \frac{1}{R_{2i}}\right]\right\} + \frac{1}{2\pi\sigma_2}\frac{1}{(1-\cos\alpha)R_{2i}}. \quad (19)$$

For $R_2 = R_{2e}$ obtains electrode resistance $R_\mathrm{e} = R_{\mathrm{e}e}$, i.e.

$$R_{\mathrm{e}e} = \frac{1}{2\pi\sigma_1}\left\{\frac{1}{r_0} - \frac{1}{r_\mathrm{t}} + \frac{1}{(1-\cos\alpha)}\left[\frac{1}{R_1} - \frac{1}{R_{2e}}\right]\right\} + \frac{1}{2\pi\sigma_2}\frac{1}{(1-\cos\alpha)R_{2e}}. \quad (20)$$

The approximate resistance value is obtained as the mean value of R_{2i} and R_{2e}

$$R_{e\,ap} = \frac{R_{ei} + R_{ee}}{2}, \text{ i.e.} \tag{21}$$

$$R_{e\,ap} = \frac{1}{2\pi\sigma_1}\left\{\frac{1}{r_0} - \frac{1}{r_t} + \frac{1}{(1-\cos\alpha)}\left[\frac{1}{R_1} - \frac{R_{2e}+R_{2i}}{2R_{2e}R_i}\right]\right\} + \\ + \frac{1}{2\pi\sigma_2(1-\cos\alpha)}\frac{R_{2e}+R_{2i}}{2R_{2e}R_i}. \tag{22}$$

3 Results

The described method is applied for $\sigma_1 = 0.01\,\text{S/m}$ and $r_0 = 5\,\text{m}$, while the rest of the parameters from Figure 1 take the following values: $\alpha \in \{45^0, 50^0, 55^0, 60^0\}$, $r_t \in \{5\,\text{m}, 10\,\text{m}, 15\,\text{m}\}$, $h \in \{5\,\text{m}, 10\,\text{m}\}$ and $\sigma_2/\sigma_1 \in \{0.5, 5\}$. The values of the parameters have been selected based on [20]-[22]. The obtained results ($R_{e\,ap}$) are validated with the values obtained from the COMSOL program package application (R_e). Number of boundary elements used during the simulation is 18562, while total number of elements is 308473. Electric potential distribution obtained in COMSOL for α=45^0, r_0=5 m, r_t=10 m, h=20 m, σ_1=0.01 S/m and σ_2=0.0001 S/m is shown in Figure 7.

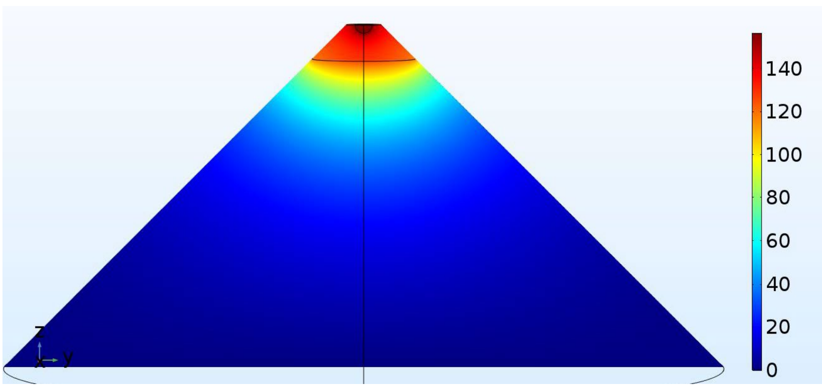


Figure 7

Electric potential distribution for α=45^0, r_0=5 m, r_t=10 m, h=20 m, σ_1=0.01 S/m and σ_2=0.0001 S/m

The results ($R_{e\,ap}$, R_e and relative error) for $\alpha \in \{45^0, 50^0, 55^0, 60^0\}$ are given in Tables 1-4, respectively. Graphics shown in Figures 8-11 correspond to Tables 1-4 respectively and contain relative error versus angle α value.

The maximum deviation of the presented results (Tables 1-4) is 15.7%, while the standard deviation value is 4.541% (based on the results from Tables 1-4), related to the median value of 5.001%.

Table 1

The results for $\alpha=45^0$

r_t[m]	h[m]	σ_2 / σ_1	$R_{e\,ap}$ [Ω]	R_e[Ω]	Relative error [%]
5	10	0.5	85.260	86.979	1.977046
		5	29.602	35.110	15.68705
	20	0.5	72.891	74.731	2.461634
		5	39.497	45.459	13.1162
10	10	0.5	66.276	65.490	1.199309
		5	24.532	26.905	8.816484
	20	0.5	58.545	57.621	1.60396
		5	30.717	33.770	9.040823
15	10	0.5	57.886	56.317	2.785968
		5	24.492	25.113	2.47417
	20	0.5	52.585	50.564	3.9976
		5	28.732	30.455	5.657585

Table 2

The results for $\alpha=50^0$

r_t[m]	h[m]	σ_2 / σ_1	$R_{e\,ap}$ [Ω]	R_e[Ω]	Relative error [%]
5	10	0.5	74.233	76.341	2.761091
		5	27.828	32.805	15.16972
	20	0.5	63.578	65.790	3.361815
		5	36.352	41.843	13.12202
10	10	0.5	60.041	59.690	0.589018
		5	24.222	26.182	7.485269
	20	0.5	53.032	52.567	0.885395
		5	29.830	32.524	8.284733
15	10	0.5	53.575	52.558	1.934003
		5	24.409	24.754	1.3937
	20	0.5	48.602	47.136	3.110419
		5	28.387	29.880	4.99731

Table 3

The results for $\alpha=55^0$

r_t[m]	h[m]	σ_2 / σ_1	$R_{e\,ap}$ [Ω]	R_e[Ω]	Relative error [%]
5	10	0.5	65.655	67.773	3.125576
		5	26.504	30.907	14.2475
	20	0.5	56.400	58.725	3.95919

		5	33.908	38.896	12.82292
10	10	0.5	55.139	54.986	0.278231
		5	24.050	25.599	6.04808
	20	0.5	48.743	48.520	0.459313
		5	29.167	31.494	7.387051
15	10	0.5	50.178	49.445	1.48177
		5	24.397	24.470	0.295781
	20	0.5	45.482	44.358	2.535321
		5	28.154	29.398	4.233505

Table 4
The results for $\alpha=60^0$

r_t[m]	h[m]	σ_2 / σ_1	$R_{e\,ap}$ [Ω]	R_e[Ω]	Relative error [%]
5	10	0.5	58.885	60.850	3.228145
		5	25.540	29.320	12.89244
	20	0.5	50.791	53.017	4.19838
		5	32.015	36.424	12.10466
10	10	0.5	51.231	51.113	0.229494
		5	23.988	25.124	4.522457
	20	0.5	45.358	45.223	0.299508
		5	28.685	30.641	6.383074
15	10	0.5	47.468	46.885	1.243428
		5	24.439	24.241	0.815542
	20	0.5	43.004	42.071	2.216863
		5	28.010	29.004	3.427549

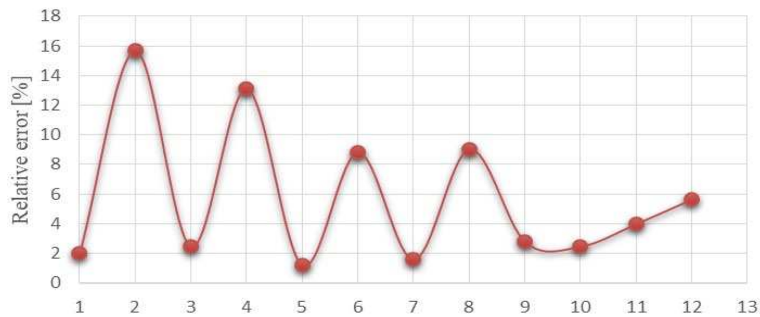


Figure 8
Relative error for different samples when $\alpha=45^0$ (Table 1)

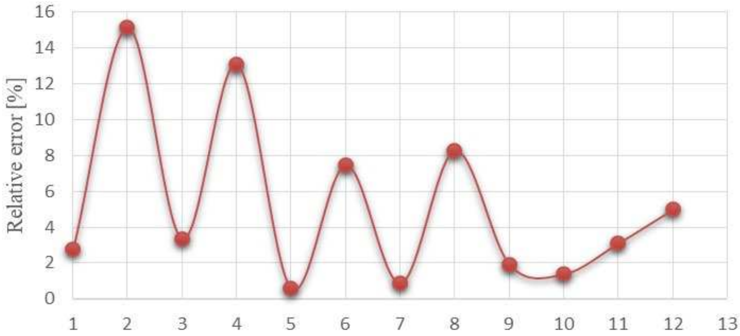


Figure 9
Relative error for different samples when $\alpha=50^0$ (Table 2)

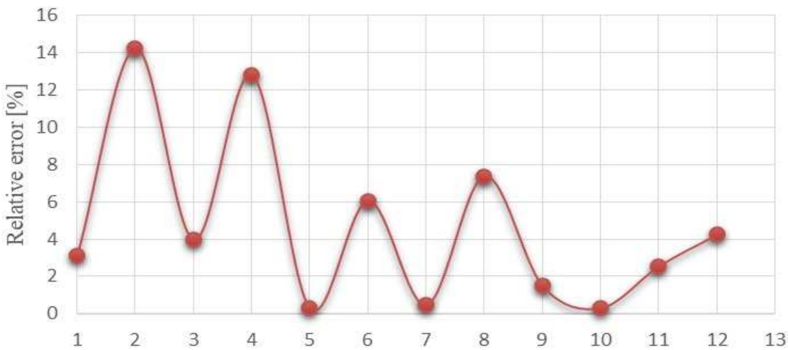


Figure 10
Relative error for different samples when $\alpha=55^0$ (Table 3)

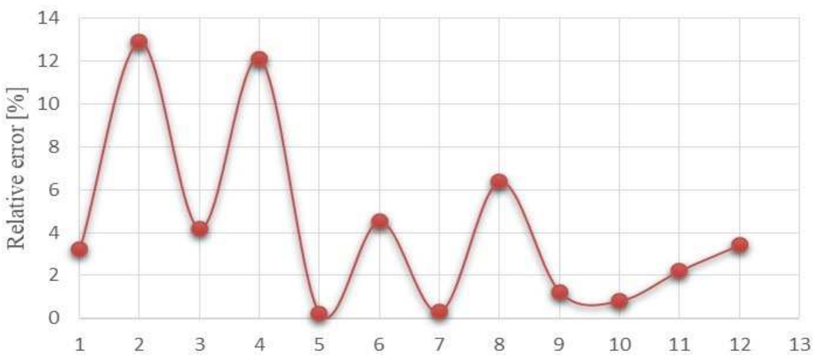


Figure 11
Relative error for different samples when $\alpha=60^0$ (Table 4)

Discussion and Conclusions

An approximate analytical procedure for determining the resistance of a hemispherical electrode, placed on top of a mountain, is presented. The mountain is modelled as truncated cone with two domains. Proposed procedure is a kind of extension of the methods given in [20] and [21]. Based on the obtained results, it can be concluded that the proposed approach is satisfactorily accurate, especially for engineering applications. The method does not involve any type of integration and reduces resistance determination to a simple arithmetic equation. It can also be extended to a hill modelled as a truncated cone consisting of three or more domains of different specific conductivity values.

Acknowledgement

This work has been supported by the Ministry Science, Technological Development and Innovation of the Republic of Serbia.

References

[1] T. Takashima, T. Nakae, R. Ishibahsi: High frequency characteristics of impedances to ground and field distributions of ground electrodes, IEEE Transaction on Power Apparatus and Systems, 1980, 100, pp. 1893-1900

[2] R. J. Heppe: Step potential and body currents near grounds in two-layer earth, IEEE Transactions on Power Apparatus and Systems, 1979, 98, pp. 45-59

[3] P. D. Rancic, L. V. Stefanovic, Dj. R. Djordjevic: A new model of the vertical ground rod in two-layer earth, IEEE Transaction on Magnetics, 1992, 28, pp. 1497-1500

[4] P. D. Rancic, L. V. Stefanovic, Dj. R. Djordjevic: An improved linear grounding system analysis in two-layer earth, IEEE Transaction on Magnetics, 1995, 32, pp. 5179-5187

[5] P. D. Rancic, Z. P. Stajic, Dj. R. Djordjevic, B. S. Tosic: Analysis of linear ground electrodes plaed in vertical three-layer earth, IEEE Transaction on Magnetics, 1996, 32, pp. 1505-1508

[6] N. N. Cvetkovic, P. D. Rancic: A simple model for a numerical determination of electrical characteristics of a pillar foundation grounding system, Engineering Analysis with Boundary Elements, Elsevier, 2009, 33, pp. 555-560

[7] N. N. Cvetkovic, P. D. Rancic: Influence of foundation on pillar grounding system's characteristics, COMPEL: The International Journal for Computation and Mathematics in Electrical and Electronic Engineering, 2009, 28, pp. 471-492

[8] D. D. Vuckovic, N. N. Cvetkovic, M. S. Stojanovic, I. Iatcheva: Approximate model for ground inhomogeneity with rectangular cross-

section: application to analysis of grounding systems. Electrical Engineering, 2018, 100, pp. 75-82

[9] C. Andreeski, D. Mechkaroska: Modelling, forecasting and testing decisions for seasonal time series in tourism, Acta Polytechnica Hungarica, 2020, 17, pp. 149-171

[10] Elena-Lorena Hedrea, Radu-Emil Precup, Raul-Cristian Roman, Emil M. Petriu: Tensor product-based model transformation approach to tower crane systems modeling, Asian Journal of Control, 2021, 23, pp. 1313-1323

[11] C. J. Kwak, K.-C. Ri, S.-I. Kwak, K.-J. Kim, U.-S. Ryu, O.-C. Kwon, N.-H. Kim: Fuzzy modus ponens and tollens based on moving distance in SISO fuzzy system, Romanian Journal of Information Science and Technology, 2021, 24, pp. 257-283

[12] N. Yapici Pehlivan, I. B. Turksen: A novel multiplicative fuzzy regression function with a multiplicative fuzzy clustering algorithm, Romanian Journal of Information Science and Technology, 2021, 24, pp. 79-98

[13] I.-D. Borleaa, R.-E. Precup, A.-B. Borlea: Improvement of K-means cluster quality by post processing resulted clusters, Procedia Computer Science, 2022, 199, pp. 63-70

[14] A.-I. Szedlak-Stineana, R.-E. Precup, E. M. Petriu, R.-C. Roman, E.-L. Hedrea, C.-A. Bojan-Dragos: Extended Kalman filter and Takagi-Sugeno fuzzy observer for a strip winding system, Expert Systems with Applications, 2022, 208, 118215

[15] D. Poljak: Advanced Modeling in Computational Electromagnetic Compatibility; Wiley-Interscience: NJ, USA, 2007

[16] F. Rachidi, M. Rubinstein, A. Smorgonskiy: Lightning Protection of Large Wind-Turbine Blades. In Wind Energy Conversion Systems; S. Muyeen, Eds.; Springer: London, UK, 2012

[17] A. Smorgonskiy, F. Rachidi, M. Rubinstein, G. Diendorfer: On the relation between lightning flash density and terrain elevation. In Proceedings of the International Symposium on Lightning Protection, Belo Horizonte, Brazil, 7-11 October 2013

[18] J. Montanyà, O. Van der Velde, E. R. Willians: Lightning discharges produced by wind turbines, Journal of Geophysical Research Atmospheres, 2014, 119, pp. 1455-1462

[19] L. D. Grcev, A. Kuhar, V. Arnautovski-Toseva, B. Markovski: Evaluation of high-frequency circuit models for horizontal and vertical grounding electrodes, IEEE Transactions on Power Delivery, 2018, 33, pp. 3065-3074

[20] A. Sunjerga, F. Rachidi, M. Rubinstein, D. Poljak: Calculation of the grounding resistance of structures located on elevated terrain, IEEE Transactions on Electromagnetic Compatibility, 2019, 61, pp. 1891-1895

[21] K. B. Tan, H. M. Lu, Y. Zhang, W. C. Zuo: Analysis of the grounding resistance of a hemispheric electrode located on a truncated cone, IEEE Transactions on Electromagnetic Compatibility, 2020, 62, pp. 1361-1363

[22] N. N. Cvetkovic, P. D. Rancic: Model for analyzing a pillar grounding system with cylindrically-shaped concrete foundation, Electromagnetics, 2009, 29, pp. 151-164

[23] P. D. Rancic, N. N. Cvetkovic: Square shaped electrode as pillar grounding system, 9th International Conference on Telecomunications in Modern Satellite, Cable and Broadcasting Services-TELSIKS 2009, October 07-09, 2009, Nis, Serbia, Proceeding of papers, Volume 2, pp. 351-354

[24] N. N. Cvetkovic, D. D. Vuckovic, M. Stojanovic, D. Krstic, D. Tasic: The grounding system of the pillar on the road, 11th International Conference on Telecomunications in Modern Satellite, Cable and Broadcasting Services-TELSIKS 2013, October 16-19, 2013, Nis, Serbia, CD Proceedings pp. 45-48

Controlling the Spindle Speed when Milling Free-Form Surfaces using Ball-End Milling Cutter

Abdulwahab Mgherony, Balázs Mikó

Institute of Material and Manufacturing Science, Óbuda University
Népszínház u. 8, H-1081 Budapest, Hungary
e-mail: mgherony.abdulwhab@uni-obuda.hu; miko.balazs@bgk.uni-obuda.hu

Abstract: Ball-end tools are widely used in industries such as die/mould, automobile, and the aerospace industry. These cutters are specially used when finishing complex three-dimensional sculptured surfaces. However, when using this cutter, the working diameter is different from the nominal diameter and varies from point to point due to the change in surface inclination. Therefore, cutting speed, which is calculated depending on the nominal diameter, is not accurate. The actual cutting speed changes resulting in several defects in the manufactured surface. This article introduces a new algorithm to improve the NC code of a CNC machine. The presented algorithm combines the STL format file with the APT language file to calculate the required spindle speed in order to provide a constant cutting speed during the entire machining process.

Keyword: Ball-end tool; surface roughness; cutting speed; STL format; APT language

Nomenclature

D_c	*tool diameter [mm]*
D_{eff}	*working diameter [mm]*
v_c	*cutting speed [m/min]*
n	*spindle speed [rpm]*
f_z	*feed per tooth [mm]*
z	*number of teeth [-]*
v_f	*feed rate [mm/min]*
A	*feed direction [°]*
a_p	*depth of cut [mm]*
a_e	*width of cut [mm]*
A_{N1}, A_{N2}	*surface inclination angles [rad]*
$\underline{P}_{t\,j}$	*j^{th} tool position point*
$\underline{P}_{tr\,i}$	*i^{th} triangle centre point*
$\underline{N}_{tr\,i}$	*i^{th} triangle normal vector*

Introduction

Free form surfaces have large importance in die and mould industry as a forming surfaces of sheet metal and plastic parts. The machining of forming insert is a piece production and generally done by milling technology or die sinking EDM technology, where the surface of the electrode machined by milling technology. The machining technology has to ensure the accurate surface geometry and fine and constant surface roughness. The surface roughness depends on many factors and circumstances, like the properties of the material, the type of machining, the geometric properties of the cutting tool, the cutting parameters, the tool path and the surface geometry.

Using ball-end tools for semi finishing and finishing free-form surfaces is common. The strong ability to adapt to different surface curvature gives this cutter an advantage over other cutting tools. Since the surface quality is not determined solely by cutting parameters, machining of the complex surfaces using ball-end tool still faces some challenges can be identified as follows: the drastic change in the geometrical features from roughing to finishing, the need to choose different strategies for the tool path, the performance of the CNC, the change in the cutting force and the change in the cutting speed. The current work focuses only on the effect of changing the cutting speed and proposes a method for controlling the spindle speed to ensure a constant cutting speed during the machining process.

During the ball-end milling process of sculpture surfaces, the working diameter of the tool changes. The working diameter value depends on the nominal diameter of the tool, the surface curvature, the feed direction, and the depth of cut. Some researcher studied the change in the working diameter, Mikó and Beňo [1] studied the effect of the inconstant effective diameter on the surface roughness, de Souza et al. [2] studied the effect of the working diameter in cutting force chip removal, and surface roughness, and found out that the surface roughness decreases by increasing the cutting speed. While Krajnik and Kopač [3] introduced a simple calculus of the working diameter when milling a surface by ball-end milling cutter and highlights the feedrate optimization in order to minimize the tool load. During the optimization, the G-code is modified directly.

Surface roughness quality and tool orientation were the focus of other studies. Sadílek et al. [4] compared the quality of the surface roughness in the case of 3-axis and 5-axis milling machines, they pointed out that small deviations of accuracy were deducted in the case of 5-axis milling. However, this deviation is higher in the case of 3-axis machine. Burek et al. [5] investigated the effect of changing tool direction on surface roughness and emphasized the importance of choosing lead and title angles. Gao et al. [6] investigated the impact of the machining inclination angles on the surface quality during machining Ti-6Al-4Vusing ball-end tool. The findings show that, the tool orientation planning has a major impact on the groove quality, and by choosing the appropriate inclination angle, the surface roughness can be reduced.

Many researchers studied the cutting speed and developed several approaches to guarantee a constant cutting speed when milling free-form surfaces. Varvruska et al. [7] proposed a method for milling using a circular cutting edge. In this method, the spindle speed and feed rate were controlled to achieve a constant cutting speed, and they found that under a constant cutting speed, the machining time is lower, and the productivity is higher compared to the traditional method. The working diameter is defined as the distance of the tool axis and the contact point of the surface and the tool. Zhang et al. [8] analysed the effect of the ball-end cutter speed and studied the relationships between the cutting speed and other important parameters such as feed direction and the angel of machining inclination. According to their results, the cutting speed has a significant effect on the machining process. Fan [9] presented in his work a mathematical model to calculate the actual cutting speed on the machined surface in 3-axis ball nose milling, based on the z-map method. Käsemodel et al. [10] presented an algorithm to control the spindle speed to keep the cutting speed constant. By using this algorithm, both the roughness and the machining time were reduced. This algorithm depends on the Grasshopper application at Rhino3D software to get the normal vector.

The aim of the research is to create a post-processing algorithm, which can modify the NC code to achieve spindle speed control and independence from CAD/CAM systems. The current article presents the general concept, the steps of the algorithm and the source of the required data. The geometric data of the surface can read from an STL file, and the tool path data can read from an APT format process file. The two standard file formats ensure the CAD/CAM independency. The work of the algorithm is demonstrated by an example.

1 Spindle Speed Control Concept

In order to the spindle speed control, the working diameter must be calculated point-by-point. Based on the applied geometric model [11], the calculation of the working diameter requires the description of the free form surface, the description of the cutting tool path and the cutting data.

In order to calculate the working diameter of the ball-end milling cutter, the surface date and the tool path data are required. The STL file format and the APT language were used to calculate the angle of the surface inclination at each cutting point, therefore, a brief description of the STL file and other file formats used to represent objects in CAD systems is given, followed by a short definition of the APT language and why it is used in this application.

1.1 STL File Format

To describe the geometric data of a free form surface, there are several possibilities from the native, CAD system-based file formats to the neutral, standard file formats. In the presented application, the method needs only the orientation of the surface normal vectors at different positions, so the different types of tessellation-based formats were analysed.

For rapid prototyping, CAD systems support many types of file formats such as STL, AMF, and OBJ. Each of these formats contain data such as texture, colour, and geometry data about the object to be manufactured. However, the data encoded in each of these types of formats is not the same and varies from file to file. For example, while an STL file only describes the surface geometry of an object without any representation of colour and texture, AMF and OBJ are capable to store information about texture and colour. Moreover, the difference between these formats is not only the type of information stored, but also the way the surface is represented. For example, STL uses planar triangles to describe the object, while AMF uses curved triangle in addition to planar straight triangles, OBJ can use more features such as free-form surfaces and free-form curves. Therefore, since STL file is the most used format and because only surface geometry is important apart from other information, STL has been used in this application [12] [13].

STL stands for Standard Triangle (or Tessellation) Language. This file format is widely used and supported by many CAD systems [14]. It consists of list of triangle facet data. Each facet includes facet normal and three vertices (Figure 1). The facet normal and the vertices are represented in three-dimensional Cartesian coordinate system. Below is an example of an STL file. The file begins with a "*SOLID*" and the name of the part, then a list of facets, and ends with an "*ENDSOLID*" followed by the name of the solid. The STL file includes only surface geometry data, while other information such as material, colour, and texture are not stored in this file format.

```
SOLID PRT0002
 facet normal 0.000e+00 0.000e+00 -1.000e+00
   outer loop
     vertex 0.000e+00 5.000e+01 0.000e+00
     vertex 5.000e+01 0.000e+00 0.000e+00
     vertex 0.000e+00 0.000e+00 0.000e+00
   endloop
 endfacet
...
ENDSOLID PRT0002
```

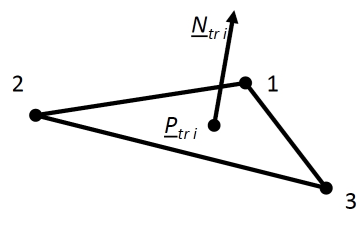


Figure 1
Definition of a triangle in the STL language

1.2 APT File

To get information about the tool path, the CNC code of the milling process should be read. Nowadays, G-code is the most common way to give instructions to numerical control machines. Although it is widely used, G-code is not that standard, every CNC controller has differences in syntax and how G-code is handled. Moreover, the location of the cutter cannot be obtained directly by reading the G-code required in this application. Therefore, the APT language was preferred to obtain the position of the tool during the milling process [15].

The Automated Programming Tool (APT) is a high-level computer programming language developed by the Massachusetts Institute of Technology (MIT) in 1958. It is used to automatically generate the program for the NC machines [16] [17]. Generally, all tool movement posture such as contact point, vector, cutting information and tool path are recorded in the APT. Through the data stored in the APT program, the NC code is generated [18].

As in the presented algorithm, there is a need to obtain the position of the cutter at each point. The structure of the APT language gives this possibility. Below is a part an APT program. The first section contains the main data of the program, like the part name (MFG0002), the machine tool ID (MILL, 1), the unit of the distances (MM), the tool ID (D10 R5) ant the direction of the coordinate system. The $$ marks comments. The next section contains the machining information, like the spindle speed (SPINDL), and the points of the tool motion. The RAPID marks the fast motion, and after the FEDRAT the working motion starts. The GOTO indicates the x,y and z coordinates of the end point of a linear motion.

```
PARTNO / MFG0002
MACHIN / MILL, 1
UNITS / MM
LOADTL / 37 $$-> D10 R5
$$-> CUTTER / 10.000000
$$-> CSYS / 1.000, 0.000, 0.000, -25.000,  $
            0.000, 1.000, 0.000, -25.000,  $
            0.000, 0.000, 1.000, 0.000
SPINDL / RPM, 2000.00, CLW
RAPID
FROM / -20.100, 25.249, 30.000
RAPID
GOTO / -20.100, 25.249, 9.994
FEDRAT / 500.00, MMPM
GOTO / -19.975, 25.249, 9.994
GOTO / -18.996, 25.249, 9.990
...
```

As can be seen from the example above, the APT program starts with some information about the part, machine and unit of measure followed by cutting parameters such as tool diameter and spindle speed. Then, there is a list of tool positions. The x, y and z coordinates of the tool position are marked in every case, not only in changing parameters, like in G coding. This is one of the advantages of the APT description of the tool path. In this example, the process is down milling, so not all the movements are cutting movements. However, the algorithm is capable to distinguish between cutting motions and rapid motions as the location of the cutter is only important when there is contact between the tool and the workpiece.

The other data of the process, such as the diameter of the cutting tool, the nominal cutting speed and nominal feed rate, the depth of cut, the width of cut, and the milling direction can be determined based on the APT file, or they can be user inputs.

1.3 The Algorithm Description

To achieve the aim of this study, an algorithm was developed to control and adjust the spindle speed point-by-point to keep the cutting speed constant. Figure 2 illustrates how this algorithm works to calculate the spindle speed required at each point of the surface.

The algorithm starts by reading the STL file. The STL file is created by the CAD system to describe the surface of a workpiece. The algorithm reads the file and calculates the incentre point for each triangle by the calculation of the average values of the coordinates of the three corner of a triangle. At the end of this stage, the algorithm will store a list of incentre points ($\underline{P}_{tr\ i}$) of the triangles with the corresponding normal vectors ($\underline{N}_{tr\ i}$). For a more efficient work, the triangles, which are not the part of the machined surface, are eliminated. In order to elimination, the normal vector of a triangle is investigated. If the normal vector is parallel with a coordinate axes, except of $Z+$ direction, the triangle is skipped.

Then, the APT file is read by the algorithm. In the current research linear interpolation was used during the generation of the APT file. The APT file contains the coordinates of the tool location. However, these points are not the contact point between the tool and the surface, but they can use for the calculation. The difference between the tool nose point and the tool/surface contact point is small, especially in case of shallow surfaces. The machining points are listed after the "*FEDRAT*" command, while the "*RAPID*" marks the points of the fast motion of the tool. So the different types of the tool location can be separated by the investigation of these key words.

The next step is to find the normal vector at each position of the tool. This can be done by connecting the position of the tool to the corresponding triangle. Therefore, the algorithm calculates the distance between the tool location and all the incentre points and returns the incenter point with the shortest distance to the tool position. That means, the tool is in the triangle which has that incenter point and it has the same normal vector as the triangle.

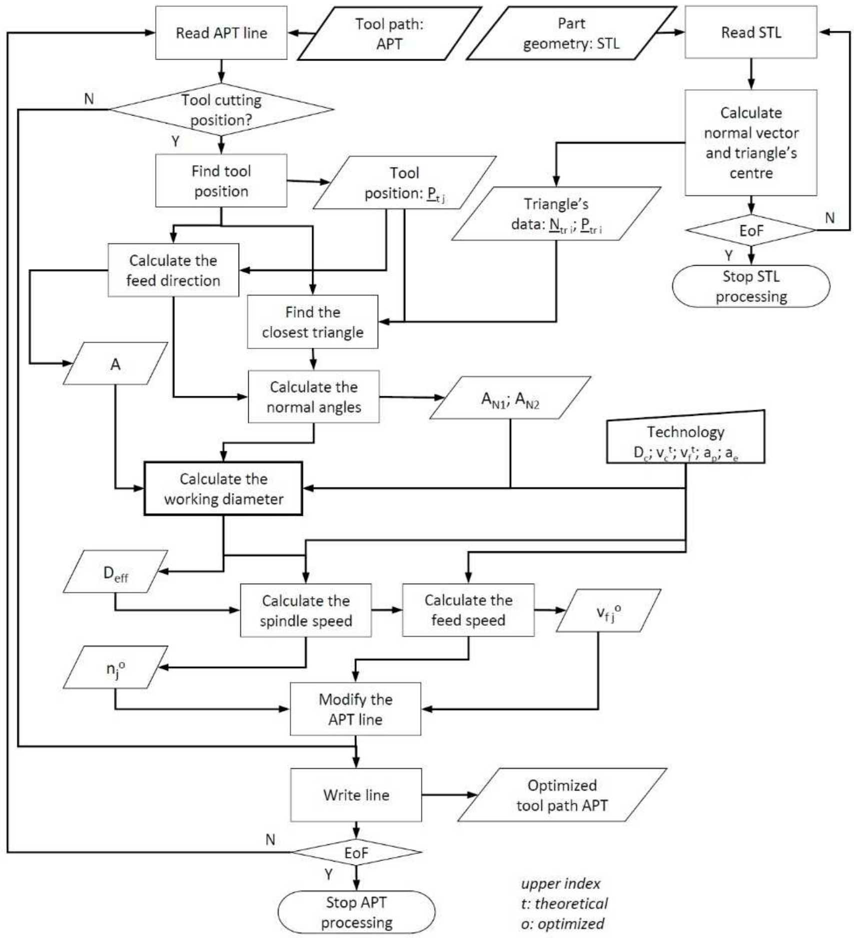


Figure 2

Chart of the algorithm

Depending on the resolution of the STL file, one triangle can be assigned to more tool position. It can be a source of the error in the calculation of the working diameter, but it has just a small effect, and the accuracy can be improved by the resolution of the surface description.

Through finding the normal vector of the surface at each tool position, the surface inclination angles AN_1 and AN_2 can be calculated based on the Figure 3.

$$A_{\mathrm{N1}} = arctg\left(\frac{N_y}{N_x}\right) \tag{1}$$

$$A_{\mathrm{N2}} = arccos(N_z) \tag{2}$$

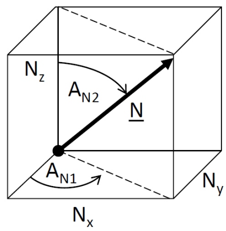


Figure 3
Calculation of the inclination angles (A_{N1}, A_{N2})

If the cutter location is known, the feed direction can be calculated as well (Fig. 4).

$$A = arctg\left(\frac{P_{y\_(j)}-P_{y\_(j-1)}}{P_{x\_(j)}-P_{x\_(j-1)}}\right) \tag{3}$$

The feed direction can be constant, like in the presented example. In this case, there is no need to calculate it point-by-point, it can be an input data.

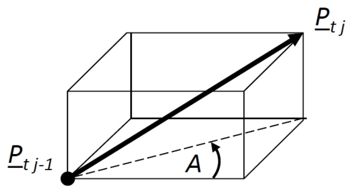


Figure 4
Calculation of the feed direction (A)

These data are necessary to calculate the working diameter at each tool position, which is an important step to calculate the required spindle speed. To calculate the working diameter, a geometric model presented by Mikó and Zentay [11] was used.

In this model, the working diameter is defined as a distance of the tool centre line and an intersection circle. The intersection circle is generated by plane, which has the same normal, as the current surface point and the distance is equal with the depth of cut. The required transformation is driven by the inclination angles A_{N1} and A_{N2}. The direction of the working diameter is defined by the feed direction. The machined part geometry of the free form surface is approximated locally in this model, which harmonizes with the use of the STL description of the surface geometry.

In order to ensure the constant cutting speed, the spindle speed has to be calculated point-by-point. The required spindle speed at each point is calculated by the following formula:

$$n_o = \frac{1000 * v_c}{D_{eff} * \pi} \text{ [1/min]} \tag{4}$$

For constant cutting condition, the feed rate has to be modified parallel with the spindle speed:

$$v_f = f_z * z * n \text{ [mm/min]} \tag{5}$$

2 The Results and the Simulation

A simulation was carried out to study the change in the cutting speed and the required change in the spindle speed to compromise the cutting speed in order to obtain a constant cutting speed using the presented algorithm. Table 1 shows the cutting parameters which were used in this simulation. While Figure 5 shows a simulation of the down milling process in CAM software.

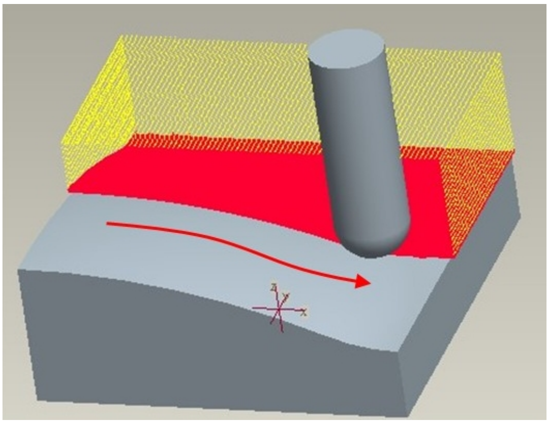


Figure 5
Down milling process in CAM simulation

Table 1
Cutting parameters

Cutting speed	v_c [m/min]	63
Spindle speed	n [rpm]	2000
Feed per tooth	f_z [mm]	0.125
Feed rate	v_f [mm/min]	500
Feed direction	A [°]	0
Depth of cut	a_p [mm]	0.30
Width of cut	a_e [mm]	0.25
Tool diameter	D [mm]	10.0
Number of teeth	z [-]	2

The STL file describes a workpiece with 50 x 50 mm (Figure 6). While the APT file contains the instructions for a down milling process, by reading this file the tool location can be determined by the algorithm as it can be seen in Figure 7. However, the resolution of the STL file is less than the resolution of the tool path description, this error is not significant from the viewpoint of the spindle speed. The coordinates of the tool position are different from the coordinates of the surface points because the tool control point is located on the nose point of the tool. This difference is not so large, than it affects the optimization goals.

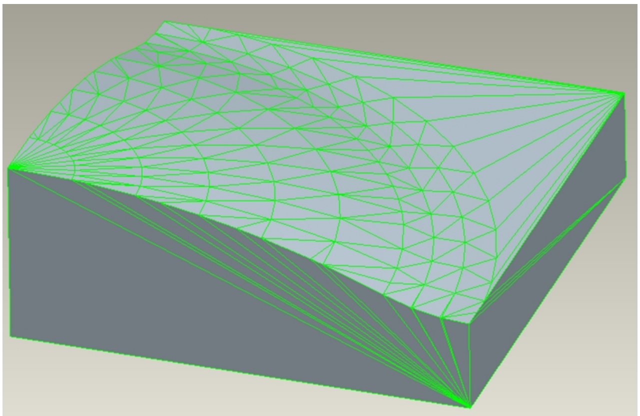


Figure 6
The STL representation of the test surface

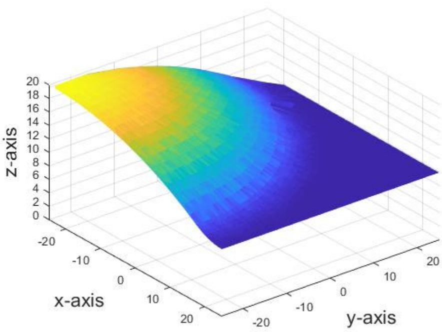


Figure 7
The tool position determined by reading the APT file

The working diameter and cutting speed were calculated in the case of feed direction $A = 0$. Figure 8 and Figure 9 show the change in the working diameter and cutting speed from point to point along the surface of the workpiece. The working diameter of the tool and the actual cutting speed change parallel of course. When the surface normal close to vertical direction (the surface is horizontal), the D_{eff} is small and the cutting speed decreases.

As can be seen in Figure 8, the working diameter drops to the lowest value (about 1.5 mm) when the surface is horizontal, in this section of the workpiece, the axis of the tool is normal on the surface. While in the middle, the diameter of the work increases as the tool engages more into the workpiece. The actual cutting speed changes from point to point along the surface, similar to the working diameter.

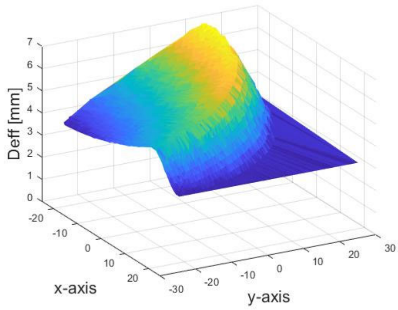


Figure 8
The working diameter

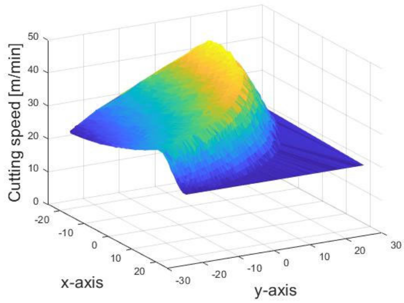


Figure 9
The cutting speed

Similarly, Figure 10 and Figure 11 show the working diameter and cutting speed as a function of the surface slope. When $A_{N1} = 0$ and $A_{N2} = 0$, the working diameter is of low value. The higher the surface slop, the greater the working diameter and more work done by the cutter. This also applies to cutting speed.

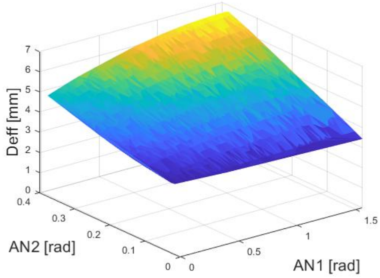


Figure 10
Working diameter versus surface inclination

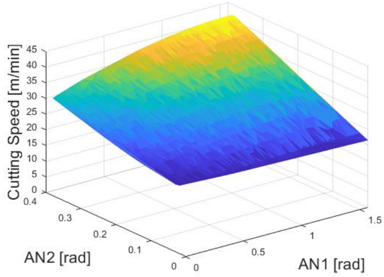


Figure 11
Cutting speed versus surface inclination

Although the cutting parameter including spindle speed is constant, the actual cutting speed changes from point to point during the milling process as a result of the change in tool working diameter which varies with different surface inclination angles. Even though, the nominal cutting speed is 63 [m/min], the actual cutting speed is much lower, and the highest speed can reach 42 [m/min], at some points, while decreases to about 21 [m/min] in other cases.

To keep the cutting speed constant, the spindle speed must be changed and in order to keep the feed per tooth value, the feed rate has to be modified too. The presented algorithm gives a solution to the problem of controlling the spindle speed as well as the feed rate. Figure 12 and Figure 13 show the modified spindle speed and feed rate at each point to keep the cutting speed constant at 63 [m/min].

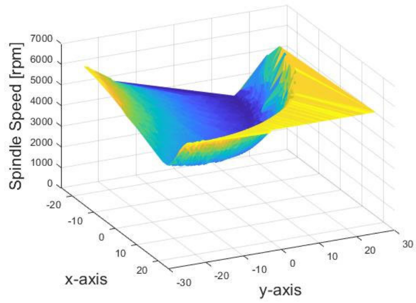


Figure 12
Required spindle speed

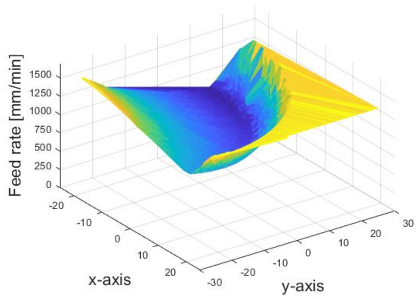


Figure 13
Required feed rate

At some points, an increase in the spindle speed is required to compensate for the decrease in cutting speed that occurs as a result of a decrease in the value of the working diameter. The spindle speed required at these points can reach 5877 [rpm] which is three times the present spindle speed. The lowest spindle speed is 3002 [rpm]. On the other hand, to achieve the same purpose of keeping the cutting speed constant, the feed rate should also be adjusted. The increase in the required feed rate is directly proportional to the spindle speed. Therefore, both parameters have the same character, and in a similar way, the increase in the feed rate reaches 1500 [mm/min], which is three times greater than the nominal feed rate. The feed rate is changed between 750 and 1469 mm/min.

By the mean of the presented algorithm, the cutting speed can be fixed to a specific value. Figure 14 shows the theoretical cutting speed as a function of surface slope. The cutting speed maintains to a constant value (63 [m/min]) regardless of the change of surface angles. In case of the simulation, a small difference can occur only in the cutting speed depending on the resolution of the calculation. The resolution is dependent on the density of the tool path points and the STL surface model.

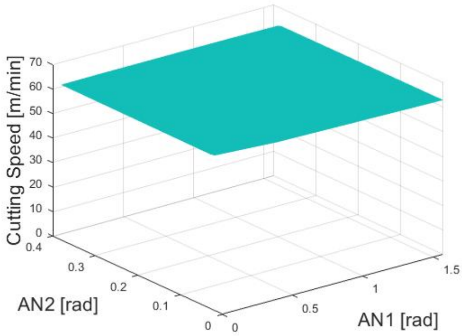


Figure 14
Theoretical cutting speed

Conclusion

The change in the working diameter of the milling cutter is the main problem faced in the manufacture of free-form surfaces using a three-axis ball end mill. Since the working diameter varies from point to point along the surface, the cutting speed also varies, which affects the cutting parameters and as a result the surface quality.

In this article, an algorithm is presented to control the spindle speed to obtain a constant cutting speed. In contrast, to the other approach to adjusting the spindle speed, the presented algorithm does not need any pre-processing and it is independent from any CAD or CAM system. The algorithm uses a new idea to get the normal vector at each point of the surface, which includes reading the STL format file and the NC file written in APT language and making a connection between the cutter locations and the triangles of the STL file. The developed algorithm based on the surface normal from STL and the feed direction from APT can determine the working diameter of the tool, and the spindle speed can be calculated in order to ensure the constant cutting speed.

A simulation was used to demonstrate how the algorithm works. The simulation shows how the working diameter and the cutting speed change during the milling process, due to the change in the angles of the surface slop. Besides, it illustrates how the spindle speed and feed rate should be adjusted to ensure a constant cutting speed through the entire milling process. The presented algorithm can implement inside a CAM system, or it can also work as an independent optimization application.

During future work the experiments have to demonstrate the effect of the constant cutting speed to the surface quality. These experiments include machine workpieces with different feeding directions in the case of constant spindle speed and in the case of constant cutting speed, measure the surface roughness in both cases and compare the results.

As the results show, there can be a large change in the spindle speed. So on one hand, the dynamic properties of the milling machine have to be investigated. The execution of the modified machining program can generate high and dynamic load of the spindle. Based on this result, the algorithm should be modified in order to decrease this dynamic load. On the other hand, the change of the spindle speed can decrease by a new type of the surface milling strategy, which can be developed based on the calculation of the working diameter of the ball-end milling cutter. The next step is to design a tool path searching algorithm, which can generate a new tool path based on the previous one, which considers the geometric information of the surface from the STL file and minimize the changing of the spindle speed.

References

[1] B. Mikó, J. Beňo: Effect of the working diameter to the surface quality in free-form surface milling. Key Engineering Materials 581:372-377 (2014) doi: 10.4028/www.scientific.net/KEM.581.372

[2] A. F. de Souza, A. E. Diniz, A. R. Rodrigues, R. T. Coelho: Investigating the cutting phenomena in free-form milling using a ball-end cutting tool for die and mold manufacturing. The International Journal of Advanced Manufacturing Technology 71(9):1565-1577 (2014) doi: 10.1007/s00170-013-5579-4

[3] P. Krajnik, J. Kopač: Modern machining of die and mold tools. Journal of Materials Processing Technology 157:543-552 (2004) doi: 10.1016/j.jmatprotec.2004.07.146

[4] M. Sadílek, L. Kousal, N. Náprstková, T. Szotkowski, J. Hajnyš: The analysis of accuracy of machined surfaces and surfaces roughness after 3axis and 5axis milling. Manufacturing Technology 18(6):1015-1022 (2018) doi: 10.21062/ujep/217.2018/a/1213-2489/MT/18/6/1015

[5] J. Burek, T. Rydzak, A. Szajna: The influence of the tool axis orientation changes and the direction of tool driving in free surface machining on surface roughness. Mechanik 91(8-9):684-686 (2018) doi: 10.17814/mechanik.2018.8-9.107

[6] P. Gao, X. Wang, Z. Liang, S. Zhang, T. Zhou, P. Yan, L. Jiao: Effects of machining inclination angles on microgroove quality in micro ball end milling of ti-6al-4v. The International Journal of Advanced Manufacturing Technology 92(5):2725-2734 (2017) doi: 10.1007/s00170-017-0305-2

[7] P. Vavruska, P. Zeman, M. Stejskal: Reducing machining time by pre-process control of spindle speed and feed-rate in milling strategies. Procedia CIRP 77:578-581 (2018) doi: 10.1016/j.procir.2018.08.216

[8] A. Zhang, X. Liu, C. Yue, R. Li, S. Y. Liang, L. Wang: Velocity effect sensitivity analysis of ball end milling Ti-6Al-4V. The International Journal of Advanced Manufacturing Technology 118:3963-3982 (2022) doi: 10.1007/s00170-021-08049-4

[9] J. Fan: Cutting speed modelling in ball nose milling applications. The International Journal of Advanced Manufacturing Technology 73(1):161-171 (2014) doi: 10.1007/s00170-014-5672-3

[10] R. B. Käsemodel, A. F. de Souza, R. Voigt, I. Basso, A. R. Rodrigues: CAD/CAM interfaced algorithm reduces cutting force, roughness, and machining time in free-form milling. The International Journal of Advanced Manufacturing Technology 107(3):1883-1900 (2020) doi: 10.1007/s00170-020-05143-x

[11] B. Mikó, P. Zentay: A geometric approach of working tool diameter in 3-axis ball-end milling. The International Journal of Advanced Manufacturing Technology 104(1):1497-1507 (2019) doi: 10.1007/s00170-019-03968-9

[12] C. Iancu: About 3d printing file formats. Annals of the Constantin Brancusi University of Targu Jiu-Letters & Social Sciences Series 2:135-138 (2018)

[13] M. Szilvási-Nagy, Gy. Mátyasi: Analysis of STL files. Mathematical and Computer Modelling 38(7-9):945-960 (2003) doi: 10.1016/S0895-7177(03)90079-3

[14] Y. Chen, C. Ng, Y. Wang: Generation of an STL file from 3d measurement data with user-controlled data reduction. The International Journal of Advanced Manufacturing Technology 15(2):127-131 (1999) doi: 10.1007/s001700050049

[15] S. Brown, C. Drayton, B. Mittman: A description of the APT language. Communications of the ACM 6(11):649-658 (1963) doi: 10.1145/368310.368322

[16] H. Y. Cheng, J. P. Tsai, Y. C. Kao: The development of an apt program interpreter for 5-axis machining. Advanced Materials Research 482:2247-2252 (2012) doi: 10.4028/www.scientific.net/AMR.482-484.2247

[17] D. T. Ross: The design and use of the apt language for automatic programming of numerically controlled machine tools. Proc. 1959 Computer Applications Symposium, Chiergo, pp. 80-99 (1959) doi: 10.1145/960118.808374

[18] D. T. Ross: Origins of the apt language for automatically programmed tools. History of Programming Languages June:279-338 (1978)

Information Basis of Digital Twins: A Quantifiable Metric for Spatio-Temporal Expressivity

Norbert Szántó[1], Ádám Csapó[2], Ildikó Horváth[2]

[1] Department of Automobile Production Technology
Széchenyi István University
Egyetem tér 1, 9026, Győr, Hungary
szanto@sze.hu

[2] University Research and Innovation Center
Óbuda University
Bécsi út 96/B, 1034, Budapest, Hungary
csapo.adam@uni-obuda.hu, horvath.ildiko@uni-obuda.hu

Abstract: This paper focuses on the concept of Digital Twins and the characterization of Digital Twins in terms of the spatio-temporal extent of the information based on which they operate. Following an in-depth literature review, we observe that even when Digital Twins (and variations of this concept) are qualified in terms of 'levels of information integration', strictly quantifiable metrics are rarely applied. To fill this gap, we propose the term 'information basis', which highlights the quantifiable extent of events occurring in space and time which have an influence on the functionality of the Digital Twin. Following a discussion on the implications of this term, we present an example Digital Twin and discuss how its different alternative implementations would fall on different points of the newly introduced information basis spectrum.

Keywords: Digital Reality; Industry 4.0; Digital Twins; Cyber-Physical Systems; Information Basis

1 Introduction

Industry 4.0 (I4.0) is an initiative that has appeared in various forms in the past decades. One recent definition in Xu et. al. states that "*Industry 4.0 represents the current trend of automation technologies in the manufacturing industry, and it mainly includes enabling technologies such as the Cyber-Physical Systems (CPS),* *Internet of Things (IoT) and cloud computing*" [1]. According to Germany Trade Invest, Industry 4.0 represents the technological evolution from embedded systems to Cyber-Physical Systems (CPS): "*In Industry 4.0, embedded systems, semantic machine-to-machine communication, IoT and CPS technologies are integrating the virtual space with the physical world, in addition, a new generation of industrial systems, such as smart factories, is emerging to deal with the complexity of production in Cyber-Physical environment*" [2]. Regardless of the context in which the term is used, there is general consensus that digitization is one of the key pillars underlying this initiative. With the rapid growth of digitization and the emergence of Cyber-Physical Systems (CPS), more and more focus is also placed on the concept of Digital Twins (DTs). In the Industry 4.0 era, the virtual copies of the system are able to interact with the physical counterparts in a bi-directional way, and are capable of replicating and analyzing production systems in real-time. [3]

In this paper, we provide an overview of the use cases underlying DTs, as well as of the various concepts based on which DTs have been commonly described in the scientific and professional literature (Section 2). Based on this short survey, we argue that one aspect in particular – the idea of spatial and temporal scope represented by a DT – is underrepresented in the literature. To compensate for this gap, we propose the concept of 'information basis' of Digital Twins, to highlight the quantifiable extent of events occurring in space and time which have an influence on the functionality of the Digital Twin (Section 3). Finally, we briefly present a digital twin developed in our lab and provide a discussion on how the concept of information basis is relevant to this particular example (Section 4).

2 History, Use Cases and Nomenclature of DTs

One of the first formulations of Digital Twins (DTs) was proposed in 2005 by M.W. Grieves, in the context of his 'Mirrored Space Model' [4], based upon which NASA began to use the term Digital Twin in 2010: "*A digital twin is an integrated multiphysics, multiscale simulation of a vehicle or system that uses the best available physical models, sensor updates, fleet history, etc., to mirror the life of its corresponding flying twin.*" [5]. Many works have built on and expanded upon this concept, up until the most recent past, e.g. [6, 3, 7].

2.1 Categorization of Digital Twins in Terms of Use Case

With the evolution of science and industry, the concept of DTs is no longer restricted to aviation and has made its way into many fields. In the field of production engineering, a common goal today is to apply digital twins toward the creation of intelligent production environments which can make use of novel technologies and novel capabilities, including those of simulation and data-driven adaptability/learning. The role of Digital Twins in this context can be broken into 3 main categories [8]:

1. Digital Twins for monitoring: whether to achieve more detailed analyses, or to support better design and maintenance processes. Examples include monitoring anomalies, fatigue, crack paths or in the physical twin [9, 10, 11, 12, 13, 14, 15]; Monitoring geometric and plastic deformations in the material composition of the physical twin [16]; Modeling reliability of the physical system [17].
2. Digital twins for lifecycle mirroring: such that the target of digital mirroring is the lifecycle of the physical system. Examples include monitoring of the long-term behavior of the system and predicting its performances by taking into account the different synergistic effects of environmental conditions [15, 18, 19, 20]; providing information continuity along the different phases of the lifecycle [21, 22]; virtual commissioning of the system [23, 24]; managing the lifecycle of Internet of Things devices [25].
3. Digital Twins for Decision Making: here, the goal is to support engineering and statistical analyses in the context of the production system. Examples include optimization of system behaviour during design phase [26, 27, 22, 28, 29]; Optimization of product lifecycle, i.e. knowing the past and present states, it is possible to predict and optimize the future performances [30, 24].

It follows from the above that existing strategies for the use of DTs naturally support the operation of the physical system being modeled. This claim is well supported by a recent survey by Liu et. al. , in which these use cases were also complemented by a quantitative assessment of their relative weight in the academic literature [31]. Thus, the authors identified the topics of Concept, Technology, Paradigm / Framework and Application, and went on to show that the prevalence of Application type contributions outweighed all three other topics combined. Further, the authors also qualified existing DT solutions based on the length of their lifecycle, within the categories of Design, Manufacturing, Service, Retirement and Full Lifecycle. The conclusion of this latter analysis was that most DT solutions focus on only a single phase of the lifecycle and only a mere 5% of solutions focus on the entire lifecycle. The most common scope of DTs was in Manufacturing and Service.

2.2 Categorization of Digital Twins in Terms of Directionality of Communication

In their paper published in 2018, Kitzinger et. al. propose to categorize DTs as belonging to one of three categories based especially on the type of communication between the physical and digital components of the DT:

- Digital Model: "A Digital Model is a digital representation of an existing or planned physical object that does not use any form of automated data exchange between the physical object and the digital object" [32].
- Digital Shadow: "If there further exists an automated one-way data flow between the state of an existing physical object and a digital object, one might refer to such a combination as Digital Shadow. A change in state of the physical object leads to a change of state in the digital object, but not vice versa" [32].
- Digital Twin: "The data flows between an existing physical object and a digital object are fully integrated in both directions, one might refer to it as Digital Twin. In such a combination, the digital object might also act as controlling instance of the physical object. A change in state of the physical object directly leads to a change in state of the digital object and vice versa" [32].

Based on the literature, it can be concluded that most integrations are being developed at the lower levels (DM, DS), and that the higher-level DTs (at least within this nomenclature) mostly appear at a conceptual level and not at the level of real use-cases [32].

In the survey mentioned earlier, Liu et. al. also uses these same levels of integration based on the communication between physical and digital parts of the given system [31]. However, while this nomenclature is perfectly valid in case of Cyber-Physical Production Systems (CPPS), it is not necessarily complete when it comes to the characterization of DTs. Thus, in the paper, Liu et. al. define Digital Twins based on a further, more detailed set of characteristics:

- Individualized: "This means that digital twin is in a one-to-one relationship with the individual physical twin. In other words, digital twin is as designed, as manufactured, as used, as maintained as the physical twin" [31].
- High-fidelity: "This means that digital twin can simulate the physical twin's behavior in the virtual space as exactly as possible, which requires multi physics modeling and continuous model updating through the whole lifecycle" [31].
- Real-time: "This means that digital twin responds to its physical twin with relatively low latency, which is made possible by recent advances in mobile communication and IoT technologies" [31].
- Controllable: "This means that changes to the digital twin or to the physical twin controls the other twin. This is the last step that closes the loop between digital and physical twins and realizes digital-physical convergence" [31].

According to Madni et. al., digital twins can be used both in the context of automated and manual processes in production engineering. Data relevant to the state of the processes, including performance, maintenance and health data can be collected and then synchronized with the digital twin. Because of this generality in applicability, and also due to the fact that practitioners often view any kind of digitized version of a physical system as a digital twin, Madni et. al. define four levels of virtual representation – each with its own purpose and scope in the underlying system's lifecycle [33]:

1. Pre-Digital Twin: Virtual system model with emphasis on technology and technical risk mitigation. Physical Twin does not exist. Data Acquisition from Physical Twin not applicable [33].
2. Digital Twin: Virtual system model of the physical twin. Physical Twin does exist. Data Acquisition from Physical Twin applicable, batch updates [33].
3. Adaptive Digital Twin: Virtual system model of the physical twin with adaptive UI. Physical Twin does exist. Data Acquisition from Physical Twin applicable, real-time updates. Machine Learning (Operator Preferences) [33].
4. Intelligent Digital Twin: Virtual system model of the physical twin with adaptive UI and reinforcement learning. Physical Twin does exist. Data Acquisition from Physical Twin is applicable, with batch / real-time updates. Machine Learning (Operator Preferences and System / Environment) [33].

In a broad sense, this categorization is similar to the previously introduced ones in that first and foremost it emphasizes the link between the physical and digital systems which form the twin. However, one novel aspect here is the appearance of machine learning, which brings Madhi et. al. 's conception of DTs closer to the question of what can be achieved using DTs, rather than how the communication is implemented. In this way, prognostications of future states become possible based on past states of the DT.

Nevertheless, we can see that one of the most salient features of DTs in in the literature seems to be the form of communication between the physical and digital components. In a recent survey of industrial applications of DTs, Negri et. al. also places a strong emphasis on this question, in the sense that the work highlights the ability of DTs to both monitoring and optimization services (the latter based on remote actuation). More generally, the many works in the literature emphasize both directions of communication; even as uni-directional communication still remains the norm in most real-world applications [32, 3]. One recent exception is an application presented by Bambura et. al. in 2020 [34]. At the same time, both Bambura et. al. and Redelinghuys et. al. have highlighted the problem of communciation latency, which is especially problematic in the case of bi-directional (closed-loop) communication, causing such DTs to be aptly characterized as close to real-time systems [34, 35].

2.3 Object-Oriented Digital Twins, Process-Oriented Digital Twins and Their Simulation Capabilities

Based on the categorization of DTs, from different nature of use and varying integration levels, the methods needed for the implementation differs greatly. The implementation and referred methods include, but are not limited to simulation (e.g. DES) methods, communication protocols, and core IT technologies in Industry 4.0 [32].

Due to the strong link between DTs and the notion of simulation, it may be worth investigation DTs from this perspective as well. In particular, there may be value in distinguishing between "object-oriented" and "process-oriented" DTs. Relevance to simulation is especially clear in the case of process-oriented DTs, and such DTs can be further qualified based on the following considerations – according to Negri et. al. [8]:

1. Some works consider DTs as representing systems based on which it is possible to create simulations, e.g. [10, 11, 12, 13, 14, 17, 21, 23, 24, 27, 36, 28, 29, 37]

2. Other works consider DTs as a simulation of the underlying system in and of themselves, e.g. [5, 9, 18, 26, 38, 39]

Using (offline) simulation, the performance of a physical system can be compared to the data (expected performance) produced by the digital twin, enabling decisions to be made that better support future interventions or developments to the system. By repeatedly updating the data in the digital twin based on measurements on the physical system, engineers can improve the digital models of the system, leading to both more precise analyses (i.e. recommendations on how to improve the performance of the physical system based on offline computations); as well as to the bootstrapping of further, more precise, even real-time simulation-based predictions on its future states and performance [33]. A possible next step in this evolution is the use of *proactive simulations*, which not only predict future states in real-time, but also carry out automated interventions to keep the physical system within the boundaries of some desired set of states.

Several recent works have highlighted this simulation-oriented aspect of DTs [35]. Schluse and Rossmann proposed the concept of "Experimentable Digital Twins" to highlight the role of simulation in DTs and its application toward the simplification of processes [40]. In all cases, simulations are viable when communication is bi-directional / closed-loop [31, 41].

3 Information Basis of Digital Twins

Based on the literature review in the previous section, we can make several claims.

First, it seems that in some cases the different levels of functionality and integration cannot be separated in a conceptually clear-cut way, due to a lack of quantifiable metrics. For example, the distinction between DTs, Adaptive DTs and Intelligent DTs is mostly a question of degree rather than a matter of clear distinction. Other aspects often highlighted in the literature, such as the notion of high-fidelity or real-time representations can be equally elusive.

In our own analyses, the nomenclature of directionality of communication (levels of integration) still seems viable. Thus, in accordance with the literature, we distinguish between the 3 levels of integration shown in Table 1 and make the claim that in order for a digital representation to be considered as a Digital Twin, there must be an existing physical system that engages in a bi-directional communication with the digital representation and that can thus be controlled by users or other systems via this digital representation.

Table 1
Levels of integration corresponding to DTs and other related concepts (DM and DS). The three levels can be clearly differentiated based on the existence of a physical counterpart and affordance for remote control.

	Physical System	**Data Communication**	**Control**	**Purpose of Model**
Digital Model	-	-	-	design tool
Digital Shadow	exists	uni-directional	-	virtual model of system
Digital Twin	exists	bi-directional	possible	virtual model of system

At the same time, the dimension of integration level leaves aside the question of how detailed a DT is in its representational and communicational capabilities. To address this question, we introduce the notion of ‘***information basis***’, which expresses the *quantifiable extent of events occurring in space and time which have an influence on the functionality of the Digital Twin*.

Table 2
Dimensions of information basis

Spatial	Possibility of virtual augmentation	no	optionally	yes
	Number of source devices	1-5	6-49	50+
	Number of locations for intervention	1	2-5	5+
Temporal	Period of state changes on the source device [s]	0-1	1-60	60+
	Period of physical-to-digital updates [s]	0-1	1-60	60+

The interpretation of this concept is further explained in Table 2.

The spatial dimension of the information basis concept is parameterized on the one hand by the number of devices providing information to the DT, when considering the level of signal sources (rather than higher-level interfaces) as the underlying reality for counting number of devices. At first, based on practical considerations, we might distinguish between 3 categories of spatial information basis in this dimension, ranging from small-sized (1-5 source devices), medium-sized (6-49 source devices) and large-sized (50+ source devices) information bases. It can be conjectured that the larger this number, the more complex the system will be, and the higher degree of automation it will represent. This latter aspect is especially important in the creation and analysis of DTs.

At the same time, another component of the spatial dimension of information basis is the question of how many "locations" are provided by a DT where external interventions are possible. If there is at least one such location, then it is already reasonable to consider the system as a DT. Also, the more locations are provided for such interventions, it can be expected that the closer the interoperability will be between the physical and digital components of the DT. Being able to answer the question of how easily the system model of the physical system can be virtualized is part and parcel of gaining an understanding of the depth of a DT.

The temporal dimension of the information basis concept, in turn, is characterized via the period (inverse frequency) of state changes in the physical component, and the period (inverse frequency) of information transfer between the physical and digital component of the DT. The specific period values of state changes will in general depend, to a large extent, on the use-case scenario for the physical system; hence, we propose to distinguish between DTs in which the temporal information basis is less than 1 second, between 1 and 60 seconds, and over 60 seconds. With respect to the period of information transfer, it is also possible to distinguish between DTs in which the physical and digital components exchange information at a close-to-real-time frequency, or at rate of seconds, hours or even days. Note that true real-time communication is impossible in a physical sense, but from an IT perspective, nanosecond-level frequencies can be regarded as real-time, whereas in the case of DTs in production engineering environments, the nominal update frequency of the communication devices used is specified at the level of milliseconds. At the same time, updates provided at a scale of milliseconds are generally perfectly suitable to practical applications and do not adversely affect the usefulness of DTs.

In addition to the above clusterizaton along the different dimensions of information basis, it can be further remarked that the two parameters considered within the temporal dimension are not completely independent and can have a decisive influence on the quality of the DT. Specifically, the frequency of state changes within the physical component can be as high as any frequency; that is, if the frequency of communication is lower, a large number of intervening physical events will remain unreported between any two communication events, which can lead to increased uncertainty as to whether the DT can reflect the underlying process with reasonable fidelity. As a result, we can make the claim that:

1. If the rate of information transfer is higher than the rate of physical state change, the DT can be considered as a **real-time-updated DT** (note that here, 'real-time' refers not to the technical capabilities of the components of the DT, but rather reflects the temporal communication between the components);

2. If conversely the rate of information transfer is lower than the rate of physical state change, the DT can be considered as a **batch-updated DT**

Thus, the ideal information transmission frequency seems to be the minimal frequency at which all state changes within the system can be communicated to the DT, as it is guaranteed that all state changes can be processed and reflected in the digital representation. Real-time-updated DTs, in turn, hold further potential for deeper inquiry using machine learning, predictive forecasting, carrying out proactive interventions etc.

4 Use-Case Example: Development of a Digital Twin

The Cyber-Physical Production Systems Lab at Széchenyi István University (SZE) is equipped with an automated FESTO didactic production line (ProLog factory), as shown in Figure 1. The production line is characterized by a linear material flow, and automated work stations with various functionalities. Capabilities include loading, form and color based separation, vacuum pick-and-place as well as other processing capabilities; and the system also includes fluid muscle stations, storing stations and commissioning stations - using purely electrical, pneumatic or in some cases electropneumatic actuators as well as inductive, capacitive or color sensors). The work stations are equipped with individual controllers, therefore they are fully operational even as separate entities. User interventions are made possible via modern Human-Machine Interfaces.

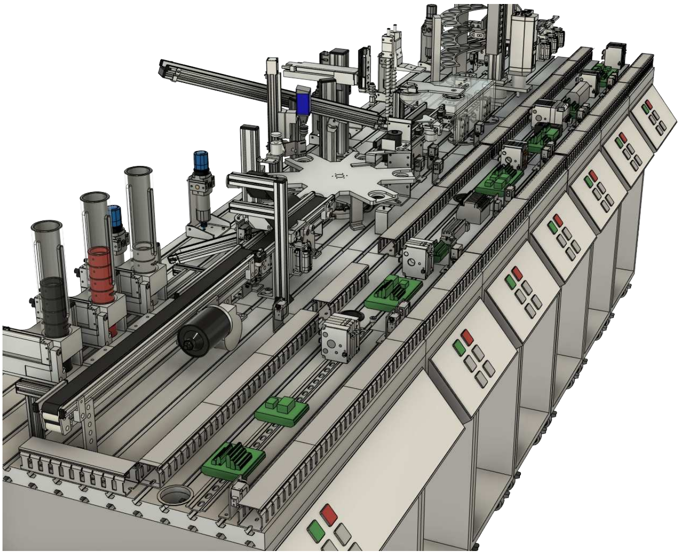


Figure 1
FESTO ProLog factory 3D model

Given the high level of automation provided by this system, it is amenable to the creation of an accompanying digital twin. The most important requirements in this context are:

- Communication protocol should be based on the OPC-UA standard
- The rate of communication should be close to real time
- Communication between physical and digital components is required to be bi-directional
- The DT should be capable of simulation using Discrete Event Simulation (DES) techniques
- The physical process should operate based on decisions made at the level of the DT (true DT intervention).

4.1 System Components and Architecture

The most important part of this development was the communication between the physical and digital comopnents, in terms of quality and speed.

The OPC-UA protocol is a platform-independent and unified standard that enables services to be modelled, connected to and exposed at a high level, without any restrictions on the domain and the way in which its key data types are modeled [42].

From the perspective of the Cyber-Physical System to be created, it was important to ensure that all components are compatible with OPC-UA. The control and external communication of the production line was implemented via an S7-1500 PLC, which supports the OPC-UA protocol as is. The DT, in turn, was implemented via the Siemens Tecnomatix Plant Simulation, which is an object-oriented, discrete event driven process simulation platform. The topology of the physical system (Figure 2) is such that the main control unit has a serial connection to the DT.

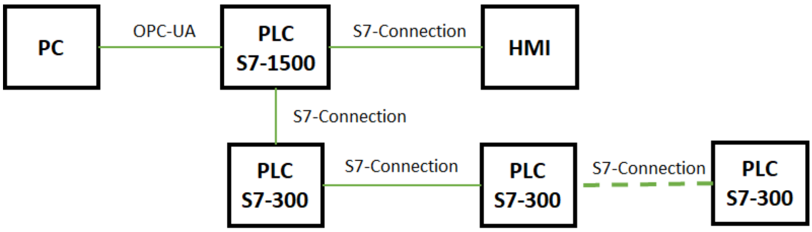


Figure 2
CPS topology

A key property of communication system is that there is a hierarchical (i.e., sender -receiver) relationship between the communicating units. Here in this case, too, the OPC-UA protocol is based on a client-server relationship. In the developed system, the control PLC on the physical side represents the server side, and the DT – i.e. the Plant Simulation software is the client (Table 3).

Table 3
Physical- Digital layer comparison

	Physical layer	**Digital layer**
Hardware	FESTO prolog Factory	PC
Software	TIA Portal	Plant Simulation
OPC-UA	Server	Client

When designing the communication, it was also necessary to create an OPC UA information model. The information model is part of the protocol, the methods and variables defined here perform a continuous data exchange during the server-client connection; a condition for the operation of the digital twin. The information model that can be integrated into the controller was created using the Siemens

SiOME software.

Another important requirement for the developed system was the implementation of close-to-real-time communication. On the server side, the minimum publishing interval was 200 ms, which meant that the DT would be a batch-updated DT. In other words, all previously selected parameters are transmitted at this frequency, regardless of whether their value has changed in the given update cycle. These parameters can be sensor values from the physical system, actuator states, or memory values in the control unit (PLC).

In the Plant Simulation platform, the minimum server read interval on the client side is 1 ms (Read interval). Functionally, this means that the received values are transmitted to the objects after this time, regardless of whether there has been a change (The OPC UA object transmits changed values to Plant Simulation after this time has elapsed.). The read interval is smaller than the server-side send interval, but it does not cause a functional difference; it would be a problem only if the set minimum read interval was greater than the server send interval. However, in our case, two-way communication was also possible, with the minimum server write interval being 10 ms (the rate at which Plant Simulation transmits the values to the OPC server).

Based on the minimum values specified for the server-client connection, it can be concluded that it was not possible to achieve real-time communication, however, near-real-time communication was possible and at a rate that is perfectly suitable in the case of automated production lines.

4.2 Cyber-Physical Processes

The goal of the process-level approach behind the Cyber-Physical System was to ensure that the decisions made in the digital twin could lead to real interventions in the physical system. In the case of a process, an important aspect is in which sub-processes and decision points it is possible to allow external interventions. The possibility of intervening in a process is also available at the security and technical level. The safety approach mainly reflects occupational safety aspects, as the personnel cannot be prepared to make interventions when the system is running in a mode outside of its usual operation at a random time, so the risk of an accident is higher. An important question is whether equipment operating in such a mode can be considered as collaborative workspace equipment and thus subject to collaborative regulation. Another important aspect of intervening in the system can be interpreted at the technical level. In the case of processes that do not operate completely manually, the sensor and control technology is already appearing, the control systems are characterized by the fact that they perform pre-written processes based on the signals coming from the system. In practice, this means that at any point where we intervene in the system, a signal, a state, changes, to which state the control system triggers a predefined response.

Table 4
WS 4 processes

Step	Process
1.	The conveyor belt is started.
2.	The product arrives at the work station.
3.	A stopwatch is triggered at the work location.
4.	The product arrives at the work location.
5.	The conveyor belt stops.
6.	The manipulator moves to pick up the part that is to be installed, the vacuum becomes active.
7.	The part is picked up (if the part is unavailable, the manipulator waits).
8.	The manipulator transports the part into the installation position, the vacuum becomes inactive.
9.	The manipulator returns to its default pose, the stopwatch at the work location becomes inactive.
10.	The conveyor belt starts.
11.	The product leaves the workstation.

However, the intervention may not have been the cause of the reaction, it was just a consequence. Unwanted intervention can result in tool breakage, control error and can interrupt the entire process. Therefore, it is very important to define where, how and for what purpose we intervene in the processes of the physical system.

On the automatic production line affected by the development, the material flow is linear, mainly in the technological and logistical processes, while the built-in part comes from a separate branch. For the vacuum manipulator on the fourth workstation (WS 4), the arrival and storage of the inserted part is implemented by means of a gravity slider. In terms of process, component supply results in an external dependency state, so this is a potential intervention point for DTs. On the fourth workstation (WS 4), the logical sequence of the processes (Table 4) that are predefined and loaded into the PLC (and cannot be changed without a restart) are as follows:

From the point of view of the process, WS 4, Step 7 depends on the condition of the presence of the part. This logical condition is a suitable place for external interventions by the DT, independent of any actuator and sensor, which can be achieved simply by assuming the presence of the component as depending on a result which can be obtained only and exclusively from the digital twin (Figure 3). In addition to the need for completeness, it is important to state that any condition or intervention coded in the PLC is suitable for defining additional conditions, but from a practical point of view, the existence of an external intervention / factor can be justified and transparent at this point. The condition can be represented as an employee confirmation or based on a stochastic process, so we can have an effect on the physical system in some defined way, which can be a simple external effect but also an intervention resulting from historical data of internal processes. Thus, the decision made in the digital twin is part of the realization of the physical system.

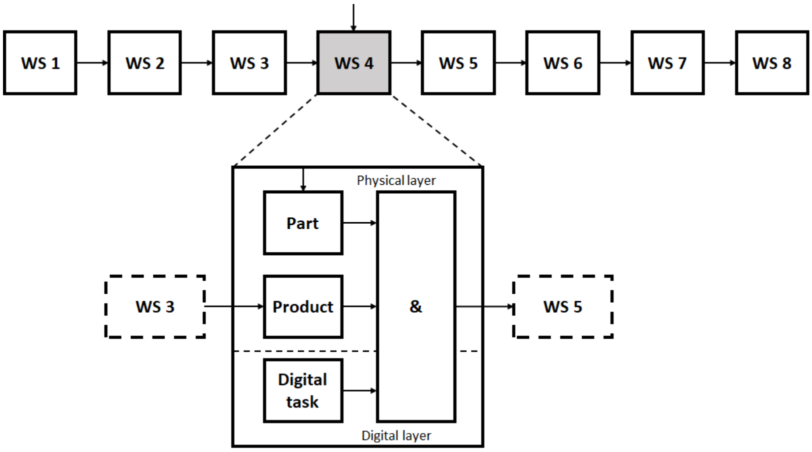


Figure 3
Cyber-physical processes

In the DT, processes operate on the basis of data from a physical process, which are extended with a digital-only task. This task represents a manual operator activity based on stochastic time values in a simulation environment. The digital task starts as a result of Step 7 detailed above, based on parameters from the physical system (Table 5). When the digital task is completed, the conditions in terms of the physical part's availability are met; thus, the required data is transferred to the physical system, and the main process continues with step 8. This method provides the possibility of virtual extension to any physical system, where the virtual extension of the factory and production line in the digital twin is almost limitless in terms of tools and processes.

4.3 Results

As a result of the development, a digital twin with two-way communication based on a close-to-real-time industry standard protocol has been implemented in a simulation environment, where decisions made at the digital level involve real intervention in the physical system. During the development, in line with the preliminary objectives, the most important parameter was to determine the

Table 5
Physical-Digital level processes

Layer	Process
Physical	7. Picking up of physical part.
Digital	7.1 Digital task is started.
Digital	7.2 Digital task is finished.
Physical	8. The manipulator transports the part to the installation location.

place, purpose and way of carrying out interventions in the physical system. The other key task was to establish communication, where on the one hand the communication time had to be realized in the close-to-real-time range using a standard protocol, and on the other hand the two-way information transfer required for the intervention had to be ensured. Once the technical conditions for communication have been determined, it was necessary to delimit the data coming from the physical system in order to transmit only the necessary information to the digital twin to avoid unnecessary communication load. The digital twin process took place in a simulation environment, where events occurred based on data from the PLC. In the environment of event-driven process simulation, this required a different model-building logic during model building, because the logic of the material flow in the software was not provided by the digital twin but by the physical system. In the digital twin, a purely digital process was developed to supplement the physical system (Figure 4), the result of which was fed back into the physical system as a function of events.

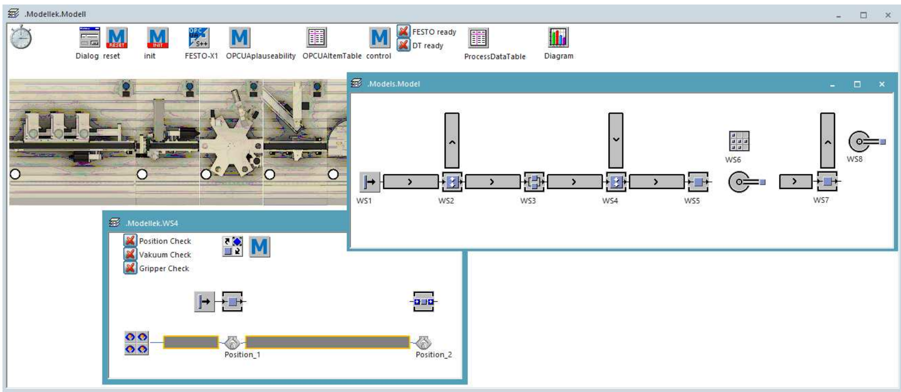


Figure 4
Plant Simulation development environment

The results of this development effort can be evaluated in terms of the scientific literature considered earlier. Based on Kitzinger et. al.'s nomenclature [32], the solution is clearly a Digital Twin. In terms of the integration levels perspective in Liu et. al. [31], the solution is also a Digital Twin, and it can further be characterized by the concepts of individualized, high-fidelity, real-time, controllable DTs. In addition, the solution also fulfills Madni et. al.'s requirements with respect to Digital Twins, but it also is equipped with all the characteristics that are necessary to attain the level of Adaptive and Intelligent DTs. Although the solution does not yet attain this level, the exact status is not a quantifiable metric in any case. In line with the observations in Bambura et. al. and Redelinghuys et. al. [34, 35], it can be shown that the developed system does not achieve real-time communication, but it nevertheless implements close-to-real-time communication. Liu et. al. and Bambura et. al. [31, 34] do emphasize the importance of bi-directional communication, whieh was implemented as part of the solution. When it comes to simulation-based DTs, there are several accepted examples of this concept in the literature, in all cases based on a bi-directional close-to-real-time communication similar to the solution developed here [40, 41]. Based on all of these considerations, the developed solution was a successful implementation of a DT with great potential. It is also characterized by the possibility of carrying out interventions in the physical system using the DT, which is uncommon in the relevant literature.

Based on the notion of information basis, relevant spatial and temporal properties can be used to evaluate the quality and the depth of the developed DT. In analyzing the spatial qualities of the system, it is clear that the virtual augmentation of the system is possible, because the system architecture and the used software support it. The number of source devices is 50+, which includes sensors, actuators, drives, manipulators and robots. The number of locations for intervention has already reached 2 during the development process. Based on the temporal dimension of the information basis, the period of state changes within the source device is between 1 and 60 seconds. The period of physical-to-digital updates is less than 1 second. Thus, it can be said that the rate of information transfer is higher than the rate of physical state change, so the developed DT can be considered as a real-time-updated DT.

Conclusions

With the spread of the I4.0 concept and the appearance of CPS, the need for Digital Twins (DTs) is becoming more and more obvious. In recent years, many theoretical and practical examples of DTs have become known, which typically use and interpret DTs in different contexts. In this paper, we provided an overview of the use cases underlying DTs, as well as of the various concepts based on which DTs have been commonly described in the scientific and professional literature.

As a result of the review, it can be said that there is no uniform interpretation and comparison method. We showed that existing nomenclatures have significant overlaps but are also often difficult to quantify, which led to our motivation to propose the concept of information basis: a quantifiable multi-dimensional metric that expresses the extent of the spatial and temporal domain that has an influence on the behavior of a DT.

In the latter part of the paper, we described a motivating example which was physically and digitally implemented in a laboratory environment.

In a practical example, the close-to-real-time bi-directional connection of the physical system with the DT was implemented, and as a result of the decisions made in the cyber space, the physical process could operate seamlessly, and it was shown to adhere to all major requirements with respect to Digital Twins.

Acknowledgements

Project no. TKP2021-NKTA-48 has been implemented with the support provided by the Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund, financed under the TKP2021-NKTA funding scheme.

References

[1] L. D. Xu, E. L. Xu, and L. Li, “Industry 4.0: state of the art and future trends,” *International Journal of Production Research*, vol. 56, no. 8, pp. 2941–2962, 2018.

[2] Germany Trade Invest, “Industries 4.0 – Smart Manufacturing for the Future,” 2014.

[3] C. Cimino, E. Negri, and L. Fumagalli, “Review of Digital Twin Applications in Manufacturing,” *Computers in Industry*, vol. 113, 2019.

[4] M. W. Grieves, “Product Lifecycle Management: The New Paradigm for Enterprises,” *International Journal of Product Development*, vol. 2, no. 1-2, pp. 71–84, 2005.

[5] M. Shafto, M. Conroy, R. Doyle, E. Glaessgen, C. Kemp, J. LeMoigne, and L. Wang, “DRAFT - Modeling, simulation, information technology & processing roadmap, Technology Area 11,” *National Aeronautics and Space Administration*, pp. 1–32, 2010.

[6] P. Baranyi, Á. Csapó, T. Budai, and G. Wersényi, “Introducing the Concept of Internet of Digital Reality–Part I,” *Acta Polytechnica Hungarica*, vol. 18, no. 7, pp. 225–240, 2021.

[7] C. M. Horváth and P. Korondi, “Supportive Robotic Welding System for Heavy, Small Series Production with Non-Uniform Welding Grooves,” *Acta Polytechnica Hungarica*, vol. 15, no. 8, p. 25, 2018.

[8] E. Negri, L. Fumagalli, and M. Macchia, “A Review of the Roles of Digital Twin in CPS-Based Production Systems,” *Procedia Manufacturing*, vol. 11, pp. 939–948, 2017.

[9] E. Glaessgen and D. Stargel, "The Digital Twin Paradigm for Future NASA and US Air Force Vehicles," in *53rd AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials conference 20th AIAA/ASME/AHS adaptive structures conference 14th AIAA*, p. 1818, 2012.

[10] E. Tuegel, "The Airframe Digital Twin: Some Challenges to Realization," in *53rd AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials conference 20th AIAA/ASME/AHS adaptive structures conference 14th AIAA*, p. 1812, 2012.

[11] K. Reifsnider and P. Majumdar, "Multiphysics Stimulated Simulation Digital Twin Methods for Fleet Management," in *54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, p. 1578, 2013.

[12] Y. Bazilevs, X. Deng, A. Korobenko, F. Lanza di Scalea, M. D. Todd, and S. G. Taylor, "Isogeometric Fatigue Damage Prediction in Large-Scale Composite Structures Driven by Dynamic Sensor Data," *Journal of Applied Mechanics*, vol. 82, no. 9, 2015.

[13] Z. Jin, "Revisiting the Meaning of Requirements," *Journal of computer science and technology*, vol. 21, no. 1, pp. 32–40, 2006.

[14] A. Cerrone, J. Hochhalter, G. Heber, and A. Ingraffea, "On the Effects of Modeling As-Manufactured Geometry: Toward Digital Twin," *International Journal of Aerospace Engineering*, 2014.

[15] E. Fourgeau, E. Gómez, H. Adli, C. Fernandes, and M. Hagege, "System Engineering Workbench for Multi-views Systems Methodology with 3DEXPERIENCE Platform. The Aircraft RADAR Use Case," in *CSDM Asia*, pp. 269–270, 2016.

[16] J. Yang, W. Zhang, and Y. Liu, "Subcycle Fatigue Crack Growth Mechanism Investigation for Aluminum Alloys and Steel," in *54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, p. 1499, 2013.

[17] B. Gockel, A. Tudor, M. Brandyberry, R. Penmetsa, and E. Tuegel, "Challenges with Structural Life Forecasting using Realistic Mission Profiles," in *53rd AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials conference 20th AIAA/ASME/AHS adaptive structures conference 14th AIAA*, p. 1813, 2012.

[18] P. K. Majumdar, M. Faisal Haider, and K. Reifsnider, "Multi-Physics Response of Structural Composites and Framework for Modeling Using Material Geometry," in *54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, p. 1577, 2013.

[19] B. Bielefeldt, J. Hochhalter, and D. Hartl, "Computationally Efficient Analysis of SMA Sensory Particles Embedded in Complex Aerostructures Using a Substructure Approach," in *Smart Materials, Adaptive Structures and Intelligent Systems*, 2015.

[20] M. Grieves and J. Vickers, "Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems," in *Transdisciplinary perspectives on complex systems*, pp. 85–113, Springer, 2017.

[21] R. Rosen, G. von Wichert, G. Lo, and K. D. Bettenhausen, "About The Importance of Autonomy and Digital Twins for the Future of Manufacturing," *IFAC-PapersOnLine*, vol. 48, no. 3, pp. 567–572, 2015. 15th IFAC Symposium onInformation Control Problems inManufacturing.

[22] M. Abramovici, J. C. Göbel, and H. B. Dang, "Semantic Data Management for the Development and Continuous Reconfiguration of Smart Products and Systems," *CIRP Annals*, vol. 65, no. 1, pp. 185–188, 2016.

[23] M. Schluse and J. Rossmann, "From Simulation to Experimentable Digital Twins: Simulation-Based Development And Operation of Complex Technical Systems," in *2016 IEEE International Symposium on Systems Engineering (ISSE)*, pp. 1–6, 2016.

[24] B. Smarslok, A. Culler, and S. Mahadevan, "Error Quantification and Confidence Assessment of Aerothermal Model Predictions for Hypersonic Aircraft," in *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, 2012.

[25] A. Canedo, "Industrial IoT Lifecycle via Digital Twins," in *Proceedings of the Eleventh IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis*, pp. 1–1, 2016.

[26] T. Gabor, L. Belzner, M. Kiermeier, M. T. Beck, and A. Neitz, "A Simulation-Based Architecture for Smart Cyber-Physical Systems," in *2016 IEEE International Conference on Autonomic Computing (ICAC)*, pp. 374–379, 2016.

[27] M. Bajaj, B. Cole, and D. Zwemer, "Architecture To Geometry – Integrating System Models With Mechanical Design," in *AIAA SPACE 2016*, 2016.

[28] J. Ríos, F. M. Morate, M. Oliva, and J. C. Hernández, "Framework to Support the Aircraft Digital Counterpart Concept with an Industrial Design View," *International Journal of Agile Systems and Management*, vol. 9, no. 3, pp. 212–231, 2016.

[29] E. Arisoy, G. Ren, E. Ulu, N. Gecer Ulu, and S. Musuvathy, "A Data-Driven Approach to Predict Hand Positions for Two-Hand Grasps of

Industrial Objects," in *ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 2016.

[30] J. Ríos, J. C. Hernández, M. Oliva, and F. Mas, "Product Avatar as Digital Counterpart of a Physical Individual Product: Literature Review and Implications in an Aircraft," in *Advances in Transdisciplinary Engineering*, vol. 2, pp. 657–666, 2015.

[31] M. Liu, S. Fang, H. Dong, and C. Xu, "Review of Digital Twin About Concepts, Technologies, and Industrial Applications," *Journal of Manufacturing Systems*, vol. 58, pp. 346–361, 2021.

[32] W. Kritzinger, M. Karner, G. Traar, J. Henjes, and W. Sihn, "Digital Twin in Manufacturing: A Categorical Literature Review and Classification," *IFAC PapersOnLine*, vol. 51, no. 11, pp. 1016–1022, 2018.

[33] A. M. Madni, C. C. Madni, and S. D. Lucero, "Leveraging Digital Twin Technology in Model-Based Systems Engineering," *Systems*, vol. 7, no. 7, 2019.

[34] R. Bambura, M. Šolc, M. Dado, and L. Kotek, "Implementation of Digital Twin for Engine Block Manufacturing Processes," *Applied Sciences*, vol. 10, no. 18, 2020.

[35] A. J. H. Redelinghuys, A. H. Basson, and K. Kruger, "A Six-Layer Architecture for the Digital Twin: A Manufacturing Case Study Implementation," *Journal of Intelligent Manufacturing*, vol. 31, pp. 1383–1402, 2020.

[36] G. Grinshpun, T. Cichon, D. Dipika, and J. Rossmann, "From Virtual Testbeds to Real Lightweight Robots: Development and Deployment of Control Algorithms for Soft Robots, with Particular Reference to Industrial Peg-in-Hole Insertion Tasks," in *Proceedings of ISR 2016: 47st International Symposium on Robotics*, pp. 208–214, 2016.

[37] H.-K. Wang, R. Haynes, H.-Z. Huang, L. Dong, and S. N. Atluri, "The Use of High-Performance Fatigue Mechanics and the Extended Kalman / Particle Filters, for Diagnostics and Prognostics of Aircraft Structures," *Computer Modeling in Engineering & Sciences*, vol. 105, no. 1, pp. 1–24, 2015.

[38] M. Shafto, M. Conroy, R. Doyle, E. Glaessgen, C. Kemp, J. LeMoigne, and L. Wang, "Modeling, simulation, information technology & processing roadmap, technology area 11," *National Aeronautics and Space Administration*, pp. 1–38, 2012.

[39] E. M. Kraft, "The Air Force Digital Thread / Digital Twin - Life Cycle Integration and Use of Computational and Experimental Knowledge," in *54th AIAA Aerospace Sciences Meeting*, 2016.

[40] M. Schluse and J. Rossmann, "From Simulation to Experimentable Digital Twins: Simulation-Based Development and Operation of Complex Technical Systems," in *2016 IEEE International Symposium on Systems Engineering (ISSE)*, pp. 1–6, 2016.

[41] A. Protic, Z. Jin, R. Marian, K. Abd, D. Campbell, and J. Chahl, "Implementation of a Bi-Directional Digital Twin for Industry 4 Labs in Academia: A Solution Based on OPC UA," *2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, pp. 979–983, 2020.

[42] OPC Foundation, "OPC Unified Architecture," `https://opcfoundation.org/about/opc-technologies/opc-ua/`, accessed: 08.03.2022.

A Novel Approach of Operation Sequencing Problem in Computer Aided Process Planning

Béla Illés[1], Imre Piller[2], Sándor Radeleczki[2], Tibor Tóth[3†], György Wagner[3]

[1]Logistical Institute, University of Miskolc, Miskolc-Egyetemváros, 3515 Miskolc, Hungary, altilles@uni-miskolc.hu

[2]Mathematical Institute, University of Miskolc, Miskolc-Egyetemváros, 3515 Miskolc, Hungary, matip@uni-miskolc.hu, matradi@uni-miskolc.hu

[3]Institute of Informatics, University of Miskolc, Miskolc-Egyetemváros, 3515 Miskolc, Hungary, toth@ait.iit.uni-miskolc.hu, wagner@iit.uni-miskolc.hu

Abstract: The paper presents a novel method for finding quasi-optimal linear sequences of operations in manufacturing processes. The objective is to produce a part by minimizing the sum of machine, setup and tool change costs. The considered approach is based on a fuzzy clustering process, where the clustering algorithm divides the problem into smaller, locally solvable parts, forming disjoint groups of manufacturing tasks. Whenever the groups are already formed, then the order between their elements is established by using a kind of sorting algorithm, inserting their elements into some chains. Joining these chains according to the induced ordering of the groups we obtain a linear ordering of the tasks of the manufacturing processes. The total cost of this manufacturing process represents a good approximation of the (expected) minimal cost. Some threshold values are used in the fuzzy classification algorithm, which results a more flexible, adaptable tool for similar kind of problems. A numerical example is also presented, which follows and illustrates the mentioned theoretical background and the calculation steps. The article provides a brief overview of a possible GNU/Octave implementation of the proposed method.

Keywords: CAM; CAPP; operation sequencing; linear order; precedence constraint; fuzzy clustering; order-congruence

1 Introduction and Preliminaries

The problem considered in our paper has an interdisciplinary character, being a problem of Computer Aided Process Planning (CAPP), which appears also in Dynamic Process-Planning and in Logistic Information Systems (see [3], [16] or [1]). CAPP is considered the key technology for computer aided design/manufacturing (CAD/CAM) integration. It is also a basic technology in robotics. See for instance a paper about the generic sequencing problem in this domain, where the sequencing of various tasks is accomplished together with making choices on how the tasks are performed [35]. In any CAPP system, selection of the operation sequence is a basic activity for manufacturing a part economically [29]. The essence of operation sequencing is determining in what order to perform a set of selected operations such that the resulting order satisfies the precedence constraints established by both the part and operations [28], to produce a part by minimizing the sum of machine, setup and tool change costs. The combination of different choices and constraints makes process sequencing a combinatorial problem [26].

To determine the optimal sequence, the literature contains various techniques like branch and bound methods, linear programming, dynamics programming (see, e.g., Lin and Wang [17] or Koulamas [18]). In [13] an approach based on particular clusters in a partially ordered set is developed. As the operations ordering problem involves various interdependent constraints, it is difficult to solve this problem using only the mentioned techniques. In fact, it is classified as an (NP)-complete problem. Recently, most works applied metaheuristics for solving process sequencing problems. Bhaskara R. et al. [19], Mojtaba et al. [25], and Li et al. [20] applied genetic algorithms to generate the optimal sequence of manufacturing operations. Foerster et al. [21] and G. Nallakumarasamy and et. al. [27] used simulated annealing for order spread minimization in sequencing cutting patterns which can be considered a kind of generalized Traveling-Salesman Problem (TSP). Li et al. [22] investigated the application of constrained-based tabu search approach for optimization of process plans. Further, the problem was investigated by Krishna et al. [23] using ant colony algorithm and found that the computational time has considerably reduced. Guo et al. [24] applied particle swarm optimization for operation-sequencing problem and concluded that there is still potential for further improvement in computation efficiency and optimality if introducing new operators and characteristics of other algorithms. In [11] a clustering algorithm based on the transition cost between two operations is combined with a precedence constraint checking algorithm to obtain an optimal operation sequencing.

The paper considers a specific case of manufacturing sequencing problems. It assumes that we can fulfil manufacturing steps in sequential order, without any overlapping between the operations, and that there is a transitional cost between the operations. In practice, this cost can mean setup or logistic cost depending on any kind of resource, (for instance time or energy, see e.g. [28]). For simplicity, we regard a cost as an aggregated value of the possible specific cost types. This can describe the forbidden precedencies in the cost matrix by infinite values. The presented algorithm can be successfully applied also to solve problems of logistic nature, replacing the setup and tool changing costs with some shipping costs.

We note that manufacturing and logistics are nowadays inseparable activities [34] due to the globalized market, globalized production and service, and related supplier and distribution activities. The large-scale expansion of the products structure, their dynamic change, the shortening of the product life cycles and the rapid increasing of the number of the suppliers on the global market, all these factors require prompt response, as well as reaching the necessary economic efficiency and improving the market position (see [31] or [32]). The resolution for the challenges of the markets from the part of manufacturing and logistics was provided by the innovations of Information Technology and Manufacturing Technology, moreover, the applications of new software solutions [30]. The Industry 4.0 and Logistics 4.0 make available high scale flexibility in manufacturing and logistics via the benefits of digitalization in the aim of fulfilling the needs of the market (see, e.g. [4], [5], [6]).

In our paper, a heuristic mathematical method is presented for realizing a rapid and economical response of small or medium-sized companies to market needs in the field of manufacturing activity sequencing. By using this method, the following production-specific parameters become manageable:

- minimization of the setup times and hence of the turnaround time of the production process based on specific criteria;
- minimization of material handling work during the production on the basis of specific technological aspects;
- optimization of the production costs incurred during the production process.

The final benefit of the presented method is that it can improve the competitive position of small and medium-sized businesses with relatively simple computational solutions. It is important to emphasize that the main advantage of our method is not the exceptional minimization of the costs, but the flexibility of the program.

The considered concrete issue is the following: We have to insert in a linear order the tasks of a manufacturing process which consists of different manufacturing operations (steps). Our goal is to form an "optimal linear order" of these tasks (cf. [33]). Let $\mathcal{W}$ stand for the n-element set of the necessary tasks and let $a \leq b$ express the fact that the task $b \in \mathcal{W}$ must be preceded in time by the task $a \in \mathcal{W}$. Of course, if $a \neq b$, then we can also use the strict order $a < b$. It is easy to see that in this way we obtain a partially ordered set $(\mathcal{W}, \leq)$. Let denote by $\prec$ the *covering relation* in this ordered set. Thus $a \prec b$ means that $a < b$ holds, and there is no $c \in \mathcal{W}$ satisfying $a < c < b$. A *solution* of the scheduling problem is a linear order of the elements of $\mathcal{W}$ (in the physical time) which extends the initially given partial order $\leq$. This order can be represented by a sequence R: $x_1, x_2, \dots, x_n$, where $x_1, x_2, \dots, x_n$ are the elements of $\mathcal{W}$. As consecutive members x_i and x_{i+1} of this order can be chosen only such operations which can be effectuated one after the other. This means that either $x_i \prec x_{i+1}$ is satisfied, or x_i and x_{i+1} are not comparable with respect to $\leq$ (i.e., their order of execution is indifferent). Conversely, if $a, b \in \mathcal{W}$ are operations which either satisfy $a \prec b$ or they are not comparable, then there exists a linear extension R of the partial order $\leq$ such that

a and b are consecutive elements in this linear order R (see e.g. [14]).

Any allowed transition from the task (manufacturing step) x_i on the task x_j always has a transaction cost (setup cost, or shipping cost) $c_{i,j}$ which can be expressed by a nonnegative real number. In case of the above linear order R let us denote these numbers by $c_{1,2}$, $c_{2,3}, \dots, c_{n-1,n}$, and let $c_{1,1}$ stand for the setup cost necessary to begin the first task x_1. Then the total cost corresponding to this linear extension R is the following:

$$T_R := c_{1,1} + c_{1,2} + c_{2,3} + \cdots + c_{n-1,n} = c_{1,1} + \sum_{i=1}^{n-1} c_{i,i+1}.$$

Our aim is to find a solution with a minimal total cost T_R, or more precisely, to find such a linear extension R of $\leq$ which corresponds approximates the minimal total cost "as much as possible" – and also satisfies at the same time some other necessary technological constraints. (The meaning of the expression "as much as possible" will be specified later.)

2 The Principles of the Proposed Solution

There are more methods for solving such a linear sequencing problem of manufacturing process planning. For instance, the application of restricted traveller salesman algorithm is a common part of these solutions (see, e.g. [13] or [12]). In this case, the costs of the allowed transitions are proportional with the distances between places (shops) visited one after the other. Unfortunately, this method has an exponential time complexity. Another common feature of these methods (which often are the parts of proprietary algorithms) is that they use a clustering process preceding the ordering (see [11], [7] or [13]). For instance, in case of this variant of restricted traveling salesman algorithm, the basic idea can be illustrated as follows.

Suppose that the shops visited by the traveling salesman are grouped in different (neighbouring) cities. For instance, suppose that the places a_1, a_2, ..., a_l are in the city A, the places $a_{l+1}, \dots, a_{l+p}$ are in another city B, ..., etc. In this case, it is worth to visit first the places situated in the same city (say A) and then the places situated in another city (say B), ..., and so on.

Analogously, we will partition the elements of the set $\mathcal{W}$ in some disjoint groups $C_1, C_2, \dots, C_k$ ensuring that two requirements are fulfilled. Namely,

- the groups contain tasks with relatively small setup costs (shipping cost) between them,
- the extension of the (initial) ordering $\leq$ induces a linear ordering between the groups $C_1, C_2, \dots, C_k$.

Of course, some further requirements can be also formulated, for instance, since more workpieces will be processed, it could be advantageous if the processing times of the manufacturing items corresponding to a group do not differ significantly (see, e.g. [3]).

We will see that this grouping can be obtained only in two steps. If the groups are already formed, then the order between their elements will be established by using a kind of sorting algorithm (Algorithm 1). In this way the elements of the groups are inserted in some chains. Joining these chains according to the induced ordering of the groups we will obtain a linear ordering of the elements of $\mathcal{W}$, whose costs represents a good approximation of the (expected) minimal cost.

3 Preliminary Steps

As a first step we generate a square table $\mathcal{M}$ the rows and columns of which are headed by the elements of $\mathcal{W}$. (Hence $i = 1, \dots, n$ and $j = 1, \dots, n$, where $n = |\mathcal{W}|$). If the order of the tasks x_i and x_j is indifferent (they being incomparable with respect to the machining order $\leq$) or x_i has to precede directly x_j-t (i.e. $x_i \prec x_j$ a holds in the partially ordered set $(\mathcal{W}, \leq)$) then in the intersection of the row of x_i and of the column x_j is inserted the setup cost $c_{i,j}$ of the transition from x_i on x_j, i.e. we set $\mathcal{M}(i,j) := c_{i,j}$. In all other cases - i.e., in each case when x_i and x_j cannot be performed consecutively (this being not possible or not allowed,) in position (i,j) is inserted ∞, i.e., we set $\mathcal{M}(i,j) := \infty$. The definition of the partial order $\leq$ and of the covering relation $\prec$ on $\mathcal{W}$ will be explained in Section Implementation.

As it was mentioned before, the order of elements in each final group C_p is established by using a kind of sorting algorithm. For this we will need a minor $Y^{(p)}$ of the matrix $\mathcal{M}$, which is determined by the elements of the group C_p. (Here $1 \leq p \leq k$.) Let $C_p = \{x_{p_1}, x_{p_2}, \dots, x_{p_{m(p)}}\}$, then $Y^{(p)}$ is a matrix of size $m(p) \times m(p)$ (with $m(p) \leq n$), which is defined as follows:

$$Y^{(p)}(i,j) = \mathcal{M}(p_i, p_j), \text{ for all } i,j = 1, \dots, m(p). \tag{1}$$

The elements $Y^{(p)}(i,j)$ of $Y^{(p)}$ are considered as some weighted edges connecting the elements of C_p. We are considering all those possible permutations π of the elements of C_p which represent a linear extension of $\leq$ (more precisely, of the restriction of $\leq$) on the set C_p; The elements $x_{k_1}, x_{k_2}, \dots, x_{k_{m(p)}}$ of a mentioned permutation are selected in the following algorithm (see, e.g. [14]):

Algorithm 1

Set $X_1 = C_p,\ P_1 = (X_1, \leq),\ M_1 = \min(P_1)$,

For $i = 1,2, \dots, m(p) - 1$

Choose $x_{k_i} \in M_i$,

Set $X_{i+1} = X_i - \{x_{k_i}\}$, $P_{i+1} = (X_{i+1}, \leq)$, and $M_{i+1} = \min(P_{i+1})$

End

Notice that $\min(P)$ is a nonempty set of all minimal elements of the poset $(P, \leq)$ and, in view of [14], any such a linear order $x_{k_1} \leq_R x_{k_2} \leq_R \dots \leq_R x_{k_{m(p)}}$ is an extension of $\leq$. The cost $c(\pi)$ of such a permutation π is computed as the sum of their consecutive edges, i.e.: $c(\pi) = \sum_{i=1}^{m(p)-1} \mathcal{M}(k_i, k_{i+1})$.

If an edge with cost ∞ is met in the above sum, then $c(\pi)$ is not computed.

4 The Fuzzy Clustering Process

The reviewed literature shows clearly that we need a flexible clustering method. In what follows, we will use a fuzzy clustering method as a heuristic estimate to guide the search to a good sequence by capturing how "cost-friendly" are the transitions between tasks.

Based on the matrix $\mathcal{M}$, we will implement a fuzzy clustering algorithm, defining first a so-called fuzzy tolerance relation on the set $\mathcal{W}$.

First, we construct a symmetric matrix $\mathcal{K}$ with the same size n, which is defined as follows:

$$\mathcal{K}(i,j) = \mathcal{K}(j,i) := \min\{\mathcal{M}(i,j), \mathcal{M}(j,i)\}, \ i,j = 1, \dots, n. \tag{2}$$

Of course, this matrix $\mathcal{K}$ will contain the value $\mathcal{M}(i,j) = \infty$ exactly in that case when the jobs x_i and x_j are comparable, but neither $x_i \prec x_j$, nor $x_j \prec x_i$ holds (i.e., they cannot be executed consecutively, one just after the other). This is because, in the case when x_i and x_j are not comparable, a finite value is assigned both to $\mathcal{M}(i,j)$ and $\mathcal{M}(j,i)$ by the construction of $\mathcal{M}$.

Now we define another symmetric matrix $\mathcal{H}$ of size $n \times n$, whose elements will be generated from the elements of the matrix $\mathcal{K}$ as follows:

$$\mathcal{H}(i,j) = \begin{cases} 1, & \text{if } i = j; \\ 0, & \text{if } \mathcal{K}(i,j) = \infty; \\ a^{-\mathcal{K}(i,j)}, & \text{otherwise,} \end{cases} \tag{3}$$

where $a > 1$ is a fixed real number. Since the elements of this matrix $\mathcal{H}$ are from the interval $[0,1]$, we can attach to it a binary fuzzy relation ρ defined on the set $\mathcal{W} = \{x_1, x_2, \dots, x_{n,}\}$, whose membership function $\mu_\rho : \mathcal{W} \times \mathcal{W} \longrightarrow [0,1]$ is defined as follows:

$$\mu_\rho(x_i, x_j) := \mathcal{H}(i,j). \tag{4}$$

Since μ_ρ is symmetric and $\mu_\rho(x_i, x_i) = 1,\ i = 1, \dots, n$, this is a so-called fuzzy tolerance relation on $\mathcal{W}$. The determining method of the least fuzzy equivalence containing relation ρ is well known and it can be effectuated in a favourable time. To explain it, we will need some related notions.

4.1 Elements of the Theory of Fuzzy Relations

A *binary fuzzy relation* ρ between the elements of a (crisp) set $X \neq \emptyset$ is defined as a pair (X, μ_ρ), where $\mu_\rho: X \times X \longrightarrow [0,1]$ is a function. The value $\mu_\rho(x, y)$ express the strength" of the fuzzy relation ρ between the elements $x, y \in X$, and μ_ρ is called the *membership function* of ρ.

A *fuzzy tolerance relation* is a binary fuzzy relation $\rho = (X, \mu_\rho)$ satisfying the properties:

$$\mu_\rho(x, x) = 1, \text{ for all } x \in X, \tag{5}$$

$$\mu_\rho(x, y) = \mu_\rho(y, x), \text{ for all } x, y \in X. \tag{6}$$

If ρ in addition satisfies the inequality

$$\mu_\rho(x, z) \geq \min\{\mu_\rho(x, y), \mu_\rho(y, z)\}, \text{ for all } x, y, z \in X, \tag{7}$$

it is called a *fuzzy equivalence* (see [2] or [9]). Let $\alpha \in [0,1]$. An α-cut of a fuzzy relation $\rho = (X, \mu_\rho)$ is a crisp relation ρ_α defined as follows:

$$\rho_\alpha := \{(x, y) \in X \times X | \mu_\rho(x, y) \geq \alpha\}.$$

If $\rho = (X, \mu_\rho)$ is a fuzzy equivalence, then ρ_α is a (crisp) equivalence relation on the set X. Let Π_α stand for the partition induced by ρ_α on X. It is easy to see that for any $\beta \in [0,1]$ with $\beta \geq \alpha$, the partition Π_β is a refinement of Π_α. Therefore, to any sequence $0 \leq \alpha_1 < \alpha_2 < \cdots < \alpha_k \leq 1$ we may attach a nested sequence of partitions $\Pi_{\alpha_1}, \Pi_{\alpha_2}, \dots, \Pi_{\alpha_k}$ of the set X. The *composition* of two fuzzy relations $\varphi = (X, \mu_\varphi)$ and $\theta = (X, \mu_\theta)$ is defined as a fuzzy relation $\varphi \circ \theta = (X, \mu_{\varphi \circ \theta})$, where for all $(x, z) \in X \times X$,

$$\mu_{\varphi \circ \theta}(x, z) := \sup\left\{\min\left(\mu_\varphi(x, y), \mu_\theta(y, z)\right) \middle| y \in X\right\}. \tag{8}$$

The *m-th power* of a fuzzy relation $\rho = (X, \mu_\rho)$ is defined inductively as $\rho^m := \rho^{m-1} \circ \rho$, where $m > 1$ and $\rho^1 := \rho$. The *transitive closure* of a fuzzy relation $\rho = (X, \mu_\rho)$ is the least fuzzy relation $\overline{\rho} = (X, \mu_{\overline{\rho}})$ satisfying the inequality (7) with $\mu_{\overline{\rho}}$ and $\rho \leq \overline{\rho}$. If ρ is fuzzy tolerance on X, then $\overline{\rho}$ always exists and it is a fuzzy equivalence. Now let X be a finite set with n elements and let $\rho = (X, \mu_\rho)$ be a fuzzy tolerance on X. It is well known that in this case there exists a number $k \leq n$ such that $\rho^k = \overline{\rho}$, and we have also $\rho^k = \rho^{k+m}$, for any positive integer m (see, e.g. [9], [15] or [10]).

In our case $X = \mathcal{W} = \{x_1, x_2, \dots, x_{n,}\}$, and we consider the fuzzy tolerance ρ

corresponding to the $n \times n$ type matrix $\mathcal{H}$, defined by the membership function $\mu_\rho: \mathcal{W} \times \mathcal{W} \longrightarrow [0,1]$,

$$\mu_\rho(x_i, x_j) := \mathcal{H}(i,j),\ 1 \leq i,j \leq n. \tag{9}$$

Computing the consecutive powers $\rho^2, \rho^3, \dots, \rho^k$, $(k \leq n)$ until $\rho^k = \rho^{k-1}$ we obtain the transitive closure $\overline{\rho} = \rho^k$ of the fuzzy tolerance ρ. We denote $\Phi = \overline{\rho} = \rho^k$. Clearly, in our case, the fuzzy equivalence relation Φ is obtained by calculating the consecutive powers $\mathcal{H}^1, \mathcal{H}^2, \dots, \mathcal{H}^k$, of the similarity matrix, where $\mathcal{H}^1 := \mathcal{H}$, and where the usual matrix product is replaced with the following "max-min product" (cf. [15]):

$$\mathcal{H}^{l+1}(i,j) = (\mathcal{H}^l \cdot \mathcal{H})(i,j) := \max_{1 \leq p \leq n} \{\min(\mathcal{H}^l(i,p), \mathcal{H}(p,j))\}. \tag{10}$$

This calculation is performed until repeating powers appear, i.e., $\mathcal{H}^k = \mathcal{H}^{k+1}$. The obtained fuzzy equivalence Φ is defined by the matrix $\mathcal{H}^k$, i.e., its membership function is:

$$\mu_\Phi(x_i, x_j) := \mathcal{H}^k(i,j),\ 1 \leq i,j \leq n. \tag{11}$$

In fact, Φ is expressing how closed are the tasks x_i and x_j, in other words, how "cost-friendly" is the transition from x_i to x_j. Clearly, any cut Φ_α of Φ corresponding to a number $\alpha \in [0,1]$ is a crisp equivalence on $\mathcal{W}$.

5 The Second Step: Generating a Linear Order-Congruence

Our aim is to insert the machining tasks (i.e., the elements of $\mathcal{W}$) into disjoint groups $C_1, C_2, \dots, C_k$, where inside these groups they might be ordered in several ways. However, the groups follow each other in a well-determined linear order, which is "compatible" with the initial partial order $\leq$. In other words, the order $\leq$ induces between the groups a linear order. Of course, in this case $\{C_1, C_2, \dots, C_k\}$ being a partition of the set $\mathcal{W}$, determines a unique equivalence relation ρ on $\mathcal{W}$-n, that equivalence whose classes are $C_1, C_2, \dots, C_k$. Hence the factor set of ρ is $\mathcal{W}/\rho = \{C_1, C_2, \dots, C_k\}$. Let us consider now that map $\varphi: \mathcal{W} \longrightarrow \mathcal{W}/\rho$, which maps any $a \in \mathcal{W}$ into the class C_a containing a. The mathematical meaning of the requirement that the order $\leq_\rho$ defined between the classes $C_1, C_2, \dots, C_k$ is "compatible" with the initial partial order $\leq$, is the condition that for any $a, b \in \mathcal{W}$ with $a \leq b$, the relation $C_a \leq_\rho C_b$ also holds, that is, $a \leq b$ implies $\varphi(a) \leq_\rho \varphi(b)$. An equivalence relation which satisfies such a property is called an order-congruence of the partially ordered set - its exact definition (see [8]) is given below:

Definition 1 (i) Let $(P, \leq)$ be a partially ordered set and $\rho \subseteq P \times P$ an equivalence on P. Denote by $[x]_\rho$ the class of ρ containing $x \in P$. If there exists a partially order $\leq_\rho$ on the factor set P/ρ with the property that the mapping a $\varphi: P \longrightarrow P/\rho$, $\varphi(x) = [x]_\rho$ is order-preserving with respect to $\leq_\rho$, that is, for any $a, b \in P$, $a \leq b$ implies $\varphi(a) \leq_\rho \varphi(b)$, then ρ is called an *order-congruence* and $\leq_\rho$ is said to be the *partial order induced on the factor set* P/ρ. In this case, $a\rho b \Leftrightarrow \varphi(a) = \varphi(b)$, in other words, the equivalence ρ is the same as the "kernel equivalence" of the function φ (i.e., $\rho = \mathrm{Ker}\varphi$).

(ii) The order congruence ρ is called *linear* if $\leq_\rho$ is a linear order on P/ρ.

In [8] is proved also that any block $[x]_\rho$ of an order congruence is convex, i.e., $a, b \in [x]_\rho$ and $a \leq c \leq b$ for any $c \in P$ imply $c \in [x]_\rho$.

In view of this definition, we are looking for linear order-congruences of the poset $(\mathcal{W}, \leq)$ of machining tasks. We have elaborated a method for generating them. This generalizes a method of László Eszes. The essence of the algorithm in [13] is the choosing of a set of minimal elements M_1. Then, a set M_2 of minimal elements is selected from $(\mathcal{W} \setminus M_1, \leq)$, and so on, until the set $\mathcal{W}$ is exhausted. It is not hard to prove that the equivalence corresponding to the partition $\mathcal{W} = M_1 \cup \ldots \cup M_k$ generated by this method is an order-congruence of $(\mathcal{W}, \leq)$. An *order ideal* of a partially ordered set $(P, \leq)$ is a nonempty subset $A \subseteq P$ with the property that for any $a \in A$ and $x \in P$, $x \leq a$ implies $x \in A$. Notice, that any order-ideal is a convex subset. The previous algorithm is modified by using the fuzzy equivalence obtained in the first step of our method as follows:

First, we consider a cut Φ_{α_1} corresponding to a number $\alpha_1 \in [0,1]$ of our fuzzy equivalence Φ, and take the $\Phi_{\alpha_1}[m_1]$ class of a minimal element m_1 from the poset $(\mathcal{W}, \leq)$. As set B_1 will be chosen such an order ideal of the poset $(\mathcal{W}, \leq)$, which contains m_1 and it is included in $\Phi_{\alpha_1}[m_1]$, i.e., $m_1 \in B_1$. The number α_1 and the class $\Phi_{\alpha_1}[m_1]$ is chosen in that way, to obtain a set B_1 as large as possible, but at least with three elements.

We consider a linear order of the elements of B_1 and denote its maximal element by m_1^*, and its predecessor element by m_1' ($m_1' \prec m_1^*$). Next, we restrict the fuzzy equivalence relation to $\mathcal{W}_1 = (\mathcal{W} \setminus B_1) \cup \{m_1^*\}$. We try to find a new α_2 value and a $\Phi_{\alpha_2}[m_1^*]$ class of $(\mathcal{W}_1, \leq)$, which provides the largest $B_2 \subseteq \Phi_{\alpha_2}[m_1^*]$. In the next step, we consider the set $\mathcal{W}_2 = \big(\mathcal{W} \setminus (B_1 \cup B_2)\big) \cup \{m_2^*\}$. We continue this procedure, while the set $\mathcal{W}$ has any element. It results the $C_1 = B_1 \setminus \{m_1^*\}, C_2 = B_2 \setminus \{m_2^*\}, \ldots, C_k = B_k$ disjoint clusters. Clearly, $\mathcal{W} = C_1 \cup C_2 \cup \cdots \cup C_k$ is a partition of $\mathcal{W}$. In Proposition 3, we will show that the sets C_i, $i = 1, \ldots, k$ are the blocks of an order-congruence of $(\mathcal{W}, \leq)$. As we mentioned before, we have determined the order of elements within the sets B_i by computing the weights of all possible linear orders (extending the initial partial order) and choosing the optimal one, and thereafter we joined the resulted chains. This means that, we join the *maximal element* m_i' of $C_i = B_i \setminus \{m_i^*\}$ (in step $i \in \{1, \ldots, k-1\}$), with the *minimal element* $m_{i+1}^{\#}$ of the chain C_{i+1}.

Clearly, in this way we obtain a linear order. Observe that choosing an order ideal B_j from the fuzzy equivalence class $\Phi[m_j]$, and ordering the elements of the group B_j in a linear order, can be done by the same algorithm, namely a variant of the Algorithm 1, as follows:

Algorithm 1a

Set $X_1 = \mathcal{W}_{j-1}$, $P_1 = (X_1, \leq)$, $M_1 = \{m_j\}$,

For $i = 1,2,\dots,|B_j|$ do

If $M_i = \emptyset$ print $x_1 < \cdots < x_{i-1}$ and Stop

Else Choose $x_i \in M_i$,

Set $X_{i+1} = X_i - \{x_i\}$, $P_{i+1} = (X_{i+1}, \leq)$, and

$M_{i+1} = \min(P_{i+1}) \cap \Phi[m_j]$

End

We note that whenever B_j has a small, limited size (e.g., $s = |B_j| \leq 6$), the linear orderings of its elements can be obtained by a simplified variant of this method: First, all the permutations of the elements of B_j are generated. Then we select those permutations $\pi = (x_{k_1}, x_{k_2}, \dots, x_{k_s})$ which have the property that each element x_{k_i}, $1 \leq i < s$ of them is a minimal one in the subposet $(P_i, \leq)$, where $P_i = \{x_{k_i}, x_{k_{i+1}}, \dots, x_{k_s}\}$. (Finally, as optimal permutation is chosen that one with the minimal cost.)

Algorithm 1b

Set $\pi = (x_{k_1}, x_{k_2}, \dots, x_{k_s}), P_1 = (\{x_1, x_2, \dots, x_s\}, \leq), d = \text{true}$

For $i = 1,2,\dots,|s| - 1$ do

If $x_{k_i} \in \min\,(P_i)$

Set $X_{i+1} = X_i - \{x_{k_i}\}$, $P_{i+1} = (X_{i+1}, \leq)$

Else Set $d = \text{false}$

End

Since the elements $x_1, \dots, x_{i-1}$ within the group B_j are generated in fact by Algorithm 1, the order $x_1 < \cdots < x_{i-1}$ always is a linear order which extends the restriction of $\leq$ to $\{x_1, \dots, x_{i-1}\}$. Let us observe that the final group B_j is an order ideal in the partially ordered set $(\mathcal{W}_{j-1}, \leq)$:

Indeed, let $B_j = \{x_1, \dots, x_s\}$, $s = |B_j|$, and take $b \in \mathcal{W}_{j-1}$ such that $b < x_k$, for some $x_k \in B_j$. Since x_k is a minimal element in the ordered set $P_k = \mathcal{W}_{j-1} \setminus \{x_1, \dots, x_{k-1}\}$, this is possible only in the case $b \in \{x_1, \dots, x_{k-1}\} \subseteq B_j$. Hence B_j is an order ideal in $(\mathcal{W}_{j-1}, \leq)$.

Now, we have to prove only that this linear order R induced by our method on $\mathcal{W}$ is an extension of the partial order $\leq$ defined initially on $\mathcal{W}$. This will be done in Proposition 3(ii). Observe, that by definition, $m_i^* R m_{i+1}^{\#}$ $i = 1, \dots, k-1$ always holds. We also claim that any element m_i^* *is maximal in the poset* $(B_i, \leq)$*, and it is minimal in the poset* $(\mathcal{W}_i, \leq)$*, where* $\mathcal{W}_i = \left(\mathcal{W} \setminus \left(\bigcup_{k=1}^{i} B_k\right)\right) \cup \{m_i^*\}$. Thus $\Phi_{\alpha_i}[m_i^*]$ is always an equivalence class of a minimal element in $\mathcal{W}_i$.

Indeed, suppose by contradiction that there is an element $b \in B_i$ such that $m_i^* < b$. Since Algorithm 1a preserves the existing order, we get $m_i^* R b$, in contradiction to the maximality of m_i^* with respect to R in B_i. Similarly, assume that there exists an element $a \in \mathcal{W}_i$ such that $a < m_i^*$. Since $m_i^* \in B_i$, $a \in \mathcal{W}_i \subseteq \mathcal{W}_{i-1}$ and B_i is an order ideal in $\mathcal{W}_{i-1}$, we obtain $a \in B_i \setminus \{m_i^*\}$, and the latter yields $a \notin \mathcal{W}_i$, which is a contradiction.

Proposition 3 (i) *The clusters* $C_1, C_2, \dots, C_k$ *are the classes of a linear order-congruence on* $(\mathcal{W}, \leq)$. (ii) R *is an extension of* $\leq$.

Proof. First, observe that our algorithm inserts the groups $C_p \subseteq \mathcal{W}$, $1 \leq p \leq k$ in a linear order $C_1 <_\rho C_2 <_\rho, \dots, <_\rho C_k$, corresponding to their indices. Denote by ρ the equivalence corresponding to the partition $\mathcal{W} = C_1 \cup C_2 \cup \dots \cup C_k$, and let $\varphi: P \to P/\rho$ be the map that assigns to any $x \in \mathcal{W}$ the cluster containing it, i.e., its $[x]_\rho$ class.

(i) Take $a, b \in \mathcal{W}$ such that $a \leq b$, and suppose that $\varphi(a) = [a]_\rho = C_i$ and $\varphi(b) = [b]_\rho = C_j$. Then either $i \leq j$ or $j < i$ holds. In the first case $\varphi(a) = C_i \leq_\rho C_j = \varphi(b)$. If this holds for each $a \leq b$, then ρ is an order-congruence. We prove that the case $j < i$ is excluded. Indeed, $b \in C_j \subseteq \mathcal{W}_{j-1}$, hence for $j < i$ we get $a \in \mathcal{W}_{i-1} \setminus C_j$ and $\mathcal{W}_{i-1} \subset \mathcal{W}_{j-1}$. Because B_j is an order ideal in $(\mathcal{W}_{j-1}, \leq)$ and m_j^* is a maximal element in B_j, we get that $C_j = B_j \setminus \{m_j^*\}$ is also an order ideal in $(\mathcal{W}_{j-1}, \leq)$. As $b \in C_j$, $a \in \mathcal{W}_{j-1}$, now the hypothesis $a \leq b$ yields $a \in C_j$, a contradiction.

(ii) Take any $a, b \in \mathcal{W}$ with $a \leq b$, and $\varphi(a) = C_i$, $\varphi(b) = C_j$. If $i = j$ i.e., $C_i = C_j$, then the order of a and b is determined within the group C_i by Algorithm 1 which preserves $\leq$. Thus, we obtain aRb. If $i \neq j$, then in view of (i) in the proof above, we get $i < j$ and $C_i <_\rho C_j$. Since $a \in C_i$ and $b \in C_j$, we get aRm_i' and $m_j^{\#} Rb$. Then aRm_i', $m_i' R m_{i+1}^{\#}$, $m_{i+1}^{\#} R m_{i+1}', \dots, m_{j-1}' R m_j^{\#}$, $m_j^{\#} Rb$ imply aRb. Thus R is an extension of $\leq$. □

6 Implementation

For experimentation, we have implemented a simple user interface. Its main purpose is to provide a more convenient way for defining the order of manufacturing operations in the technological matrix. The user interface has been implemented in

JavaScript with HTML5 Canvas. It helps check that the provided relation is a partial order or not. The proposed algorithm has implemented in GNU/Octave. Therefore, it is compatible with MATLAB. The relations are represented in a matrix. The language makes available the infinity symbol (inf) for denoting the infinite weights in the matrix.

At the first step, we must set a partial order. In fact, as a first input a binary relation ρ is defined between the elements of $\mathcal{W}$: $x_i \rho x_j$ means that the task x_i must precede the task x_j in the technological process. Hence, ρ can be represented by a directed graph. Its vertices are the elements of $\mathcal{W}$. Then the reflexive- transitive closure R of this relation ρ is formed – this will be a partial order whenever the graph is acyclic. Thus, the input is considered correct only in this acyclic case (see, e.g. [7]). Now, by using the adjacency matrix of the partial order R, generating the covering relation (given in a so-called Hasse-diagram) corresponding to R is a well-known function.

We have to insert the technologically justified weights (or we can generate random weights for testing purposes) for the covering pairs of tasks. We must also provide some weights for the incomparable tasks. In our example, we have generated uniformly distributed random values on the interval $[1,120] \subset \mathbb{R}$.

We calculate the order from the weights of the technological matrix. We assume that the transitions between manufacturing operations are possible, where the weights are finite. All further steps of the algorithm are organized into separate functions. The purposes of them are the following.

- Calculate the symmetric matrix from the initial technological matrix.
- Calculate the similarity matrix from the symmetric matrix.
- Calculate the transitive closure.
- Calculate the appropriate threshold value according to the number of the resulted clusters.
- Use the threshold value for clustering. Find equivalence classes to find minimal indices.
- Eliminate row and column (when removing a task from matrix representation).
- Find shortest path starting from the given (minimal) task.
- Evaluate the cost of the linear order.

6.1 Some Illustrative Examples

We must note that, the proposed algorithm works properly without the elimination of unnecessary weights (which are not part of the Hasse diagram). We can see the matrix representation of the partial order in Table 1 and the corresponding weight matrix in Table 2. (We note that there is no restriction on the effectuation order of the task O, therefore, it appears in the Hasse diagram as an element incomparable with the others.) The example below is from [12], pages 160-161, and it is related to a manufacturing process of a ventilation screw, where the effectuated operations were the following:

A, B: Longitudinal turning
(roughing respectively smoothing)
C, E, N: Insertion
D, J: Edge chamfer
F: Stabbing
G, H, I, O: Drilling
K: Thread cutting
L: Thread rolling
M: Thread turning
Q: Thread drilling
P: Rubbing
R: Dosage-collision

Table 1
The adjacency matrix of the considered partial order

	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**	**O**	**P**	**Q**	**R**
A	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	0
B	0	1	1	1	1	1	0	0	1	1	1	1	1	0	0	1	1	0
C	0	0	1	0	1	1	0	0	1	1	1	1	1	0	0	1	1	0
D	0	0	0	1	1	1	0	0	1	1	1	1	1	0	0	1	1	0
E	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0
F	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
G	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0
H	0	0	0	0	1	1	0	1	1	1	1	1	1	0	0	1	1	0
I	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	1	0
J	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	1	0
K	0	0	0	0	1	1	0	0	1	1	1	0	0	0	0	1	1	0
L	0	0	0	0	1	1	0	0	1	1	0	1	0	0	0	1	1	0
M	0	0	0	0	1	1	0	0	1	1	0	0	1	0	0	1	1	0
N	0	0	0	0	1	1	0	0	1	1	1	1	1	1	0	1	1	0
O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
P	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0
Q	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0
R	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1

Table 2
Edge weights in matrix form

	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**	**O**	**P**	**Q**	**R**
A	∞	4	∞	∞	∞	∞	19	5	∞	∞	∞	∞	∞	6	47	∞	∞	96
B	∞	∞	11	12	∞	∞	∞	71	∞	∞	∞	∞	∞	47	88	∞	∞	∞
C	∞	∞	∞	67	∞	∞	∞	86	∞	∞	13	14	15	16	76	∞	∞	∞
D	∞	∞	55	∞	∞	∞	∞	7	∞	∞	16	17	18	34	102	∞	∞	∞

E	∞	∞	∞	∞	∞	∞	∞	∞	103	100	∞	∞	∞	∞	74	41	40	∞
F	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	60	∞	∞	∞
G	14	7	∞	∞	∞	∞	∞	8	∞	∞	∞	∞	∞	9	90	∞	∞	∞
H	∞	25	51	68	∞	∞	∞	∞	∞	∞	19	21	22	6	45	∞	∞	∞
I	∞	∞	∞	∞	90	∞	∞	∞	∞	44	∞	∞	∞	∞	120	45	44	∞
J	∞	∞	∞	∞	38	∞	∞	∞	43	∞	∞	∞	∞	∞	31	43	42	∞
K	∞	∞	∞	∞	30	∞	∞	∞	28	29	∞	65	54	∞	24	∞	∞	∞
L	∞	∞	∞	∞	33	∞	∞	∞	31	32	20	∞	17	∞	32	∞	∞	∞
M	∞	∞	∞	∞	38	∞	∞	∞	35	36	88	59	∞	∞	108	∞	∞	∞
N	∞	107	75	9	∞	∞	∞	107	∞	∞	24	25	26	∞	68	∞	∞	∞
O	14	48	37	96	6	38	87	36	58	19	109	75	18	17	∞	60	57	88
P	∞	∞	∞	∞	∞	47	∞	∞	∞	∞	∞	∞	∞	∞	107	∞	49	∞
Q	∞	∞	∞	∞	∞	48	∞	∞	∞	∞	∞	∞	∞	∞	44	51	∞	∞
R	119	∞	∞	∞	∞	∞	3	∞	∞	∞	∞	∞	∞	∞	16	∞	∞	∞

The *dual of the Hasse diagram* of the considered partially ordered set (that is, the down-directed graph of the covering relation $\prec$) is from [12] (see Figure 1).

The result of the proposed algorithm for this numerical example is the following linear extension:

$$A \prec R \prec G \prec B \prec D \prec H \prec N \prec C \prec K \prec$$
$$L \prec M \prec E \prec J \prec O \prec I \prec P \prec Q \prec F.$$

The total cost of the resulted linear extension is 670.

6.2 Comparing with Simulated Annealing

The Simulated Annealing (SA) heuristic provides an alternative approach for finding the optimal path regarding to the cost matrix [21]. The method uses a temperature value (denoted by T) for limiting the magnitude of the local search. The SA method uses the following basic operations at parameters.

- It sets the initial T value and decrease it in each iteration. It chooses an initial permutation (as the path) denoted by π_0. It contains a permutation generator, which alters the actual permutation to achieve a new state in the search space.
- When the cost of the generated permutation is lower than the previous one ($f(\pi_{i+1}) < f(\pi_i)$) the algorithm updates the minimal cost and the path estimation to $f(\pi_{i+1})$ and π_{i+1}.
- In other case, it decides (based on the T value) to accept a worse cost and path.
- The stop condition of the SA algorithm is a limit for the T value.

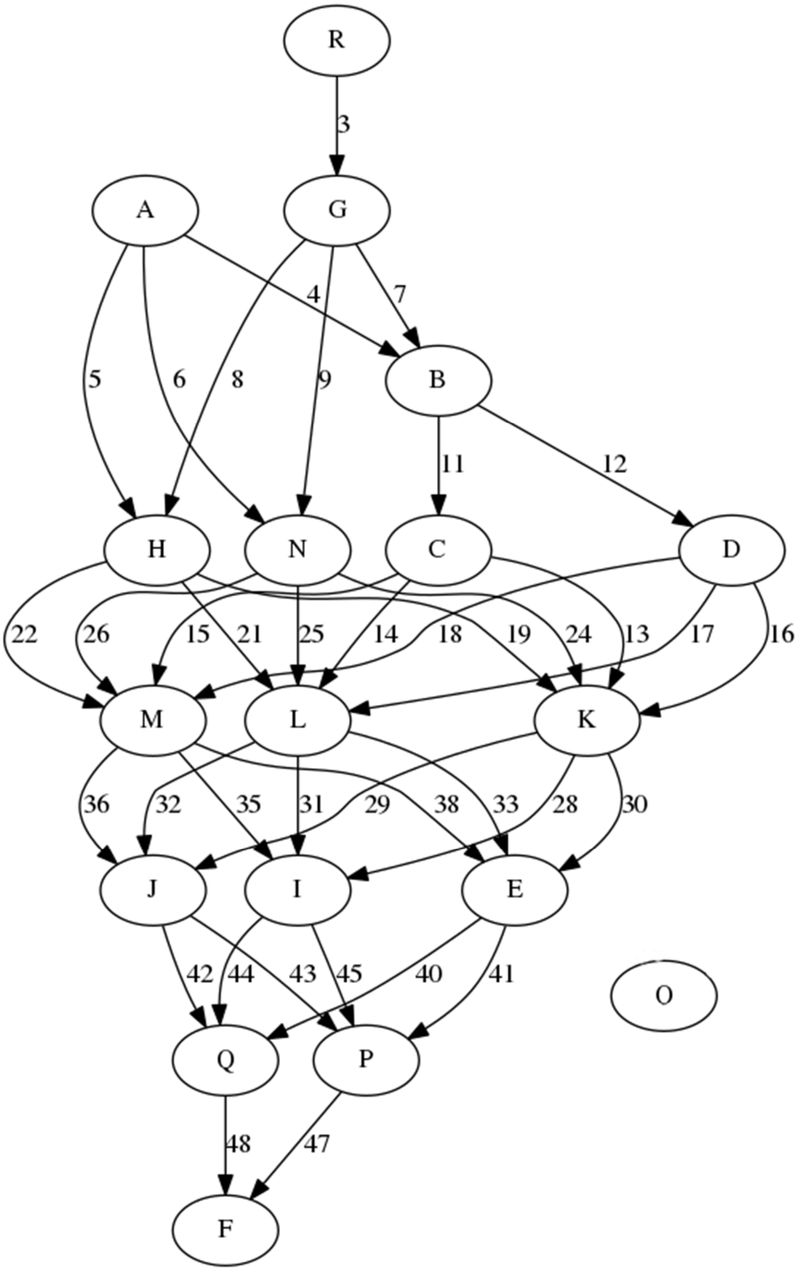


Figure 1
The dual of the Hasse diagram of the considered numerical example.
It also contains the weights on the edges of the graph.

At first as an experimentation, we tried to use a random search and check the provided $f(\pi_i)$ values. Unfortunately, it was unable to provide viable solution in a reasonable time. (The reason is the sparse cost matrix, where the probability of random linear extension is low.)

Next, we have improved the algorithm by providing a sub optimal linear extension. Instead of random permutations, the algorithm chooses two different indices in the path randomly and tried to swap them for a lower cost value. The resulted path is the following: $A, R, G, B, D, H, C, N, L, K, J, E, O, M, I, P, Q, F.$ The cost of this path is 573.

The experimentation with the SA algorithm shows that a relatively simple algorithm with local search can improve the result, but it had difficulties with finding the initial solution.

It is important to check the effectiveness of the algorithm by some statistical experimentations. In this way, we can see the results of the mentioned algorithms for larger set of examples, and the nondeterministic nature of the simulated annealing method also can be managed properly. In the followings, we show measurements which try to demonstrate how the algorithms performs for different sequencing problems.

We have chosen three other examples where the main operations were the following: Various kinds of turning and drilling, coarse grinding, transporting between workplaces and rubbing. Their adjacency matrix had sizes: 10, 12 and 16.

For comparing the efficiency of the Fuzzy Clustering and Simulated Annealing method, we generated 100 random weight matrices for each sample. (The count 100 is an experimental value.)

The generation of the weight matrix uses uniform distributed values from the range [0, 10). The algorithm sets infinity weights for the transitions which are not possible according to the adjacency matrix.

We have estimated the minimal costs both by using the proposed Fuzzy Clustering and the Simulated Annealing methods. We can see the distributions of the minimal costs on Figure 2. (The ranges of estimated minimal costs and the frequencies have been set to the same interval for each histogram, for helping the visual comparison easier.)

This statistical experiment shows that the Fuzzy Clustering method provides better estimation of the minimal costs.

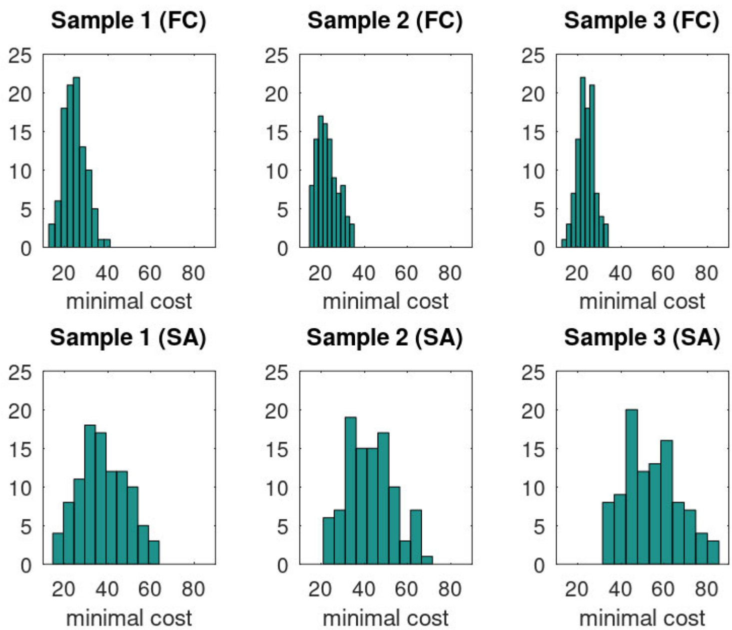


Figure 2

Histograms of the distribution of estimated minimal costs for our method using the Fuzzy Clustering (FC) and for the Simulated Annealing (SA) method

Conclusions

The method discussed in the paper is a heuristic procedure which can solve the job scheduling problem. It consists of a fuzzy clustering algorithm and a procedure for finding the optimal linear extensions of the given partial order inside the clusters. If the size of a cluster K is not too big (for instance, $|K| \leq 6$), then the second phase can be effectuated by a complete enumeration, i.e., in an exact way, as given in Algorithm 1b. Because the procedure is intended for medium-sized manufacturing processes, the size of the formed clusters is small, in general – therefore, our method (which uses Algorithm 1b) is a locally exact heuristic method.

The method has been implemented in GNU/Octave, and has been extended by a JavaScript application, which can help the engineers by guaranteeing that the input matrix is mathematically correct (antisymmetric and acyclic). The Octave source code contains many functions for making possible to reuse or improve some parts of the proposed algorithm.

The results of the proposed algorithm have been illustrated on a small-sized problem; however, this size is typical in real-world applications. Some logistical scheduling problems belong also to the application area of our algorithm.

It is hard to compare the achieved results with state-of-the-art solutions of the manufacturing industry based on some modifications of the travelling salesman method. In fact, the restricted travelling salesman method has several semi-heuristic variants which are applied to large scale scheduling problems, but most of them are proprietary. Their time complexity is generally better than our method, however, for medium-sized job sequences the latter is more flexible. This means that in the case of modification of the technological process the input of our algorithm can be adapted easily. This flexibility is due to the fuzzy clustering procedure, which makes possible to change some threshold values appropriately during the optimization process.

Comparing our method with Eszes procedure, we can state that the results obtained by our method represent better approximations of the optimal solutions because the formed clusters are more realistic. Both proposed fuzzy clustering method and the SA method have many parameters. We must conclude that, our preferred method is more robust in general, and suitable for problems where the manual tuning of the parameters (fuzzy clustering) is beneficial. It does not seem to be relevant to measure the running times of the program, because the Octave software has been chosen for gaining experiments of the algorithm conveniently, and it has terminated in few seconds on modern desktop computers.

Acknowledgment

We dedicate our paper to the memory of Professor Tibor Tóth who passed away in 2020. This research was initiated by him, and each of the authors is grateful to profit from numerous inspiring and friendly discussions with this outstanding researcher. In particular, the fifth author is indebted to Tibor as PhD supervisor.

The authors would like to thank the reviewers for their valuable comments and suggestions which significantly improved the presentation of the paper.

References

[1] Al-wswasi, M., Ivanov, A., and Makatsoris, H.: A survey on smart automated computer-aided process planning (ACAPP) techniques. The International Journal of Advanced Manufacturing Technology, 97(1), (2018) pp. 809-832

[2] Dubois, D., Prade, H.: Fuzzy Sets and Systems, Theory and Applications; Mathematics in Science and Engineering, Vol. 144, Academic Press, New-York, p. 393 (1980) ISBN-13: 978-0122227509

[3] Halevi, G. and Weill, R. D.: Principles of Process Planning, A logical approach, Springer Science+Business Media, Dordrecht (1995) ISBN: 978-0-412-54360-9

[4] Hompel, M., Vogel-Heuser, B. and Bauernhansl, T.: Industry 4.0 handbook Bd.1: Produktion, Springer Reference Technik, Berlin (2017) ISBN 978-3-662-45278-3 [in German]

[5] Hompel, M., Vogel-Heuser, B. and Bauernhansl, T.: Industry 4.0 handbook

Bd.2: Automatisierung, Springer Reference Technik, Berlin (2017) ISBN 978-3-662-53247-8 [in German]

[6] Hompel, M., Vogel-Heuser, B. and Bauernhansl, T.: Industry 4.0 handbook Bd.3: Logistik, Springer Reference Technik, Berlin (2017), ISBN 978-3-662-53250-8 [in German]

[7] Huang, W., Lin, W. and Xu, S.: Application of graph theory and hybrid GA-SA for operation sequencing in a dynamic workshop environment. Computer-Aided Design and Applications, 14(2), (2017) pp. 148-159

[8] Körtesi, P., Radeleczki, S. and Szilágyi, Sz.: Congruences and isotone maps on partially ordered sets, Math. Pannonica, 16/1 (2005), pp. 39-55

[9] Radeleczki, S.: A classification method based on fuzzy contexts, Bul. Stiint. Univ. Baia-Mare, Ser. B, Matematica-Informatica, Vol. XVIII (2002), Nr. 2, pp. 319-324

[10] Radeleczki, S, Tóth, T.: Fuzzy methods in Group Technology.: Research report, OTKA 2348/91, (4.2/6) Univ. of Miskolc, Inst. of Informatics (1992), pp. 1-34 (In Hungarian)

[11] Phung, L. X., Van Tran, D., Hoang, S. V., and Truong, S. H.: Effective method of operation sequence optimization in CAPP based on modified clustering algorithm. Bulletin of the JSME, Journal of Advanced Mechanical Design, Systems, and Manufacturing, 11(1), (2017), JAMDSM0001-JAMDSM0001

[12] Tóth, T.: Production Systems and Processes, Fundamentals of Production Informatics, University of Miskolc Publishing House (Miskolci Egyetemi Kiadó) (2004) p. 464, ISBN: 9636616302 (in Hungarian)

[13] Tóth, T., Detzky, I. and Eszes, L.: Optimization of Discrete Technology Processes Using a Method Traced Back to Constrained Travelling Salesman Problem. Proceedings of the 27th International MATADOR Conference, Manchester, 21. April (1988), editor Davis B. J., edited by University of Manchester in collaboration with MacMillan Education Ltd. pp. 145-153

[14] Trotter, W. T.: Combinatorics and Partially Ordered Sets, Dimension Theory, John Hopkins University Press, Baltimore and London, (1992), ISBN-13: 978-0801869778

[15] Xu, H. and Wang, H. P.: Part family formation for GT applications based on fuzzy mathematics, International Journal of Production Research, Vol. 27, No. 9 (1989), pp. 1637-1651

[16] Xu, X., Wang, L. and Newman, S. T.: Computer-aided process planning - a critical review of recent developments and future trends. International Journal of Computer Integrated Manufacturing, 24(1) (2011) pp. 1-31

[17] Lin C. J, Wang H. P.: Optimal operation planning and sequencing minimization of tool change overs. International Journal of Production

Research, 31 (1993) pp. 311-324

[18] Koulamas C.: Operations sequencing and machining economics. Int J Prod Res 31(4), (1993) pp. 957-975

[19] Bhaskara Reddy SV, Shunmugam MS, Narendran TT.: Operation sequencing in CAPP using genetic algorithms. Int J Prod Res 37(5), (1999), pp. 1063-1074

[20] Li L, Fuh JYH, Zhang YF, Nee AYC.: Application of genetic algorithm to computer-aided process planning in distributed manufacturing environments. Robot Comput-Integr Manuf. 21 (2005) pp. 568-578

[21] Foerster H, Wischer G.: Simulated annealing for order spread minimization in sequencing cutting patterns. Eur J. Oper. Res. 110 (1998) pp. 272-281

[22] Li WD, Ong SK, Nee AYC.: Optimization of process plans using a constraint-based tabu search approach. Int J. Prod Res, 42 (10), (1998), pp. 1955-1985

[23] Krishna AG, Rao KM.: Optimisation of operations sequence in CAPP using an ant colony algorithm. Int J Adv Manuf Technol 29, (2006), pp. 159-164

[24] Guo YW, Mileham AR, Owen GW, Li WD: Operation sequencing optimization using a particle swarm optimization approach. Proc Inst Mech Eng, B, J. Eng Manuf. 220 (B12), (2006), pp. 1945-1958

[25] Mojtaba S, Reza T-M.: Application of genetic algorithm to computer-aided process planning in preliminary and detailed planning. Eng Appl Artif Intell. 2009.04.05, doi:10.1016/j.engappai

[26] Korde UP, Bora BC, Stelson KA, Riley DR (1992) Computer aided process planning for turned parts using fundamental and heuristic principles. J Eng Ind, (1992), pp. 114:31-40

[27] Nallakumarasamy, G., Srinivasan, P. S. S., Raja, K. V., and Malayalamurthi, R.: Optimization of operation sequencing in CAPP using simulated annealing technique (SAT). Int J Adv Manuf Technol, 54(5-8), (2011), pp. 721-728

[28] Ssemakula, M. E., and Rangachar, R. M.: The prospects of process sequence optimization in CAPP systems. Computers and Industrial Engineering, 16(1), (1989) pp. 161-170

[29] Rétfalvi, A., and Stampfer, M.: The key steps toward automation of the fixture planning and design. Acta Polytechnica Hungarica, 10(6), (2013), pp. 77-98

[30] Azemi, F., Šimunović, G., Lujić, R., Tokody, D., and Rajnai, Z.: The use of advanced manufacturing technology to reduce product cost. Acta Polytechnica Hungarica, 16(7), (2019), pp.115-131

[31] Bechler, G.: Demand-oriented product line optimization, eBook, Springer,

(2019), ISBN 978-3-658-25542-8 [in German]

[32] Krapp, S.: The successful design of logistics in cooperative groups: An empirical study, eBook, Springer, (2018), ISBN: 978-3-658-23888-9 [in German]

[33] Scholz, D.: Mixed-integer linear layout planning model, In: Innerbetriebliche Standortplanung, Gabler Verlag, (2010) ISBN 978-3-8349-2277-9 [in German]

[34] Wehking, K.-H., Yousefifar, R., Nortwyck, R.: Computer-aided planning, technical handbook for logistics, Springer Verlag, (2010) ISBN 978-3-662-60868-5 [in German]

[35] Zahorán, L. and Kovács, A.: ProSeqqo: A generic solver for process planning and sequencing in industrial robotics. Robotics and Computer-Integrated Manufacturing, 78, (2022), 102387, (DOI: 10.1016/j.rcim.2022.102387)

On the Use of Quaternions, in the Translated Reference Frame Formalism

Claudiu Pozna

Department of Automation, Széchenyi István University,
Egyetem tér 1, 9026 Győr, Hungary

Department of Automation and Information Technology, Transilvania University of Brasov, Str. Mihai Viteazu 5, Corp V, et. 3, 500174 Brasov, Romania

E-mail: cp@unitbv.ro

Radu-Emil Precup

Department of Automation and Applied Informatics, Politehnica University of Timisoara, Bd. V. Parvan 2, 300223 Timisoara, Romania

Center for Fundamental and Advanced Technical Research, Romanian Academy – Timisoara Branch, Bd. Mihai Viteazu 24, 300223 Timisoara, Romania

E-mail: radu.precup@upt.ro

Abstract: The approaches dedicated to the geometric model of the object, in the object-oriented modeling framework, specific to robots, are focused on the computation of the position of the elements, relative to a global reference frame. One important problem, in this regard, is the direct geometry, which starts with the initial data represented by the generalized coordinates and parameters of each element (in its own reference frame) and calculates the configuration of the structure in the global reference frame. This paper extends the Authors' geometrical modeling approach, referred to as the translated reference frame formalism, by changing the mathematical apparatus of homogeneous transformations with that specific to quaternions. This new approach has three main advantages over the state-of-the art, namely, it simplifies substantially the configuration modeling, the transformation parameters are intuitive, and the computation is substituted with the operation of choosing a suitable transformation that belongs to a set of six homogeneous transformations. The suggested approach is validated by the computation of the geometric models of two illustrative robotic examples. The source codes are available in a public repository.

Keywords: direct geometry; geometric modeling; homogeneous operators; quaternions; translated reference frame formalism

1 Introduction

Modeling is a process of knowledge, that uses two mechanisms of thinking, approximation and conceptualization. The approximation refers to the simplification of the phenomenon by eluding the features considered insignificant (by the model designer), and the conceptualization refers to the replacement of the particular by the general, by the concept. It is widely acknowledged that the model is an abstraction of reality; the difference between the model and the reality is inherent and called perturbation. In order to be aware of this difference, the model design is preceded by the specification of the hypotheses, i.e., the conditions that state when the model fully reflects the phenomenon.

In the case of robotics, the object-oriented modeling framework organizes the model construction through a succession of increasingly precise model classes. The analysis and design principles specific to object-oriented framework are set in [1]. These principles are applied in [2] and [3] to create a software library that is used in engineering and teaching applications as suggestively illustrated in [4]. An application of object-oriented modeling to delta robots is described in [5]. The analysis and design principles are implemented in Modelica in [6], where predefined objects of Lego type are assembled in different configurations.

In the context of object-oriented modeling, a succession of three classes of models, namely geometric, kinematic and dynamic ones, is formulated in [7]. Each class contains the related assumptions (hypothesis), the properties they use, as well as the methods they offer. The mentioned classes are in a parent-child relationship, which means that they can inherit their methods and properties. The basis of this succession is the geometric model, which is the father of the kinematic one, which in turn is the father of the dynamic one.

The geometric model object contains the richest collection of hypotheses (eliminated, to a certain extent, by kinematic and dynamic models). Such examples of hypotheses are: the phenomena are out of time, there is no movement, bodies are rigid, interactions with the environment are non-existent, and there are no inertial effects. The properties of the geometric model refer to abstractions of the size and orientation of the bodies obtained by attaching reference (or coordinate) frames (or systems) to each element, as, for example, the lengths of the elements, the angles of rotation, and the displacements in the joints (i.e., the generalized coordinates).

The approaches specific to the geometric model object deal with the computation of the position of the elements relative to a global reference frame. Two types of problems are important in this regard:

(i) *The direct geometry*, which starts with the initial data represented by the generalized coordinates and parameters of each element (in its own reference frame), and calculates the configuration of the structure in the global reference frame.

(ii) *The inverse geometry*, which starts with the input data represented by the position of the gripper (effector), and calculates the generalized coordinates of the robot joints.

Several formalisms have been proposed to solve the direct geometry problem. They will be briefly discussed as follows. The Denavit-Hartenberg (DH) formalism [8] [9] is the most popular one, it is imagined for a structure with arbitrary axes (the general case), but it proves to be difficult to be applied to structures with parallel or perpendicular axes. The DH formalism model describes the technological dimensions of robot's elements through a sequence of four parameters, three constant ones and one variable one. The DH formalism is followed by Paul's approach [10], which uses DH parameters but is focused on structures with translation and rotation couplets, and Khalil and Kleinfinger's approach [11], which uses the same DH parameters but generalizes the formalism to closed and tree structures. A significant change is made by Craig in [12] in terms of moving the origin of the reference frame of the element from the upstream joint to the downstream joint. However, it is difficult to define the three parameters and to mathematically compute the homogeneous transformation. The DH and Craig's formalisms are included in software packages as that developed by Corke in [13] for Matlab and in the Robot Analyzer coordinated by Saha [14]. Another formalism, proposed by Gogu and his co-authors in [15] and [16], conceives a significant simplification that refers to particular structures with parallel or perpendicular joints axes. It is also important to mention the screw theory-based formalism [17-19], which proves to be a modern and efficient alternative to the DH formalism.

The paper is an extension of authors' recent approach given in [20] and referred to as translated reference frame formalism by changing the mathematical apparatus of homogeneous transformations with that specific to quaternions. This novel approach to solve the direct geometry problem is important with respect to the state-of-the-art discussed above because of three reasons, (a), (b) and (c). (a) It simplifies substantially the configuration modeling. (b) The transformation parameters are intuitive. (c) The computation is substituted with the operation of choosing a suitable transformation that belongs to a set of six homogeneous transformations. In addition, the quaternion-based approach proposed by this paper is advantageous as it is attractive for rotational modeling and for generalization to structures with arbitrary axes.

The paper is organized as follows: the stages of the translated reference frame formalism proposed in [20] are briefly recalled in the next section. The presentation of the mathematical apparatus associated to the formalism is prepared in Section 3 focused on quaternions. Sections 4 and 5 describe the proposed approach, in a particular version applied to structures with parallel and or perpendicular axes, and its generalization to arbitrary axes as well, respectively, and an algorithm is formulated in this regard. Section 6 gives the use of the suggested mathematical tools to calculate the Jacobian matrix. The theoretical results are validated in Section 7 in terms of two illustrative examples, their analytical solution and the software implementation and application of the algorithm. The conclusions highlighted in Section 8 conclude the paper.

2 Overview of Translated Reference Frame Formalism

Since each element of the robot's structure is defined in its own reference (or coordinate) frame (or system), the formalisms mentioned in the previous section are expressed as algorithms that ultimately determine the transformation operators from one reference frame to another one. More precisely, the parameters of each element are defined in its own (local) reference frame, and it is of interest to express their values in other reference frames (local or global ones). In this context, the position of the effector in the global reference frame is of interest, i.e.

$$\begin{bmatrix} \boldsymbol{p} \\ \boldsymbol{o} \end{bmatrix} = \boldsymbol{f}(q_1, \ldots, q_n) \tag{1}$$

where $\mathbf{p}$ is the position of the gripper, $\mathbf{o}$ is the orientation of the gripper (in the Cartesian space), and q_1, …, q_n are the generalized coordinates (variables of joint space). The result in (1) is important in the representation of the configuration, the solution to the kinematic model (the calculation of the Jacobian matrix), and the construction of the dynamic model of the structure.

The translated reference frame formalism [20] is described and organized in terms of several rules that involve the variables illustrated in Figure 1. For example, for the joint i, the generalized variable is q_i, the origin of the reference frame is O_i, and the unit vector of the axis i is $\hat{\mathbf{u}}_{i-1}^{i}$ (defined in the reference frame $\{i–1\}$).

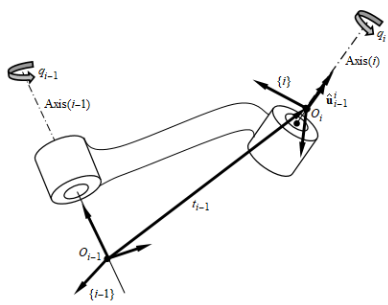


Figure 1
The link between the reference frames $\{i–1\}$ and $\{i\}$ (adapted from [20])

The formalism is organized in the following steps [20]:

1. The reference frames are attached to each element.

1.a. The origins of the reference frames are defined as the points O_1, …, O_n, on the axes of rotation or translation of each joint:

$$O_i \in Axis(i). \tag{2}$$

1.b. The generic reference frame of the first joint is defined so that one of the axes of the system coincides with the axis of rotation and / or translation of the joint:

$$\{0\} \mid \hat{\mathbf{x}}^0 \vee \hat{\mathbf{y}}^0 \vee \hat{\mathbf{z}}^0 \equiv \hat{\mathbf{u}}_0^1. \tag{3}$$

1.c. The chosen reference frame is translated to each chosen origin, obtaining a set of n reference frames. Since the joints have mutually parallel or perpendicular axes, each axis of the joints will overlap on one of the axes of the reference frame corresponding to that axis. The reference frame of the element i, namely $\{i\}$, is located on the i axis:

$$\{0\} \parallel \{i\}, \forall i = 1...n. \tag{4}$$

2. The transformation parameters from the reference frame $\{i\}$ to the reference frame $\{i–1\}$ is defined.

2.a. The position vector of the point O_i is expressed in the reference frame $\{i–1\}$.

2.b. The unit vector of the i axis is expressed in the reference frame $\{i–1\}$.

2.c. The generalized variable q_i is expressed, i.e., the angle of rotation or the length of translation at joint i.

3. The mathematical formulae of the formalism are applied.

3.a. Six homogeneous transformations have been identified to solve all configurations:

$${}^{i-1}_{i}\boldsymbol{T} = \begin{bmatrix} {}^{i-1}_{i}\boldsymbol{Q} & {}^{i-1}_{i}\boldsymbol{P} \\ \mathbf{0}_{1\times 3} & 1 \end{bmatrix} \tag{5}$$

where $\mathbf{Q} \in \{\mathbf{R}_x, \mathbf{R}_y, \mathbf{R}_z, \mathbf{I}\}$ is the rotation component of the operator, $\mathbf{R}_x$, $\mathbf{R}_y$ and $\mathbf{R}_z$ are the elementary angle rotations around the X, Y and Z axes attached to q_i:

$$\mathbf{R}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(q_i) & -\sin(q_i) \\ 0 & \sin(q_i) & \cos(q_i) \end{bmatrix}, \mathbf{R}_y = \begin{bmatrix} \cos(q_i) & 0 & \sin(q_i) \\ 0 & 1 & 0 \\ -\sin(q_i) & 0 & \cos(q_i) \end{bmatrix},$$
$$\mathbf{R}_z = \begin{bmatrix} \cos(q_i) & -\sin(q_i) & 0 \\ \sin(q_i) & \cos(q_i) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{6}$$

If the joint i is prismatic, three translations in the directions X, Y, Z are possible, $\mathbf{Q} = \mathbf{I}$, and the translation component of the transformation is

$$ {}^{i-1}_{\ \ i}\mathbf{P} = \mathbf{t}_{i-1} + q_i \varepsilon_i \hat{\mathbf{u}}^{i}_{i-1}. \tag{7} $$

If the joint i is revolute, then $\varepsilon_i = 0$ in (7).

3.b. The generalization of the formalism, its use for structures with arbitrary axes (not necessarily parallel or perpendicular) is conducted by means of a homogeneous operator that rotates the axis i–1 to the axis i:

$$ {}^{i-1}_{\ \ i}\boldsymbol{T} = \begin{bmatrix} \boldsymbol{R}(\alpha, \hat{\boldsymbol{v}}) & {}^{i-1}_{\ \ i}\boldsymbol{P} \\ \mathbf{0}_{1\times 3} & 1 \end{bmatrix} \begin{bmatrix} {}^{i-1}_{\ \ i}\boldsymbol{Q} & \mathbf{0}_{3\times 1} \\ \mathbf{0}_{1\times 3} & 1 \end{bmatrix} \tag{8} $$

where $\mathbf{R}(\alpha, \hat{\mathbf{v}})$ is a Rodriguez-type rotation matrix, α is the angle of rotation and $\hat{\mathbf{v}}$ is the unit vector of the axis of rotation (also illustrated in Figure 2).

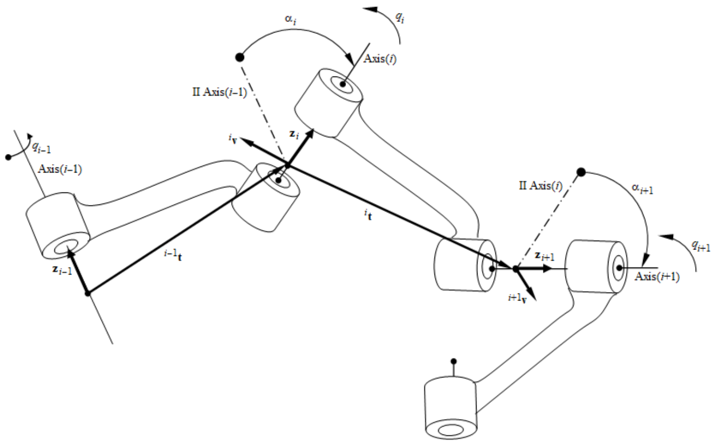


Figure 2
The general form of the formalism (adapted from [20])

This formalism, which actually carries out the transformation from the reference frame $\{i\}$ to the reference frame $\{i–1\}$, can be intuited as a journey of the reference frame $\{i–1\}$ to the reference frame $\{i\}$. In other words, it performs the translation of the reference frame $\{i–1\}$ to the origin of the reference frame $\{i\}$, followed by the rotation with the angle α around the unit vector $\hat{\mathbf{v}}$, and finally by the rotation or translation corresponding to the generalized variable q_i.

The subject of this paper is the transformation of the proposed formalism by modifying the mathematical formulae given above. This modification concerns the relations that use homogeneous matrices with new ones that manipulate quaternions. The steps 1 and 2 of the formalism will be kept, and the step 3 will be modified in terms of introducing new relations.

3 Definitions and Operations on Quaternions

Quaternions are hypercomplex numbers defined by Hamilton in [21] to describe the three-dimensional rotations of objects. The notation ***q*** will be used as follows for a quaternion, and it employs the following equivalent forms:

$$\begin{aligned} &\mathbf{q} = a + b\,i + c\,j + d\,k, \\ &\mathbf{q} = (a \quad b \quad c \quad d), \\ &\mathbf{q} = (a \quad \mathbf{v}), \end{aligned} \tag{9}$$

where a, b, c and d are real numbers, i, j and k are the imaginary elements (a generalization of the imaginary element i of a complex number), a is the real part of the quaternion, **v** is its vector part

$$\mathbf{v} = b\,i + c\,j + d\,k, \tag{10}$$

and the imaginary elements fulfill

$$\begin{aligned} &i^2 = j^2 = k^2 = -1, \\ &i \times j = k,\; j \times k = i,\; k \times i = j, \\ &j \times i = -k,\; k \times j = -i,\; i \times k = -j. \end{aligned} \tag{11}$$

Two operations are defined on the set of quaternions, namely addition and multiplication (non-commutative):

$$\begin{aligned} &\mathbf{q_1} + \mathbf{q_2} = (a_1 + a_2 \quad \mathbf{v_1} + \mathbf{v_2}), \\ &\mathbf{q} + \mathbf{0} = \mathbf{q}, \\ &\mathbf{q_1 q_2} = (a_1 a_2 - \mathbf{v_1 v_2} \quad a_1 \mathbf{v_2} + a_2 \mathbf{v_1} + \mathbf{v_1} \times \mathbf{v_2}) \neq \mathbf{q}_2 \mathbf{q}_1, \\ &\mathbf{q0} = \mathbf{0}, \\ &\mathbf{q_1}(\mathbf{q_2 q}_3) = (\mathbf{q}_1 \mathbf{q}_2)\mathbf{q}_3, \end{aligned} \tag{12}$$

where:

$$\begin{aligned} &\mathbf{0} = (0 \quad 0 \quad 0 \quad 0), \\ &\mathbf{1} = (1 \quad 0 \quad 0 \quad 0), \end{aligned} \tag{13}$$

which means that they belong to a non-commutative algebraic structure.

Rotations are described by rotation operators. For example, the rotation of the vector **p**, its transformation into the vector ***q*** around the unit vector **v** with the angle α is expressed as

$$\mathbf{q} = \left(\cos\frac{\alpha}{2} \quad \mathbf{v}\sin\frac{\alpha}{2}\right)(0 \quad \mathbf{p})\left(\cos\frac{\alpha}{2} \quad \mathbf{v}\sin\frac{\alpha}{2}\right)^{*}. \tag{14}$$

The operations will be integrated in matrix computation in the next section.

4 Integration of Quaternions in the Translated Reference Frame Formalism

The proposed formalism preserves the first two steps that define the local reference frame (or coordinate systems) but modifies the mathematical apparatus. The transformation from the reference frame $\{k\}$ to the reference frame $\{i–1\}$ is carried out using

$$ {}^{i-1}_{k}\mathbf{T} = [\,{}^{i-1}_{i}\mathbf{r} \quad \mathbf{1}] \cdot \begin{bmatrix} {}^{i}_{k}\mathbf{T} & \mathbf{0} \\ \mathbf{0} & {}^{i-1}\mathbf{t} \end{bmatrix} \cdot \begin{bmatrix} {}^{i-1}_{i}\mathbf{r}^{*} \\ \mathbf{1} \end{bmatrix}, \tag{15} $$

where:

$$ \begin{aligned} &{}^{i-1}_{i}\mathbf{r} = (\cos(q_i/2) \quad \hat{\mathbf{u}}^{i}_{i-1}\sin(q_i/2)), \\ &{}^{i-1}_{i}\mathbf{t} = (0 \quad \mathbf{t}_{i-1} + \varepsilon_i q_i \hat{\mathbf{u}}^{i}_{i-1}) = (0 \quad t_{x,i-1} \quad t_{y,i-1} \quad t_{z,i-1}) + \varepsilon_i q_i (0 \quad \hat{\mathbf{u}}^{i}_{i-1}), \\ &\varepsilon_i = \begin{cases} 0 & \text{for revolute joint}, \\ 1 & \text{for prismatic joint}, \end{cases} \end{aligned} \tag{16} $$

q_i is the generalized variable at joint i, $\hat{\mathbf{u}}^{i}_{i-1}$ is the unit vector on which the movement of the joint i takes place, expressed in the reference frame $\{i–1\}$, $t_{x,i–1}$, $t_{y,i–1}$ and $t_{z,i–1}$ are the components of the vector linking the origin of the reference frame $\{i–1\}$ to the reference frame $\{i\}$ expressed in the reference frame $\{i–1\}$. Pointing out that

$$ \begin{aligned} &{}^{k-1}_{k}\mathbf{T} = [\mathbf{1} \quad \mathbf{1}] \cdot \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & {}^{k-1}\mathbf{t} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix}, \\ &{}^{k}_{k}\mathbf{T} = \mathbf{0}, \end{aligned} \tag{17} $$

with **1** and **0** defined in (13), the relationship (15) is iterative in i, it starts with the reference frame k=1…n and iterates downstream, at the limit, to the global reference frame $\{0\}$.

The relationship (15) can be reversed as follows:

$$ {}^{i}_{k}\mathbf{T} = [\,{}^{i-1}_{i}\mathbf{r}^{*} \quad \mathbf{1}] \cdot \begin{bmatrix} {}^{i-1}_{k}\mathbf{T} - {}^{i-1}\mathbf{t} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \cdot \begin{bmatrix} {}^{i-1}_{i}\mathbf{r} \\ \mathbf{1} \end{bmatrix}, \tag{18} $$

with the interpretation that knowing the transformation from the reference frame $\{k\}$ to the reference frame $\{i–1\}$, the transformation to the reference frame $\{i\}$ can be calculated. Specifically, an upstream iteration is carried out.

If the orientations of the effector are intended to be computed, making use of the calculation of the unit vectors of its reference frame in the basic reference frame, the following formula can be used:

$$\mathbf{w}_0^n = \left(\prod_{i=1}^{n} {}_{i}^{i-1}\mathbf{r} \right) \cdot \mathbf{w}_n^n \cdot \left(\prod_{i=1}^{n} {}_{i}^{i-1}\mathbf{r} \right)^*, \tag{19}$$

where $\mathbf{w}_0^n$ is the quaternion corresponding to the unit vectors in the basic reference frame, and

$$\mathbf{w}_n^n = \begin{cases} (0 \quad 1 \quad 0 \quad 0) & \text{for} \quad \hat{\mathbf{n}}, \\ (0 \quad 0 \quad 1 \quad 0) & \text{for} \quad \hat{\mathbf{o}}, \\ (0 \quad 0 \quad 0 \quad 1) & \text{for} \quad \hat{\mathbf{a}}, \end{cases} \tag{20}$$

where $\hat{\mathbf{n}}$, $\hat{\mathbf{o}}$ and $\hat{\mathbf{a}}$ are highlighted in Figure 3. It is underlined that the cuaternion product is associative but not commutative.

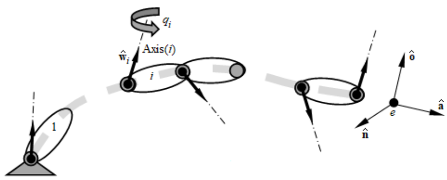


Figure 3
Jacobian matrix computation

5 Generalization to Arbitrary Axes

The translated reference frame formalism can be generalized to any type of structure as illustrated in Figure 2. The structures of serial robots have, generally, parallel or perpendicular to the axes of rotation or translation. This substantially simplifies the calculation of the rotation quaternion

$${}_{i}^{i-1}\mathbf{r} = (\cos(q_i/2) \quad \hat{\mathbf{u}}_{i-1}^{i} \sin(q_i/2)), \tag{21}$$

where $\hat{\mathbf{u}}_{i-1}^{i}$ is the unit vector of axis i, expressed in the reference frame $\{i–1\}$, with the possible expressions

$$\hat{\mathbf{u}}_{i-1}^{i} = \begin{cases} [1 \quad 0 \quad 0] & \text{if} \quad \text{Axis}(i) \parallel \hat{\mathbf{x}}^0, \\ [0 \quad 1 \quad 0] & \text{if} \quad \text{Axis}(i) \parallel \hat{\mathbf{y}}^0, \\ [0 \quad 0 \quad 1] & \text{if} \quad \text{Axis}(i) \parallel \hat{\mathbf{z}}^0. \end{cases} \tag{22}$$

But if a more general situation with two rotations is considered, the first overlaps the axis of the joint i–1 with the axis of the joint i, and the second rotates the joint with the generalized variable q_i in terms of

$$ {}^{i-1}_{i}\mathbf{r} = (\cos(\alpha_i/2) \quad \hat{\mathbf{v}}_{i-1}\sin(\alpha_i/2))(\cos(q_i/2) \quad \hat{\mathbf{u}}^{i}_{i-1}\sin(q_i/2)), \tag{23} $$

with:

$$ \begin{aligned} &\hat{\mathbf{v}}_{i-1} = \hat{\mathbf{z}}_i \times \hat{\mathbf{z}}_{i-1}, \\ &\alpha_i = \cos^{-1}(\hat{\mathbf{z}}_i \cdot \hat{\mathbf{z}}_{i-1}), \end{aligned} \tag{24} $$

where:

$$ \hat{\mathbf{z}}_{i-1} = [0 \quad 0 \quad 1]^T, \tag{25} $$

and $\hat{\mathbf{z}}_i$ is the unit vector of the axis of rotation, expressed in the reference frame $\{i–1\}$.

The proposed formalism and the steps presented in previous section along with the information in this section are organized systematically in the algorithm presented in Figure 4.

6 Jacobian Matrix Computation

The Jacobian matrix is the fundamental concept of the kinematic model

$$ \begin{bmatrix} \mathbf{v} \\ \boldsymbol{\omega} \end{bmatrix} = \mathbf{J}_0(\mathbf{q})\dot{\mathbf{q}}, \tag{26} $$

where $\mathbf{v}$ and $\boldsymbol{\omega}$ are the linear and the angular velocities, respectively, of the point of interest, respectively (at the limit of the origin of the effector the coordinate system), $\mathbf{J}_0 = \begin{bmatrix} {}^{v}\mathbf{J}_0 \\ {}^{\omega}\mathbf{J}_0 \end{bmatrix}$ is the Jacobian matrix expressed in the global reference frame, consisting of the two elements, the upper one referring to linear velocities, and the lower one corresponding to angular velocities, $\mathbf{q} = [q_1 \quad ... \quad q_n]^T$ is the vector of generalized variables, T indicates matrix transposition, and n is the number of degrees of freedom of the structure.

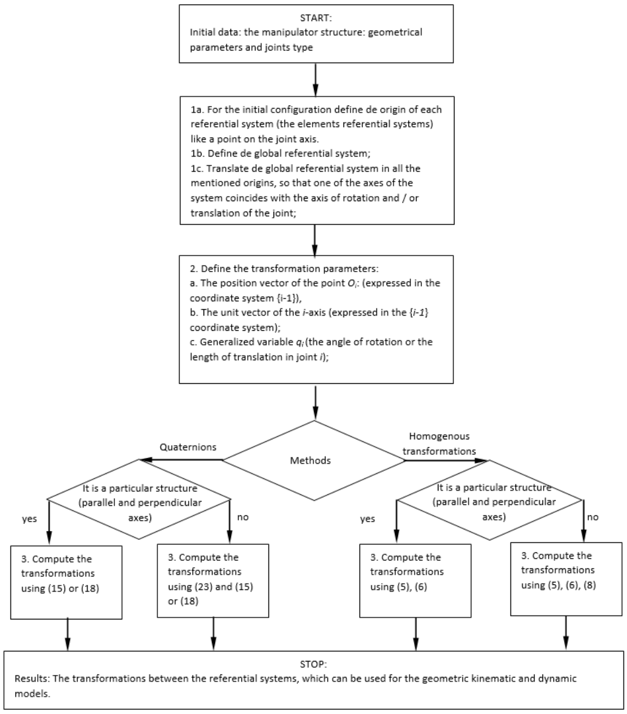


Figure 4

The block diagram of the proposed formalism

The computation of the Jacobian matrix is done as follows separately for the two elements mentioned above in relation with Figure 3. The upper element, corresponding to the linear velocities, is calculated using the partial derivatives of the components of the vector included in the quaternion ${}^{\mathbf{0}}_{\mathbf{e}}\mathbf{T} = (0 \quad {}^{0}p_{xe} \quad {}^{0}p_{ye} \quad {}^{0}p_{ze})$

$$ {}^{v}\mathbf{J}_0 = \begin{bmatrix} \dfrac{\partial p_{xe}}{\partial q_1} & \dots & \dfrac{\partial p_{xe}}{\partial q_n} \\ \dfrac{\partial p_{ye}}{\partial q_1} & \dots & \dfrac{\partial p_{ye}}{\partial q_n} \\ \dfrac{\partial p_{ze}}{\partial q_1} & \dots & \dfrac{\partial p_{ze}}{\partial q_n} \end{bmatrix}. \tag{27} $$

The lower element, corresponding to the angular velocities, is computed using the rotation axes

$$ {}^{\omega}\mathbf{J}_0 = [\varepsilon_1 \cdot \hat{\mathbf{w}}_0^1 \quad \dots \quad \varepsilon_n \cdot \hat{\mathbf{w}}_0^n], \tag{28} $$

where $\hat{\mathbf{w}}_0^i$ is the unit vector of the $\{i\}$ axis expressed in the base reference frame, and

$$\varepsilon_i = \begin{cases} 1 & \text{for} \quad \text{revolute joint}, \\ 0 & \text{for} \quad \text{prismatic joint}, \end{cases}$$
$$(0 \quad \hat{\mathbf{w}}_0^k) = \left(\prod_{i=1}^{k} {}_{i}^{i-1}\mathbf{r} \right) \cdot (0 \quad \hat{\mathbf{w}}_k^k) \cdot \left(\prod_{i=1}^{k} {}_{i}^{i-1}\mathbf{r} \right)^{*}. \tag{29}$$

Once again it is underlined that the cuaternion product is associative but not commutative.

7 Examples

Two simple preparatory cases are first presented. They refer to planar structures such as Rotation Rotation (RR) and Rotation Translation (RT).

Since the principles from which the two alternatives of the proposed formalism start (included in steps 1 and 2 of the algorithm) are identical, this allows the idea of carrying out their comparison. The solution to the direct geometry problem obtained by both approaches will be presented as follows.

Figure 5 illustrates an RR type structure, which according to the first two steps of the presented formalism contains the four coordinate systems {0 = 1}, {2}, {e} attached to the base, the elements and the gripper, respectively. There are also illustrated here the components of the translation vector that connects the mentioned coordinate systems, l_1 and l_2. The axes of the two joints are perpendicular to the plane of the structure $\hat{z} \equiv \hat{u}^1 \equiv \hat{u}^2$. Figure 5 also shows the two generalized angular variables q_1 and q_2. The use of the homogeneous transformations for the RR structure is synthesized in Table 1.

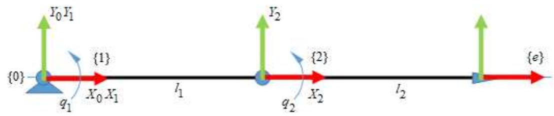


Figure 5
The RR structure

The approach proposed in this paper modifies the mathematical apparatus, i.e., it uses quaternions instead of homogeneous transformations in step 3. Table 2 gives the application of the transformation equations (15) to the RR structures. The transformations as well as the elements that are included in these transformations are defined. The quaternions are actually computed for each transformation.

Table 1
Homogeneous transformations for the RR structure

No.	The parameters	The transformations
1	${}^0_1\mathbf{Q} = \mathbf{R}_\mathbf{z}(q_1)$ ${}^0_1\mathbf{P} = [0 \quad 0 \quad 0]$	${}^0_1\mathbf{T} = \begin{bmatrix} c_{q_1} & -s_{q_1} & 0 & 0 \\ s_{q_1} & c_{q_1} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$
2	${}^1_2\mathbf{Q} = \mathbf{R}_z(q_2)$ ${}^1_2\mathbf{P} = [l_1 \quad 0 \quad 0]$	${}^1_2\mathbf{T} = \begin{bmatrix} c_{q_2} & -s_{q_2} & 0 & l_1 \\ s_{q_2} & c_{q_2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$
e	${}^2_e\mathbf{Q} = \mathbf{I}$ ${}^6_e\mathbf{P} = [l_2 \quad 0 \quad 0]$	${}^6_e\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & l_2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Table 2
Use of quaternions for the RR structure

The transformation	**r** quaternions	**t** quaternions
Transformation e-0		
${}^0_e\mathbf{T} = [{}^0_1\mathbf{r} \quad \mathbf{1}] \cdot \begin{bmatrix} {}^1_e\mathbf{T} & \mathbf{0} \\ \mathbf{0} & {}^0\mathbf{t} \end{bmatrix} \cdot \begin{bmatrix} {}^0_1\mathbf{r}^* \\ \mathbf{1} \end{bmatrix}$	${}^0_1\mathbf{r} = \left(\cos\left(\frac{q_1}{2}\right) \quad \hat{\mathbf{u}}^1_1 \sin\left(\frac{q_1}{2}\right) \right)$ $\hat{\mathbf{u}}^1_1 = [0 \quad 0 \quad 1]$	${}^0\mathbf{t} = \mathbf{0}$
Transformation e-1		
${}^1_e\mathbf{T} = [{}^1_2\mathbf{r} \quad \mathbf{1}] \cdot \begin{bmatrix} {}^2_e\mathbf{T} & \mathbf{0} \\ \mathbf{0} & {}^1\mathbf{t} \end{bmatrix} \cdot \begin{bmatrix} {}^1_2\mathbf{r}^* \\ \mathbf{1} \end{bmatrix}$	${}^0_1\mathbf{r} = \left(\cos\left(\frac{q_2}{2}\right) \quad \hat{\mathbf{u}}^2_2 \sin\left(\frac{q_2}{2}\right) \right)$ $\hat{\mathbf{u}}^2_2 = [0 \quad 0 \quad 1]$	${}^1\mathbf{t} = (0 \quad l_1 \quad 0 \quad 0)$
Transformation e-2		
${}^2_e\mathbf{T} = [\mathbf{1} \quad \mathbf{1}] \cdot \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & {}^2\mathbf{t} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix}$		${}^2\mathbf{t} = (0 \quad l_2 \quad 0 \quad 0)$
Transformation 2-0		
${}^0_2\mathbf{T} = [{}^0_1\mathbf{r} \quad \mathbf{1}] \cdot \begin{bmatrix} {}^1_2\mathbf{T} & \mathbf{0} \\ \mathbf{0} & {}^0\mathbf{t} \end{bmatrix} \cdot \begin{bmatrix} {}^0_1\mathbf{r}^* \\ \mathbf{1} \end{bmatrix}$	${}^0_1\mathbf{r} = \left(\cos\left(\frac{q_1}{2}\right) \quad \hat{\mathbf{u}}_1 \sin\left(\frac{q_1}{2}\right) \right)$ $\hat{\mathbf{u}}_1 = [0 \quad 0 \quad 1]$	${}^0\mathbf{t} = \mathbf{0}$
Transformation 2-1		
${}^1_2\mathbf{T} = [\mathbf{1} \quad \mathbf{1}] \cdot \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & {}^1\mathbf{t} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix}$		${}^1\mathbf{t} = (0 \quad l_1 \quad 0 \quad 0)$

Figure 6 illustrates a RT type structure, which according to the first two steps of the presented formalism, contains the four coordinate systems {0 = 1}, {2}, {e} attached to the base, the elements and the gripper, respectively. In addition, Figure 6 highlights the components of the translation vector that connects the mentioned coordinate systems, l_1 and l_2, the axes of the two joints $\hat{\mathbf{z}} = \hat{\mathbf{u}}^1$ and $\hat{\mathbf{x}} = \hat{\mathbf{u}}^2$, and the two generalized angular variables q_1 and q_2. The use of the homogeneous transformations for the RT structure is synthesized in Table 3.

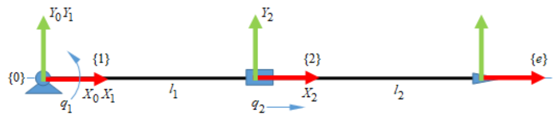


Figure 6
The RT structure

Table 3
Use of quaternions for the RT structure

No.	The parameters	The transformations
1	${}^0_1\mathbf{Q} = \mathbf{R}_z(q_1)$ ${}^0_1\mathbf{P} = [0 \quad 0 \quad 0]$	${}^0_1\mathbf{T} = \begin{bmatrix} c_{q_1} & -s_{q_1} & 0 & 0 \\ s_{q_1} & c_{q_1} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$
2	${}^1_2\mathbf{Q} = \mathbf{I}$ ${}^1_2\mathbf{P} = [l_1 + q_2 \quad 0 \quad 0]$	${}^1_2\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & l_1 + q_2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$
e	${}^2_e\mathbf{Q} = \mathbf{I}$ ${}^6_e\mathbf{P} = [l_2 \quad 0 \quad 0]$	${}^6_e\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & l_2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Table 4 gives the application of the transformation equations (15) to the RT structures. The transformations as well as the elements that are included in these transformations are defined. The quaternions are computed, as in the previous example, for each transformation.

Table 4
Use of quaternions for the RR structure

The transformation	**r** quaternions	**t** quaternions
Transformation e-0		
${}^0_e\mathbf{T} = [{}^0_1\mathbf{r} \quad 1] \cdot \begin{bmatrix} {}^1_e\mathbf{T} & \mathbf{0} \\ \mathbf{0} & {}^0\mathbf{t} \end{bmatrix} \cdot \begin{bmatrix} {}^0_1\mathbf{r}^* \\ \mathbf{1} \end{bmatrix}$	${}^0_1\mathbf{r} = \left(\cos\left(\frac{q_1}{2}\right) \quad \hat{\mathbf{u}}^1_1 \sin\left(\frac{q_1}{2}\right) \right)$ $\hat{\mathbf{u}}^2_2 = [0 \quad 0 \quad 1]$	${}^0\mathbf{t} = \mathbf{0}$

Transformation e-1		
${}^{1}_{e}\mathbf{T} = [{}^{1}_{2}\mathbf{r} \quad \mathbf{1}] \cdot \begin{bmatrix} {}^{2}_{e}\mathbf{T} & \mathbf{0} \\ \mathbf{0} & {}^{1}\mathbf{t} \end{bmatrix} \cdot \begin{bmatrix} {}^{1}_{2}\mathbf{r}^{*} \\ \mathbf{1} \end{bmatrix}$	${}^{0}_{1}\mathbf{r} = (1 \quad 0 \quad 0 \quad 0) = \mathbf{1}$	${}^{1}\mathbf{t} = (0 \quad q_2 + l_1 \quad 0 \quad 0)$
Transformation e-2		
${}^{2}_{e}\mathbf{T} = [\mathbf{1} \quad \mathbf{1}] \cdot \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & {}^{2}\mathbf{t} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix}$		${}^{2}\mathbf{t} = (0 \quad l_2 \quad 0 \quad 0)$
Transformation 2-0		
${}^{0}_{2}\mathbf{T} = [{}^{0}_{1}\mathbf{r} \quad \mathbf{1}] \cdot \begin{bmatrix} {}^{1}_{2}\mathbf{T} & \mathbf{0} \\ \mathbf{0} & {}^{0}\mathbf{t} \end{bmatrix} \cdot \begin{bmatrix} {}^{0}_{1}\mathbf{r}^{*} \\ \mathbf{1} \end{bmatrix}$	${}^{0}_{1}\mathbf{r} = \left(\cos\left(\frac{q_1}{2}\right) \quad \hat{\mathbf{u}}_1 \sin\left(\frac{q_1}{2}\right) \right)$ $\hat{\mathbf{u}}_1 = [0 \quad 0 \quad 1]$	${}^{0}\mathbf{t} = \mathbf{0}$
Transformation 2-1		
${}^{1}_{2}\mathbf{T} = [\mathbf{1} \quad \mathbf{1}] \cdot \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & {}^{2}\mathbf{t} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix}$		${}^{2}\mathbf{t} = (0 \quad l_2 \quad 0 \quad 0)$

The translated reference frame formalism is described in [20] along with its comparison to the DH formalism for a PUMA type robot structure. This time, considering the same structure illustrated in Figure 7, a comparison will be carried out as follows between the two mathematical formulations of the same formalism, namely that proposed in [20] and the formulation proposed in this paper. Therefore, an answer is given to the question of comparing the mathematical apparatus of homogeneous transformations with that of quaternions.

Table 5 given in [22] describes the parameters and homogeneous transformations in accordance with (2). The second column specifies the elements of the homogeneous transformation, i.e., the rotation matrices and the translation vectors. Each row in the table refers to the transformation from the reference frame $\{i\}$ to the reference frame $\{i{-}1\}$.

Table 6 given in [22] gives details on the transformations defined using quaternions. The rows in the table offer, similar to the rows in Tables 2 and 4, the transformation applied, and the quaternions used by it as well.

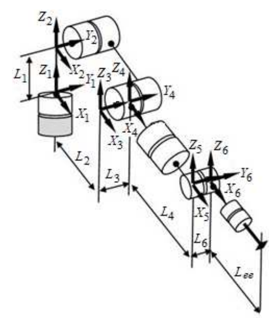


Figure 7
The RT structure

8 Simulation Results

The Matlab program used in the automation of the formalism proposed is freely available at https://autocarsim.com/use-of-quaternions/. The program refers to the PUMA type robot structure. The use of the program is described in the dedicated script.

The transformations specified in Table 6 are implemented in a Matlab program, which allows determining the position of each element, validating the theoretical results, and obtaining the graphical representation of the structure configuration. Table 7 given in [22] describes a sample of the results of several simulations conducted for different angular variables.

The figures included in Table 7, describe the reference frame of the base and that of the effector (the CAD convention was used to color the axes), the elements of the structure and the joints (circles). The values of the angular variables are specified in the title of each figure.

Conclusions

This paper further developed the authors' previous approach, regarding a new formalism for solving the direct geometry problem of manipulators suggested in [20]. The advantages of the proposed formalism are the simplicity to define the reference frames of the elements, and the simplicity of defining the mathematical transformations. More precisely, it has been shown in [20] that a set of six generic transformations is available to be easily customized to any transformation used by the formalism. The paper brings novelties in the mathematical apparatus of the formalism, i.e., it suggests an alternative to replace the computation of homogeneous transformations with that of quaternions.

The quaternions are employed in solving the direct geometry problem and the Jacobian matrix calculation. The proposed formalism was exemplified and compared in three structures, namely RR, RT and PUMA type. The examples also include the conversion of algorithms into computer programs, which allowed the simulation of solving the direct geometry problem.

Future research will be focused on the application of authors' formalism to the design of control systems for robots in real-world applications. Such applications are popular in several fields including path planning [23] [24], electric vehicles [25] [26], haptic interfaces [27], manufacturing processes [28-30]. Several optimization algorithms can be used in these applications for performance improvement and also reducing the heuristics in the design, as, for example, cellular genetic algorithms [31], tabu search based on quantum computing [32], Bacterial Foraging Optimization Algorithms [33] [34], Clonal Selection Algorithms [34], Grey Wolf Optimizers [35], Particle Swarm Optimization Algorithms [36], Slime Mold Algorithms [29], classical optimization algorithms

[37], multi-parametric quadratic programming [38], and Metaheuristic Algorithms with parameter adaptation [39].

Acknowledgement

This work was supported by a grant of the Romanian Ministry of Education and Research, CNCS - UEFISCDI, project number PN-III-P4-ID-PCE-2020-0269, within PNCDI III, the Széchenyi István University, Győr, Hungary, and a grant of the Romanian Ministry of Research, Innovation and Digitization, CNCS/CCCDI – UEFISCDI, project number PN-III-P2-2.1-PED-2019-4366 (431PED) within PNCDI III.

References

[1] G. Booch, R.A. Maksimchuk, M. W. Engle, B. J. Young, J. Conallen, K. A. Houston: Object-Oriented Analysis and Design with Applications, 3rd Edition. Upper Saddle River, NJ: Addison-Wesley Professional, 2007

[2] Q. Zhang, Y.-F. Zhang, S.-Y. Qin: The object oriented analysis and modeling for obstacle avoidance of a behavior-based robot, in Proceedings of 2007 IEEE International Conference on Systems, Man and Cybernetics, Montreal, QC, Canada, 2007, pp. 1545-1550

[3] M. Rickert, A. Gaschler: Robotics library: An object-oriented approach to robot applications, in Proceedings of 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems, Vancouver, BC, Canada, 2017, pp. 733-740

[4] P. Ramaraj: Robots that help humans build better mental models of robots, in Proceedings of 2021 ACM/IEEE International Conference on Human-Robot Interaction, Virtual, 2021, pp. 595-597

[5] S. A. Bortoff: Object-oriented modeling and control of delta robots, in Proceedings of 2018 IEEE Conference on Control Technology and Applications, Copenhagen, Denmark, 2018, pp. 251-258

[6] P. Fritzon: Principles of Object Oriented Modeling and Simulation with Modelica 3.3: A Cyber-Physical Approach, 2nd Edition. Piscataway, NJ: IEEE Press-Wiley, 2015

[7] C. Pozna: Modelarea robotilor cu post fix (in Romanian). Bucharest: MatrixRom, 2013

[8] J. Denavit, R. S. Hartenberg: A kinematic notation for lower-pair mechanisms based on matrices, Journal of Applied Mechanics, Vol. 22, No. 2, 1955, pp. 215-221

[9] R. S. Hartenberg, J. Denavit: Kinematic Synthesis of Linkages. New York, NY: McGraw-Hill, 1964

[10] R. P. Paul: Robot Manipulators: Mathematics, Programming, and Control. Cambridge, MA, London: The MIT Press, 1981

[11] W. Khalil, J. F. Kleinfinger: A new geometrical notation for open and closed loop robots, in Proceedings of 1986 IEEE International Conference on Robotics and Automation, San Francisco, CA, USA, 1986, pp. 1174-1180

[12] J. J. Craig: Introduction to Robotics: Mechanics and Control, 3rd Edition. Upper Saddle River, NJ: Pearson Education, 2005

[13] P. Corke: Robotics, Vision and Control. Berlin, Heidelberg: Springer-Verlag, 2011

[14] RoboAnalyzer: 3D Model Based Robotics Learning Software, posted at http://www.roboanalyzer.com/ and accessed at July 12, 2021

[15] G. Gogu, P. Coiffet: Representation du movement des corps solids (in French). Paris: Hermes, 1996

[16] G. Gogu, P. Coiffet, A. Barraco: Representation des deplacements des robots (in French). Paris: Hermes, 1997

[17] R. Murray, Z. Li, & S. Sastry: A Mathematical Introduction to Robotic Manipulation. CRC Press, 1994

[18] J. J. Craig: Introduction to Robotics: Mechanics and Control. Pearson Education, 2004

[19] B. Siciliano, L. Sciavicco: Robotics: Modelling, Planning and Control. Springer, 2008

[20] C. Pozna, R.-E. Precup: On a translated frame-based approach to geometric modeling of robots, Robotics and Autonomous Systems, Vol. 91, 2017, pp. 49-58

[21] W. R. Hamilton: Elements of Quaternions. Cambridge: Cambridge University Press, 1899

[22] C. Pozna, R.-E. Precup: Appendix of the paper C. Pozna and R.-E. Precup, "On the Use of Quaternions in the Translated Reference Frame Formalism", Acta Polytechnica Hungarica, 2023, posted at https://www.aut.upt.ro/~rprecup/Appendix_paper_APH_2023_Pozna_Precup.pdf and accessed at November 12, 2022

[23] G. Klančar, S. Blažič: Optimal constant acceleration motion primitives, IEEE Transactions on Vehicular Technology, Vol. 68, No. 9, 2019, pp. 8502-8511

[24] J. Vaščák, J. Hvizdoš, M. Puheim: Agent-based cloud computing systems for traffic management, in Proceedings of 2016 International Conference on Intelligent Networking and Collaborative Systems, Ostrava, Czech Republic, 2016, pp. 73-79

[25] Z. C. Johanyák: A modified particle swarm optimization algorithm for the optimization of a fuzzy classification subsystem in a series hybrid electric

vehicle, Technicki Vjesnik - Technical Gazette, Vol. 24, No. 2, 2017, pp. 295-301

[26] A. Lucchini, S. Formentin, M. Corno, D. Piga, S. M. Savaresi: Torque vectoring for high-performance electric vehicles: a data-driven MPC approach, IEEE Control Systems Letters, Vol. 4, No. 3, 2020, pp. 725-730

[27] N. Ando, P. Korondi, H. Hashimoto: Networked telemicromanipulation systems “Haptic Loupe”, IEEE Transactions on Industrial Electronics, Vol. 51, No. 6, 2004, pp. 1259-1271

[28] R. E. Haber, J. R. Alique, A. Alique, J. Hernández, R. Uribe-Etxebarria: Embedded fuzzy-control system for machining processes: Results of a case study, Computers in Industry, Vol. 50, No. 3, 2003, pp. 353-366

[29] R.-E. Precup, R.-C. David, R.-C. Roman, E. M. Petriu, A.-I. Szedlak-Stinean: Slime mould algorithm-based tuning of cost-effective fuzzy controllers for servo systems, International Journal of Computational Intelligence Systems, Vol. 14, No. 1, 2021, pp. 1042-1052

[30] V. M. Nejković, M. S. Milićević, G. Janaćković, M. Grozdanović: Application of fuzzy analytic hierarchy process to inductive steel tube welding, Romanian Journal of Information Science and Technology, Vol. 25, No. 1, 2022, pp. 3-19

[31] E. Osaba, J. Del Ser, A. D. Martinez, J. L. Lobo, F. Herrera: AT-MFCGA: An adaptive transfer-guided multifactorial cellular genetic algorithm for evolutionary multitasking, Information Sciences, Vol. 570, 2021, pp. 577-598

[32] E. Osaba, E. Villar-Rodriguez, I. Oregi, A. Moreno-Fernandez-de-Leceta: Hybrid quantum computing-tabu search algorithm for partitioning problems: preliminary study on the traveling salesman problem, in Proceedings of 2021 IEEE Congress on Evolutionary Computation, Kraków, Poland, 2021, pp. 351-358

[33] S. I. Bejinariu, H. Costin: Medical image registration by means of a bio-inspired optimization strategy, Computer Science Journal of Moldova, Vol. 20, No. 2, 2012, pp. 178-202

[34] S. I. Bejinariu, H. Costin, F. Rotaru, C. Niţă, R. Luca, C. Lazăr: Parallel processing and bio-inspired computing for biomedical image registration, Computer Science Journal of Moldova, Vol. 22, No. 2, 2014, pp. 253-277

[35] I. A. Zamfirache, R.-E. Precup, R.-C. Roman, E. M. Petriu: Neural network-based control using actor-critic reinforcement learning and grey wolf optimizer with experimental servo system validation, Expert Systems with Applications, Vol. 225, 2023, paper 120112

[36] D. Sánchez, P. Melin, O. Castillo: Comparison of particle swarm optimization variants with fuzzy dynamic parameter adaptation for modular

granular neural networks for human recognition, Journal of Intelligent and Fuzzy Systems, Vol. 38, No. 3, 2020, pp. 3229-3252

[37] D. Singh, A. Shukla: Manifold optimization with MMSE hybrid precoder for Mm-Wave massive MIMO communication, Romanian Journal of Information Science and Technology, Vol. 25, No. 1, 2022, pp. 36-46

[38] Z. Preitl, R.-E. Precup, J. K. Tar, M. Takács: Use of multi-parametric quadratic programming in fuzzy control systems, Acta Polytechnica Hungarica, Vol. 3, No. 3, 2006, pp. 29-43

[39] O. Castillo, P. Melin, E. Ontiveros, C. Peraza, P. Ochoa, F. Valdez, J. Soria: A high-speed interval type 2 fuzzy system approach for dynamic parameter adaptation in metaheuristics, Engineering Applications of Artificial Intelligence, Vol. 85, 2019, pp. 666-680

Quality Improvement in Education, based on Student Feedback

Gabriella Farkas, Erzsébet Ancza, Georgina Nóra Tóth

Óbuda University, Bécsi út 96/b, 1034 Budapest, Hungary; farkas.gabriella@uni-obuda.hu, ancza.erzsebet@uni-obuda.hu, toth.georgina@uni-obuda.hu

Abstract: This work deals with the improvement of higher education processes and the identification of solution guidelines, for problem areas analyzed from student feedback. The purpose of this article is to analyze and evaluate student opinions and suggestions. To this end, on the one hand, we conducted an online survey, with the aim of collecting educational improvement proposals, from the student's perspective. On the other hand, we investigated the applicability of this method, which is well known and effectively used in the field of social science. Using the method, we examined what student expectations can be formulated, in order to improve and develop the quality of education. In the course of factor analysis, we planned to define those expectations and attitudes that will be of effective help, in the successful improvement and development of the teaching work, which can ultimately mean long-term success for the student.

Keywords: Quality improvement; Student feedback; Q-method; Factor analysis

1 Introduction and Purpose

Determining quality is basically a difficult task, especially in the field of education or training. According to the requirements of Standards and Guidelines for Quality Assurance in the European Higher Education Area (ESG), quality can be formulated as a result of the interaction of the instructor, the student and the institutional learning environment. Quality assurance in higher education means among other things, providing a learning environment in which the content of the trainings, learning opportunities and infrastructure are fit for purpose. To do this it is essential to take into account the needs and expectations of students and other stakeholders as well as society (stakeholders). [1] There are several types of opinion surveys in our institution. Students are given the opportunity to express their opinions both during training (Opinion request sheet) and directly upon graduation (final examination student opinion survey). Due to the character of the service, it is also important to know how graduate students can benefit from their acquired knowledge in the longer term. Therefore, the 1st, 3rd and 5th years of the student's follow-up (Graduate Career Tracking System), is important feedback for the higher education institution. To improve the quality of education we can receive important information from student and teacher feedbacks but at the same time we need to know the expected needs. There are no needs assessments of this kind and it is hard to imagine that those entering higher education could describe this. There are methods, tools which can help to determine these expectations (student attitudes). In our research we chose the Q-method which is becoming increasingly popular in social science research and has already been applied in several ways in educational and pedagogical fields.

2 Literature Review

The Q-method is suitable for researching the range and diversity of subjective experiences, points of view and opinions. However, it facilitates the identification of similarities, the construction of broader categories of the phenomenon under study as well as the exploration of patterns and relationships within and between these categories. The method has been used in various fields and its application is increasing from time to time. [2] Therefore, the main goal of the Q-method is to form types and attitudes from subjective points of view towards a thing, e.g., towards political attitudes, advertisements, environment. The creation of the Q-method is associated with the name of the English psychologist-physicist William Stephenson. The development of the method was based on the desire to provide a scientific framework for determining subjectivity. [3] The Q-method assists in the systematic examination of human subjective opinion-forming and decision-making. [4] A qualitative analysis evaluating subjective opinions that uses a quantitative approach to factor analysis based on statements ranked by respondents. [5]

The Q-method is not a tool designed to reach agreement but a means of examining the diversity of opinions on topics about which there has been a more or less mature debate. [6] This is due to forcing respondents to choose or rank between opinions and statements. [7] The Q-method is also an effective tool for managing conflicts and identifying expectations for future directions. [8]

It can be considered a reverse factor analysis, according to which it analyzes the people themselves not their characteristics. The method does not focus on differences between individuals but on differences within individuals. The mathematical basis of the method is the same as the mathematical basis of factor analysis. With the Q-method a relatively large number of statements can be evaluated with the involvement of a very small number of individuals. The correlation coefficients calculated by this method show what the correlation is between people. [9]

The mathematical background of the procedure is provided by calculation of the correlation and modified (inverse) factor analysis with the help of which we can create common groups and factors from similar opinions [10]. Relatively few, usually from 10 to 50 individuals are involved in the study, who are selected on the basis of specific criteria. Due to its distinctive features, the Q-methodology combines qualitative and quantitative research procedures, combining the advantages of both research methods. [11] [12] The method is a scientifically based philosophical and statistical framework and approach with the help of which various individual subjectivities can be quantitatively analyzed and evaluated using objective scientific means. [13] In a sense the Q-methodology combines the power of qualitative and quantitative research traditions. [14]

The Q-method uses a set of statements (Q sample) describing a certain topic, formulating statements. [15] Within this framework study participants (P-set) should be selected and invited to participate in a group from among these statements. They are asked to evaluate a heterogeneous set of claims in terms of agreement and disagreement. [16]

The application areas of the Q-method are wide-ranging but it is typically used in areas where customer expectations can be analyzed qualitatively and quantitatively:

- In opinion polls, e.g., the definition of political leanings
- For behavioral studies of health patients in clinical psychology, for analysis of social work [17] [18]
- In education, pedagogical analyses, qualitative examination of opinions, expectations [19]
- In media research, marketing research, examining customer attitudes

3 Methodology and Approach

Our studies are based on the results of surveys conducted among students who passed the final exam in the 2021 and 2022 academic years. During this period, there were typically attendance and online final exams, so the surveys were conducted by students both online and personal. In the case of our institution the surveys were carried out with a completion result of 14-63%. The questionnaire used is structured as follows:

1. Evaluation of educational activities of specialization (satisfaction [1-6] and importance [1-6]); 8 questions
2. Evaluation of general educational infrastructure (satisfaction [1-6] and importance [1-6]); 7 questions

3. Evaluation of the means to support the teaching of basic and vocational subjects (satisfaction [1-6] and importance [1-6]); 5 questions
4. Evaluation of activities related to the organization of education (satisfaction [1-6] and importance [1-6]); 11 questions
5. Overall, I got what I expected from the training [1-6]
6. If I were to start my studies again in this field, I would choose the same University again [1-6]
7. Strengths-Weaknesses

The results evaluation sheets are prepared in a separate way for Institutes, which does not allow for the separate evaluation of individual courses (BSc, MSc, special postgraduate training scheme) or specializations. The formulation of quality improvement options that can be formulated on the basis of the results of the surveys can be done at different levels. In the questionnaire compiled for student review, the individual questions and the evaluations given to them can be linked to different levels of intervention (responsibility), these are: instructor, lecturer of the subject (L), Institute (I), head of the training (H), Faculty (F), University (U).

In our research, we treated the results of the surveys carried out during the period under review as a whole and in general and based them on our further studies. Of the listed intervention levels, only the subject areas at teaching (L) and Institute (I) level were analyzed. In general, the following areas were identified by students as problems:

- Professional and pedagogical preparedness of teachers
- The relationship between faculty and students of the training/specialization
- Organization of classes (lectures, seminars, labs)
- Effective filling of training time with subjects
- Interdependence of subjects
- Organization of exams
- The matching of the final exam with what they have learned

We conducted an online survey among our students about what educational development opportunities and solutions they find useful to improve the listed problems. During the filling out, they could choose more than one of the development options defined, and they could also formulate their own proposal, A total of 88 students participated in the survey. Of those surveyed, 53.4% are BSc in mechanical engineering, 35.2% are BSc in mechatronics engineering and 11.4% are BSc in safety engineering. 70.5% are full-time students and 29.5% are correspondence students. The results of the questions and answers in the questionnaire on the development of teaching activities, evaluation and collection of suggestions were as follows:

1. In your opinion, what are some ways the instructors could improve communication with students?
2. In your opinion, what are some ways the instructors could improve their professional and pedagogical skills?
3. In your opinion, how can the instructors improve the educational support tools (notes, aids, presentations, etc.) of their subjects?
4. In your opinion, what are some ways the instructors can improve the educational tools and techniques that help to get to know the profession and transfer knowledge during the training?
5. In your opinion, what are some ways the instructors can improve the organization of classes (lecture, practice, lab)?
6. In your opinion, what are some ways to improve the accountability of subjects and the organization of exams?
7. In your opinion, what are some ways to improve the structure of subjects?
8. In your opinion, what are some ways to improve the interdependence of subjects?
9. In your opinion, what are some ways to improve student involvement in research?
10. In your opinion, what are some ways to improve the educational infrastructure in each training program/specialization?
11. In your opinion, what are some ways to improve the infrastructure related to services at the Faculty?
12. In your opinion, what are some ways to improve student counseling and administration at the Faculty?

Based on the student responses, in the rest of our research we would like to examine the expectations and attitudes of the students using the Q-method. The steps in this method are shown in Figure 1.

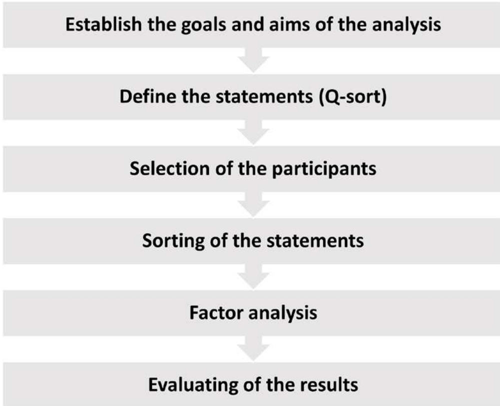


Figure 1
Steps of the practical application of the Q-method

In the first step the purpose of our research was determined. In the course of our studies, we are looking at what groups of student expectations can be defined using the Q-method along the lines of educational development. To do this in the next step we have identified 34 statements, these are:

1. It is useful to create and operate a complaint interface where students can make their comments.
2. For all subjects, it is expected that the assessment method is in accordance with the requirements.
3. There should be exam topics for the subjects and they should fit the learning materials.
4. Involving external lecturers in education effectively improves knowledge of the profession.
5. Organizing and holding small group sessions and exercises is an effective way to improve communication.
6. As a student I am satisfied to participate in company visits, where you can get acquainted with industrial practices.
7. In a well-functioning educational infrastructure, there are open labs available to students at any time.
8. In a well-functioning educational infrastructure, it is essential to have software that is also used in industry.
9. In a well-functioning educational infrastructure, it is necessary that the technical equipment of the lecturers and classrooms is at a high level.
10. In a well-functioning educational infrastructure, a room provided for project tasks is essential for students.

11. In an effective educational infrastructure, it is necessary to have modern machines, tools, instruments in laboratories.
12. In an effective educational infrastructure, it is necessary to have labs suitable for practical sessions.
13. The basis of the organized lessons (lectures, practices) is that the instructor knows and properly manages the technical devices.
14. Online tests effectively help to complete the tests for the purpose of practice and self-checking.
15. Instructors should provide with an adequate number of consultation opportunities for students.
16. Instructors should regularly participate in educational methodology training.
17. The professionalism of instructors can be effectively developed if they strengthen their industrial relations.
18. The theoretical curriculum should be supported by appropriate textbooks and recommended literature.
19. Notes, aids and examples should be prepared for each subject, which are also available in digital form.
20. Transparent, outlined, illustrative presentations help to understand the curriculum, therefore it is an effective educational support tool.
21. The curriculum of the given subject should be adapted to the training goals of the given specialty/training/specialization.
22. Study administration is facilitated by the mobile application developed for this purpose.
23. For academic administration, it is important to have up-to-date information on the websites, which is easily accessible.
24. The schedule of the subject must include the learning materials covered during the semester.
25. The interdependence of subjects is facilitated by increasing the number of practical lessons.
26. Learning materials will help you learn the profession much better if they contain real, industrial examples.
27. The theoretical and practical parts of the curriculum should be balanced.
28. The learning materials to be uploaded to Moodle should be uploaded on time, and their availability should be ensured by the instructor.
29. Learning materials often contain outdated, old examples or worse, no examples. Their topicality contributes to the development of educational support tools.

30. Students would prefer to participate in research if there was a thesis topic they could develop.
31. One of the best ways to motivate students is if they can earn extra points during lessons.
32. Students could be more involved in research if there were suitable labs for this.
33. When teaching the practical curriculum, it should be a truly practical (lab) lesson.
34. Create a learning material of examples for practical tasks with real, industrial examples.

Then we defined the Q-table in which the 34 statements can be placed (Figure 2).

Most disagreed				Neutral	Most agreed			
-4	-3	-2	-1	0	+1	+2	+3	+4

Figure 2
The structure of the Q-table

We asked 15 participants to settle the statements. Respondents were selected to reflect the opinions of those interested in a particular topic, but they did not represent the entire student community. The respondent students were full-time undergraduate students. Their task was to place the statements in the Q-table. First it was necessary to divide the statements into three groups: agree, disagree, indifferent (neutral). Each score then had to be sorted according to how much you agree or disagree with that statement. After sorting, we used PQMethod - 2.35 (Mar 2014) by Peter Schmolck to perform factor analysis. [20] Our program settings were:

- Title of Study: My first analysis of the development of education
- Column Range: -4 TO 4
- Depth of Columns: 2 3 4 5 6 5 4 3 2
- Sorts Entered: 15

To perform a factor analysis, the program guide was used. Based on this the main component analysis was carried out, the factors were determined and analyzed with the factor analysis. [21] In the following we present our results for both the online survey and the Q-method.

4 Results

The 15 completed Q-sorts (individual evaluation table) were then processed using the already referenced Schmolz's PQMethod computer program. The program trained typical Q-classes (factors), based on their similarity or difference in individual preference sequences (Q-classes). Using the Q-method, we originally created eight factors from individual rankings, which were reduced by rotations to 4 factors. The Q factor analysis identified the following types of students:

- Students who require practice-oriented instruction typically consider it important to have real-world hands-on (labs) sessions that include learning materials that contain real industry examples. In this expectation, there is a need for external (industrial partner) involvement in education, supporting the theory, with applications presented in industrial practice.
- Students who require an appropriate learning environment and who consider it important to have computer labs with the appropriate software and to use open labs during their studies. They expect labs equipped with machines and devices that also perform tasks for teaching and research purposes.
- Students who require well qualified teachers formulate as expectations the organization of classes, the timely availability of learning materials, the alignment of requirements with the subject and the teacher's educational methodological preparation. They require a combination of traditional and modern forms of education, in which they can feel more motivated. They consider it important that the exam requirements are aligned with the curriculum.
- Students who need up-to-date, informational (communication) background. They place great emphasis on the quick, easy access to real-world information that is important to them. They require up-to-date information on the websites operated by the institutions, easy, fast search possibilities and descriptions to help with certain administrative procedures. Not only academic administration, but also communication with instructors is considered important, and in the event of a problem, they should receive appropriate help and quick and efficient operation of complaint handling.

In the following we analyzed student opinions based on responses to the educational development questionnaire:

1. In your opinion, what are some ways instructors could improve communication with students?

The largest number of votes was 58% for small, group lessons and practices, followed by 51.1% by the functioning, fast complaint interface, where students can indicate their problems, 29% for more face-to-face consultations with students, 26.1% for educational-methodological training for instructors, and also 26.1% for class visits in which the instructor's ability as a lecturer is assessed.

2. In your opinion, what are some ways instructors could improve their professional and pedagogical skills?

In this issue 64.8% of the votes were given to strengthening industrial relations, followed by 44.3% participation in educational-methodological training for trainers, 42% participation in professional conferences, 38.6% participation in pedagogical training for trainers, and 33% in the processing of professional journals and textbooks.

3. In your opinion, how can instructors improve the educational support tools (notes, aids, presentations, etc.) of the subjects?

In this topic 75% of the votes were given to the learning materials available on Moodle, followed by 70.5% if possible, a book of problems, 65.9% learning materials available in electronic and printed form, 56.8% notes to be updated with current events, including follow-up questions and 52.3% to provide transparent, outline and illustrative presentations.

4. In your opinion, what are some ways instructors can improve the educational tools and techniques that help to get to know the profession and transfer knowledge during the training?

The most votes 80.7% was given by real-world industrial examples and practices in the teaching materials, followed by 69.73% for the involvement of an external lecturer and 65.9% for the organization of study trips to industrial partners.

5. In your opinion, what are some ways instructors can improve the organization of classes (lecture, practice, lab)?

In this topic 65.9% of the votes were given to the proper use of technical means, followed by 64.8% for student motivation, 38.6% for group work and consultation opportunities and 33% for accurate start and end of classes.

6. In your opinion, what are some ways to improve the accountability of subjects and the organization of exams?

70.5% of the votes were given to the accountability for the requirements of the subject or in the case of exams for the exam topics to fit the curriculum, followed by 63.6% for the use of online tests in Moodle and 45.5% for the correction of midterm tests within a fixed deadline.

7. In your opinion, what are some ways to improve the structure of subjects?

In this case the advices were in order: the teaching materials should be aligned with the training goals of the given training program/specialization (69.3%), the topics should include the teaching materials involved during the semester (65.9%), the industrial and professional expectations should be taken into account in the structure of the curriculum (65.9%), the teaching of the practical curriculum should be a truly practical (lab) occupation (63.6%), the theoretical and practical parts of the curriculum should be balanced (51.1%).

8. In your opinion, what are some ways to improve the interdependence of subjects?

Student opinions for the development of the teaching were in order: increasing the number of practical lessons (55.7%), reducing the pre-requirements for subjects (37.5%), increasing the number of exam courses (35.2%), taking subjects without pre-requirements (34.1%).

9. In your opinion, what are some ways to improve student involvement in research?

In this issue the suggestions were in order: providing thesis topics related to research tasks (71.6%), providing laboratories for students for research tasks (59.1%), providing several Scientific Student Circle topics related to research topics (33%).

10. In your opinion, what are some ways to improve the educational infrastructure in each training program/specialization?

The most votes (84.1%) were given to the provision and education of modern software also used in industry, followed by the development and modernization of laboratory equipment and instruments by 73.9%, the development of laboratories suitable for practical sessions and the provision of open laboratories for students (65.9%).

11. In your opinion, what are some ways to improve the infrastructure related to services at the Faculty?

In this case the most evaluated things were in order: improving the technical equipment of lecturers and classrooms (67%), providing a room suitable for project tasks (62.5%), providing a computer room outside the classroom (50%).

12. In your opinion, what are some ways to improve student counselling and administration at the Faculty?

In this question the order was between the devices: development of the Faculty website, easier access of information (60.2%), provision of a room suitable for project tasks (62.5%), provision of a computer room outside the classroom (50%).

Conclusions

In this work, we presented the educational development opportunities, based on the assessment of student opinions. With our studies we focused on examining the most problematic areas identified by students. On the one hand, we have completed a series of questions in which we collected suggestions for quality improvement by interviewing students. The questionnaire was compiled by nearly a hundred students and reflects the opinions and insights of all three undergraduate, full-time and correspondence students in our institution. After that we examined the applicability of the method, which is popular in the fields of social science, made the necessary statements for quality improvement and asked students currently studying to prepare Q tables. We evaluated and analyzed the vessels and then compared them with each other.

Overall, we can say that the applied Q-method is well suited for exploratory and analytical examination of qualitative opinion types, expectations, evaluations. The method usually works with a small sample (10-50 test subjects) and representativeness is not required. In our study there were 15 students who are currently pursuing their studies in all three undergraduate courses. The Q-procedure is usually used in cases where there are not yet conscious standard opinions and points of view. The Q-method can be used in this context as an exploratory tool as a preliminary or complementary procedure for various quantitative and qualitative research. In contrast to the usual review methods that reveal problems and satisfactions, the Q-method is also suitable for identifying trends and expectations that can be detected based on reviews. For this reason, it can be used to complement traditional methods and be a step forward in defining trends for quality improvement. It can prevent large-sample, statistically determined quantitative research but it is by no means an alternative to survey research, where representativeness is an important requirement. The results of the Q-method cannot be generalized to the population in the same way as the results of qualitative studies in general, i.e., it is not possible to determine from the types explored how much of the sample is included, nor what percentage of the basic population may be characterized by that type of behavior. Further studies would be needed to establish this, which could be formulated as the next step in our research. Based on its results, it can be said that it is worth using this method either to analyze the temporality of expectations or to analyze a community of opinions in more detail. Finally, we can state that our research has also shown, that opinion assessment and the Q-method, are suitable and complementary procedures, for determining subjective value judgments and expectations.

References

[1] Standards and Guidelines for Quality Assurance in the European Higher Education Area (ESG) (2015) Brussels, Belgium, pp. 6-8, ISBN: 978-9-08-168672-3

[2] P. Shinebourne: Using Q Method in Qualitative Research. International Journal of Qualitative Methods Volume 8, Issue 1, March 2009, pp. 93-97, https://doi.org/10.1177/160940690900800109

[3] Bruce McKeown and Dan Thomas: Q Methodology. Quantitative Applications in the Social Sciences SAGE Publications, Inc. 1988, doi: https://dx.doi.org/10.4135/9781412985512

[4] Brown, Steven R.: Q methodology and qualitative research. Qualitative Health Research, Vol. 6, No. 4, 1996, pp. 561-567, tps://doi.org/10.1177/104973239600600408

[5] Kougias, I., Nikitas, A., Thiel, C., Szabo, S. (2020) Clean energy and transport pathways for islands: A stakeholder analysis using Q method. Transportation Research Part D: Transport and Environment 78, 102180 ISSN 1361-9209, https://doi.org/10.1016/j.trd.2019.11.009

[6] Watts, S. and Stenner, P.: Doing Q methodology. Qualitative Research in Psychology, 2, 2005, pp. 67-91, DOI: 10.1191/1478088705qp022oa

[7] R. Wolbertus, S. Jansen, M. Kroesena: Stakeholders' perspectives on future electric vehicle charging infrastructure developments. Futures 123 (2020) 102610, pp. 1-16, https://doi.org/10.1016/j.futures.2020.102610

[8] Hofmeister-Tóth Á., Simon J.: Theory and application of the Q method in marketing research/A Q-módszer elmélete és alkalmazása a marketingkutatásban. Vezetéstudomány - Budapest Management Review, 37 (9) pp. 16-26, DOI 10.14267/VEZTUD.2006.09.02

[9] Somoskovi Csilla: The attitude examination of workers of Psychiatry with the help of q-method. Theses of doctoral (PhD) dissertation. 2020, doi: 10.17166/KE2020.008

[10] Ernest, J. M.: Using Q methodology as a mixed methods approach to study beliefs about early childhood education. International Journal of Multiple Research Approaches, 5, 2011, pp. 223-237, http://dx.doi.org/10.5172/mra.2011.5.2.223

[11] Ramlo, S.: Facilitating a faculty learning community: Determining consensus using Q methodology. Mid-Western Educational Researcher, 24 (1), 2011, pp. 30-38

[12] Danielson, S., Tuler, Seth P., Santos, S. L., Webler, T., Chess, C.: Three Tools for Evaluating Participation: Focus Groups, Q Method, and Surveys. Environmental Practice. Vol. 14, No. 2, 2012, pp. 101-109

[13] G. J. Miller, M. L. Whicker: Handbook of research methods in public administration analytics. Marcel Dekker, Inc. New York, Basel, Hong Kong, 1999, pp. 599-637, ISBN: 0-8247-0213-1

[14] Anthony J. Onwuegbuzie, Using Q Methodology in the Literature Review Process: A Mixed Research Approach. Journal of Educational Issues, 2015, Vol. 1, No. 2, pp. 90-109, ISSN 2377-2263

[15] Stephenson, W.: Introduction to Q-methodology. Operant Subjectivity. Vol. 17, 1993, pp. 1-13

[16] Ásványi K., Hager Ch.: Guest preferences for sustainable hotel attributes/A vendégek preferenciái a fenntartható szállodai attribútumokról. Közép-Európai Közlemények 14(2) pp. 83-105, https://ojs.bibl.u-szeged.hu/index.php/vikekkek/article/view/34600

[17] I. T. Ellingsen, I. Størksen & P. Stephens: Q methodology in social work research, International Journal of Social Research Methodology, Vol. 13, 2010, Issue 5, pp. 395-409, https://doi.org/10.1080/13645570903368286

[18] Watts, S., Stenner P.: Doing Q methodology: theory, method and interpretation. Qualitative Research in Psychology. Volume 2, 2005, pp. 67-91, doi: https://doi.org/10.1191/1478088705qp022oa

[19] Valenta, A., Therriaul, D., Dieter, M., Mrtek, R.: Identifying student attitudes and learning styles in distance education jaln. Volume 5, Issue 2, 2001, pp. 111-127, DOI: 10.24059/olj.v5i2.1882

[20] Schmolck, P.: PQMethod-2.11d, 2002, http://www.rz.unibw-muenchen.de/~p41bsmk/qmethod

[21] Schmolck, P.: PQMethod Manual, The QMethod Page, 2014, http://schmolck.org/qmethod/

Fighting Insider Threats, with Zero-Trust in Microservice-based, Smart Grid OT Systems

Marina Stanojevic, Darko Capko, Imre Lendak, Sebastijan Stoja, Bojan Jelacic

Faculty of technical sciences, University of Novi Sad
Trg Dositeja Obradovica 6, 21000 Novi Sad, Serbia
E-mail: marina.stanojevic@uns.ac.rs; dcapko@uns.ac.rs; lendak@uns.ac.rs; sebastijan.stoja@uns.ac.rs; bojan.jelacic@uns.ac.rs

Abstract: The Operational technology (OT) systems, utilized in critical infrastructure systems, can largely benefit from microservice-based control center architectures, by lowering upfront investment and maintenance costs. Many system operators are cautious and do not choose such modern system architectures, citing cybersecurity as a major concern. We intend to tackle that challenge and, in this paper, we investigate the threats to such mission critical systems and propose mitigation strategies aimed at lowering the likelihood of cyber-attacks. We developed a threat model focused on both external and insider threats and we group them. We utilize Microsoft's STRIDE methodology to analyze the threats on a per-service level, in a specific use case, in the smart grid sector. We propose mitigations for each threat, by putting the zero-trust principle, at the core of our proposal. We calculate the resulting risks, for each threat, based on impact and likelihood, and show that it is significantly reduced when all proposed measures are applied.

Keywords: Microservices; Security; Threat modeling; Smart Grid; Cloud computing; Zero Trust principle; Insider threats

1 Introduction

Every critical system is target to cyber-criminals and the energy industry is no exception. According to statistics, the energy industry became a prime target in the last decade [1]. We can identify several types of attackers, based on their motive for an attack. The most common motives behind corporations or even states are sabotage, where the goal is to present specific utility as unreliable which can result in losing customers and money. Another type of attacker is an individual motivated by a desire for revenge because of some personal reason or desire to prove themself as capable, to hack systems. This individual can be an insider that already has access and knowledge of the system's vulnerabilities making it easier to perform an attack.

Smart grids are considered as a critical infrastructure, whose continuous operating is mandatory. With introducing new functionalities, new vulnerabilities are introduced, as well [2]. Smart grid is consisted of two systems, information technology (IT) which is responsible for business and operational technology (OT) taking care of real-time operations. As the demand for electricity is growing, it is also becoming clear that not all current ways of energy supply are sustainable. Humanity needs to start leveraging renewable energy sources, such as solar panels or windmills. Smart grid OT system is mainly developed in monolithic and service-oriented (SOA) architecture which makes it robust and non-scalable. Traditional, monolithic solution for smart grid OT was deployed on utility-owned computing resources and security over a decade relied on physical isolation and security by obscurity.

Microservices architecture is known for over a decade [3] but is not yet leveraged in all industries. It is based on building systems as sets of multiple independent services, each running as separate processes, which communicate with each other through lightweight mechanisms. The main advantage, compared to traditional, monolithic architectures is that the services are deployed separately. Another main advantage is out of the box solution for vertical scalability. As one node in a microservice architecture is considered to be responsible for one function, if the number of requests for that function is increased, more instances of that specific service could be started to respond to the increased load. Another advantage of breaking up the monolith into multiple services is improved fault tolerance. If a single function of a monolith fails, the complete system might become unusable. For microservices that is not the case. As only the affected functions are unavailable and if there are no dependencies between them, the system can continue to provide functions that are still up and running while failure is investigated and fixed.

Deploying system such as smart grid OT in cloud environment is tempting because of all the benefits cloud brings, but there is also a big concern regarding security that should not be neglected. Although it is important to protect system on perimeters, which is mostly present in traditional monolithic smart grid OT solution as well, breaking the monolith to multiple microservices introduces new threats. Implementing a concept of zero-trust principle between microservices can reduce the risk of attacks to be perform as it is required for all users or nodes in the system to be authenticated and continuously validated while using the system features. It is also a good practice when it comes to protecting the system from insiders.

In this paper, an accent will be on building a threat model for smart grid OT systems whose architecture is built on a microservices platform [4]. After the threat model is built, we propose a novel architecture with appropriate security controls following zero-trust principle. The proposed architecture was tested and verified with Microsoft's Spoofing, Tampering, Repudiation, Information

Disclosure, Denial of Service, Elevation of Privilege (STRIDE) methodology. Risks for threat exploitation were identified and measured.

The paper consists of six sections. In the following section, related works are analyzed. Section 3 decomposes reference architecture into components and explains them together with a diagram representing application's data flow. Section 4 contains the threat model and proposes mitigations for each threat. Each component is analyzed using the STRIDE methodology and risks for threat exploitation are calculated and that is presented in Section 5. Section 6 contains our conclusions.

2 Related Work

In this section, we give an overview of the related literature on the researched fields: security in microservice-based architecture, insider threats and zero-trust model.

2.1 Security in Microservice-based Architecture

An overview on the current state of security in microservice-based systems is given in paper [5]. The authors analyzed 70 articles and gray literature on the topic and presented summed security ideas, principles, analyses, mechanisms and designs used to protect microservice-based systems. By decomposing components, the need for network connectivity is introduced and with it attack surface expanded. As microservices are designed to trust their peers, compromising one, all the others become exploited.

In [6] authors answer the question "What are the risks and how they can be addressed in an early phase or minimized after an attack?", giving the list of recommendations. As presented in literature survey [7], most studies focus on the stopping or mitigating attacks and not much on recovering from them. Intrusion detection systems can be used in container environments, as well [8]. Communication is the biggest concern when it comes to securing microservice-based architecture and authorization and authentication turned out to be the most used security mechanisms [9]. Authors present literature review and they found that mechanisms such as OAuth 2.0, OpenID Connect are used to overcome this problem. The authors of reference [10] developed a framework for establishing trust and making communication secure. Testing showed that due to communication overhead, system performance is slightly impacted.

2.2 Insider Threats

Even though 85% of damage comes from insiders [11] who can be either employee or trusted third party with appropriate privileges (e.g., cloud vendor), the most studies cover protection from attacks coming from external threats. The first step to improve defense from insiders is to establish strong security policies and constantly monitor employee activities [12]. One way to approach a better understanding of insider threat is to identify the problem space, technical and behavioral events and indicators and to analyze potential attackers and their motivation [13]. Reference [14] reports similar research with a focus on Internet-of-Things (IoT) environments. Authors of reference [15] analyzed 120 real case studies and defined attack patterns that could help in detecting insider-threats.

Insider threats are present in microservice-based systems as well especially because in deploying microservices, involvement of special governance tools is needed. In reference [16] authors identify integrity threats and define set of security requirements for microservice-based systems. They propose a framework for insider-resistant integrity protection.

2.3 Zero-trust Model

One example of a security-as-a-service solution is presented in reference [17] in which the authors introduced a flexible monitoring and policy enforcement infrastructure for network traffic. Cloud applications can leverage this solution to detect and block external and internal threats. Although shifting responsibility to others is tempting, zero-trust principle is gaining popularity among the majority of companies [18] which is quite the opposite security concept. As the word says itself, this principle is based on treating all network traffic as hostile, where it is not important if it came from inside or outside of perimeter.

Implementation of Kubernetes and Istio service mash gives out of the box solution for zero-trust model in containerized environment. Reference [19] proposes an additional set of tools whose usage can protect data that are transferred between microservices. An interesting study was done on the impact of the zero-trust model implementation on system performance [20]. Results showed that Istio reduced latency variability in responding to sequential HTTP requests and that the CPU and memory usage can be increased. Another paper on protection of security data is [21] where authors implemented new policies in the zero-trust model. An access control proxy is introduced whose task is to analyze access request, user type, device type, application type and data type. Overall strategy for establishing zero-trust model in cloud computing environment is given in [22]. As cloud environment cannot be trusted due to its dynamic and shareable landscape main challenge is to protect resources from data breaches. The authors propose implementation of trust engine, that will dynamically calculate trust which is used later in transaction requests. Although the benefits of implementing zero-trust have been well researched, exploring the disadvantages and costs of zero-trust is neglected [23].

The division of smart grid OT system into multiple microservices gives more flexible, scalable architecture and better performance results [4]. As the proposed architecture has not been analyzed from the security point of view, that will be done in this paper. As the zero-trust model seems to be very good for mitigating insider threats, we will prove that by implementing zero-trust, risks for exploiting threats, even from insiders, is significantly reduced.

3 Reference Architecture

Smart grid OT system is considered critical, and it is expected that it provides real-time service 24 hours a day. Some of its functions is to give overview of the distribution network's health, connectivity, status of equipment on the field, etc. Potential network outage could endanger people's lives so fast detection and quick utility personnel response are crucial.

In this research, smart grid OT system is built on microservice-based architecture, deployed in cloud environment. Using this architecture gives better system performance [4] and response because if the demand is higher, the number of microservice instances can be increased. Deploying the system in the cloud environment lowers the needed upfront costs and overall cost-consumption. The main disadvantage in this approach is that new security vulnerabilities are introduced and that the system is now exposed to the public internet and to the cloud vendor.

Reference architecture used in this paper is consisted of seven services, each deployed separately in the same microservice cluster and two databases. Client application and SCADA are not considered as part of the microservice cluster as they are applications that utilities use as an access point to the services. Client application is usually deployed in control room of utility that is monitored and controlled, while SCADA is used to monitor and control devices in the field. In the following subsections, each component and system's data flow are described.

3.1 System Components

Architecture of the system is shown in Figure 1.

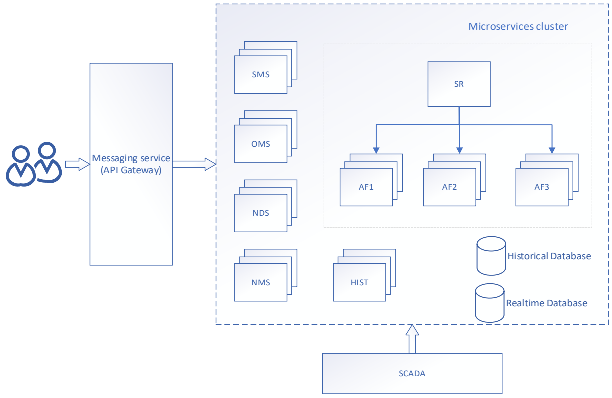


Figure 1
Reference architecture

System is consisted of the components which are divided in four groups:

1. Perimeter components – The entry point of the microservice cluster is Messaging service (MS) and its responsibility is to forward client requests to the appropriate service. Every change coming from process environment (SCADA) is processed in the Network dynamics service (NDS) and stored in RD. Communication is two-way, status of remote points can be changed through this service.

2. Core services – Network management service (NMS) is responsible for maintaining the distribution network's model and for orchestrating model updates. Outage management service (OMS) keeps track of outages in the power grid, defines recovery plans, performs automatic actions to power restoration. Service responsible to keep records of repairs and to create work orders to send crews to the field is Switching management service (SMS). Historian service (HIST) is component used as proxy to Historical database. It allows adding new entries and reading existing.

3. Services for analytic functions – Service for function requests (SR) upon receiving request from MS, divides it to multiple requests based on number of roots and forward them to appropriate analytic function. Analytic function (AF) engages algorithm for specific energetic calculation. Since calculations for multiple roots can be done in parallel, more instances of this service are used.

4. Data storage – Realtime database (RD) contains information about the current state of the distribution network. Every change in the system is preserved in Historical database (HD) database.

3.2 Inputs, Outputs and Data Flow

Data flow diagram in Figure 2 represents how data is sent through an application and its components. It is important input for any threat analysis.

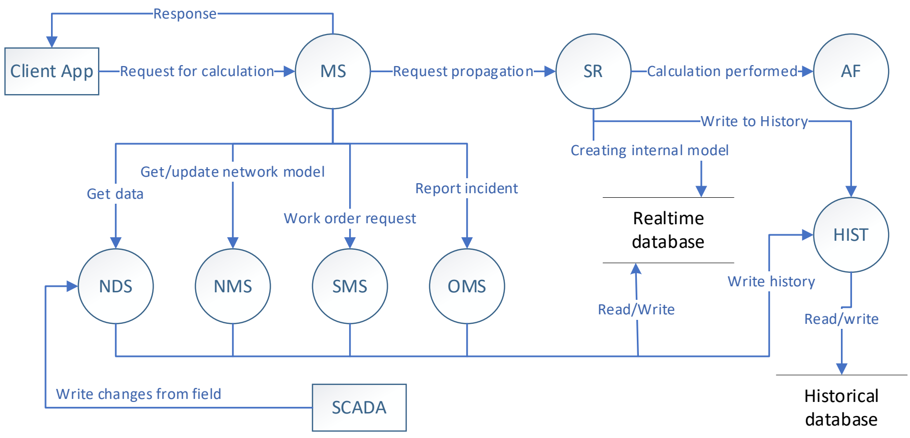


Figure 2
Data flow diagram

Each service has a list of methods it exposes to other services and to applications outside the microservice cluster (e.g., SCADA and client application). Services have read and write access to Realtime database and to HIST. SR is receiving requests from MS, obtaining information from Realtime database that are necessary for calculations and propagating request to appropriate AF. This service is also communicating with HIST and writing history to Historical database through it. Regarding client application, it is communicating with system through MS whose role is to be a mediator and to forward request to appropriate service. SCADA is communicating with NDS service directly and is sending updated status of field equipment.

4 Threat Modeling and Mitigation

First step in threat modeling is to decompose application to its components, which is easier in microservice architecture because application is already decomposed to multiple services. Every connection point between components should be observed from the attacker's point of view and analyzed how it could be exploited to gain leverage. The threat model diagram of application's data flow is presented in Figure 3. It focuses on detecting treats coming from insiders. For threat analysis, Microsoft Threat Modeling tool [24] was used. As it can be seen on data flow diagram in Figure 2, the services NMS, NDS, MDM, WOM and OMS have the same flows, so they are modeled as one Service in Figure 3, to simplify diagram.

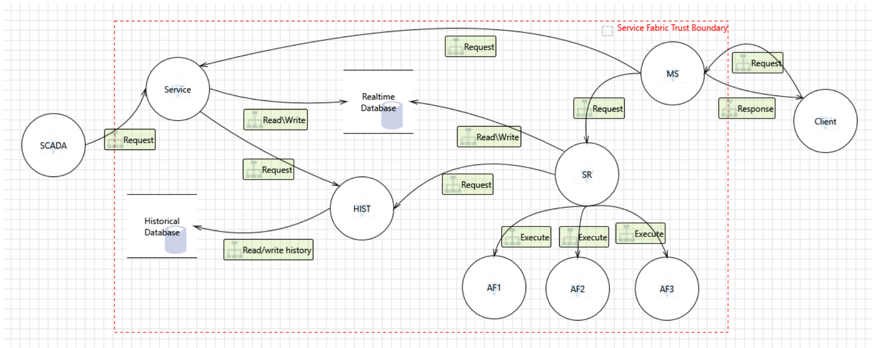


Figure 3
Threat modeling diagram

After making the diagram, presented in Figure 3, a report was generated which gave overview on threats between each component and proposed mitigation. After analyzing report, we came to conclusion that threats for all services are the same. Proposed mitigation list is given in the following subsections and it is necessary to apply them to all services in the system. Figure 4 contains graphical representation of proposed mitigation list.

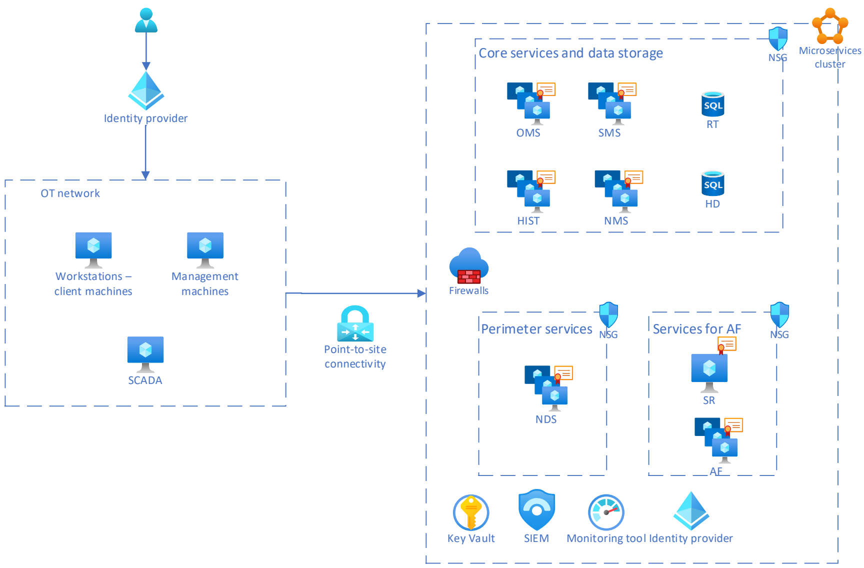


Figure 4
Proposed mitigation

4.1 Establishing Strict Business Rules

Appropriate periodical security education and repetitive psychological (background) checks are essential for all employees. To grant admin rights to the user, an additional paperwork should be requested, such as that he has additional responsibility and will be liable in court if his account is used for malicious activities. Workstations must be in physically isolated room that is constantly monitored and access to this room should be given only to employees with privileges. External devices, such as cameras, USB or hard drives are not allowed. This way, possibility of stealing information is significantly decreased. System deployment and update is considered critical and will be supervised and approved by personnel with highest privileges. Adding new nodes to Service Fabric cluster outside the deployment procedure is prohibited.

4.2 Authentication

The first step is to restrict anonymous access to the Service Fabric cluster. That can be done with implementing authentication process, where users must prove that they are who they claim to be. To secure client-to-node access, identity provider is used to which users must authenticate before gaining access privilege to application. Besides requesting strong passwords from clients through policy, multi-Factor authentication (MFA) must be applied. This way the adversary can’t login even if he steals user's password.

To secure system from insider threats, implementing zero-trust principle is required so nodes authentication is also needed. Recommendation is to use certificates (e.g., Cluster X.509). This will prevent nodes that are not certified to join Service Fabric cluster and to gain unauthorized access to other nodes. It must be ensured that Service Fabric client-to-node certificate is different from node-to-node certificate. Service Fabric certificates should be obtained from an approved Certificate Authority (CA) and self-signed or test certificates cannot be allowed in production. The same principle applies when it comes to scaling up, each service instance must have a valid certificate.

4.3 Authorization

There are different system use cases, client access to cluster, access to cluster for deployment purpose, admin access and access from SCADA to update status of equipment. To implement zero-trust principle, it is not enough to analyze and protect only system perimeters so authorization methods will be applied to each microservice individually as they all interact with each other. Role-based access control (RBAC) must be implemented by following defense in depth principles. For each service, there is a separate group of privileges assigned and defined for each user what privileges he has. Recommendation is that only administrator account has full access to management capabilities. After establishing RBAC for client's access, authorization controls will be enabled, to prevent user from achieving privilege escalation. This will enable verification of caller's permissions and establishing whether he has or does not have enough privilege to execute method on the server.

Access Control Lists (ACL) must be implemented on files (e.g., XML files) and prevent their unauthorized manipulation. This is especially important in between nodes communication when service is trying to manipulate with data of other service and change its state. When updating critical configuration, additional check for admins must be applied, for example to repeat MFA.

4.4 Protecting Secrets

Binaries which contain sensitive information (e.g., code for topology calculation) must be obfuscated. Sensitive data stored by services (e.g., internal model created by service for requests) will be encrypted and disk-level encryption will be used on nodes in the cluster. Message security protection level will be set to encrypt and sign. This way even if an adversary intercepts a message, it will be encrypted and unusable. A key vault will be used as a solution which has its own authentication and authorization for certificate and key storage. Important condition is that the data should not be seen or extracted by anyone without rights (e.g., vendor) and that all secrets are guarded using industry algorithms and appropriate key lengths.

4.5 Network Security

To additionally harden network security, firewall that will allow network traffic only from specific IP addresses and ports will be deployed. Regarding communication between services, throttling will be enabled. Network segregation is important to separate services that have communication with external systems (NDS and SR) from the rest of the system. Each network group have defined inbound and outbound network rules, which is another layer of communication restriction between services. Preventing over-consumption of resources by limiting concurrent calls, instances or sessions is also a good practice.

Any machine from the cluster shouldn't be exposed publicly or have public endpoints. Here we propose that for accessing machines that are in the Service Fabric cluster, the company set up point-to-site VPN and dedicate one host machine in corporate network to be used for remote connections for administration purposes. Access to that host machine, must be restricted only to personnel, with high authority, whose job is to run the business. Introducing point-to-site VPN, additional level of authentication is required. To access machine in cluster, user must authenticate to corporate network and to identity provider whether through installing certificate on machine or providing credentials. Point-to-site VPN will be used also for client access, each client will have dedicated workstation on which client application is installed. Same applies to communication between SCADA which is on machine from OT network.

4.6 Database Security

Database access will be configured with roles and least-privileged accounts will be used to connect to the database. Regarding service login, instead of direct access to the tables, which must be forbidden, there is a list of selected stored procedures which are allowed to be executed by each service. Members of the database admin server role will be very limited and never contain accounts used by services. Firewall is configured for Database Engine Access. Login auditing will be enabled on DB Server. Digital signature will be added to critical database securables. Strong encryption algorithms must be used to encrypt data in the database. In microservice architecture, good practice is that every service has its own database so that risk of interfering with each other's data is not possible. If that is not possible, Row Level Security (RLS) must be applied. RLS enables implementation of restrictions on data row access. For example, ensuring that services can access only those data rows that are pertinent to their scope.

4.7 Logging and Monitoring

Proper logging of all security events and user actions builds traceability in a system and denies any possible repudiation issues. Logging successful and failed authentication attempts must be enabled. Besides that, application has its own logs and audits every request. These log files are considered sensitive information and are protected from unauthorized access by restricting view/write privileges only to administrators who will do any necessary inspections. Another good practice is to disable deletion of these files, they are archived periodically instead.

Monitoring metrics can be useful in detecting system's anomalous behavior. If we take denial of service attack as an example, monitoring system could detect it at the beginning of an attack, because number of requests is increased, and the system would start increasing the number of service instances. A sudden increase in the number of instances could trigger a rule which would create incident. An external monitoring tool is recommended, because it would reduce the need to connect and monitor directly. When it comes to what should be monitored, besides number of requests and number of service instances, successful/unsuccessful logins to the identity provider, key vault accesses, request solve duration, CPU/memory usage for each machine, traffic between system components, accesses to database could be helpful. A Security Information and

Event Management (SIEM) solution is recommended tool which in addition brings threat detection by analyzing collected security events. According to the zero-trust principle, access to these tools is restricted only to admins.

4.8 Protecting from Outside

Special attention is needed for field devices that communicate with the system through the SCADA. Although the SCADA is outside of the microservice cluster and its security is out of scope for this paper, some measures need to be taken. The NDS service is tracking frequency of value changes and if some field device is sending more changes than usual, incident is reported with high priority. SIEM can be used as a detection tool in this scenario. On the other hand, there is whitelist for clients who have access to the system. MS detects and forbids clients that act suspicious and are sending a lot of requests in short time.

4.9 Cost of Novel Architecture

It must be noted that deploying the proposed system architecture in a smart grid OT environment certainly introduces costs and impacts performance. Firewall, VPN tunnel, key vault and SIEM are tools that can be quite resource-intensive, depending on cloud vendor. Controls like authorization and encryption lead to higher CPU utilization which can be resolved with upgraded hardware or running additional micro-service instances.

5 Component Level STRIDE Analysis

Keeping in mind the nature of microservice architecture (e.g., components can change context or accessibility rapidly) and the fact that the attacker could be an insider, besides investigating threats for connection points on microservices trust boundary, STRIDE analysis for each component is also needed. That way, it can be determined how attacker can target each component. In this section, each component service is analyzed individually using the STRIDE methodology [25]. Risk for each threat is calculated following the Federal Information Processing (FISP) [26] and the European Network and Information Security Agency (ENISA) [27] standard, as shown on Table 1 and based on impact and likelihood. Impact, presented in Table 2, determines how negatively the exploitation of the threat would affect the business and clients. On the other hand, likelihood (Table 3) stands for the probability of the threat's exploitation by attackers.

Table 1
Risk matrix

	Impact					
Likelihood		Very high	High	Moderate	Low	Very low
	Very high	Very high	High	Moderate	Low	Very low
	High	Very high	High	Moderate	Low	Very low
	Moderate	High	Moderate	Moderate	Low	Very low
	Low	Moderate	Low	Low	Low	Very low
	Very low	Low	Low	Very low	Very low	Very low

Table 2
Impact level description

Impact	Description
Very high	Expected severe or catastrophic adverse effects on operation, assets, individuals: • Serious personnel injuries or loss of lives. • Long-lasting system unavailability and unusability resulting in blackouts. • Valuable asset destruction.
High	Any action which could lead to losing customers due to system unreliability and major financial loss: • The system loses capability to perform one or more of its primary functions (or connectivity with SCADA) and gives wrong calculations. • Leakage of secret customer information. • Expensive equipment damage.
Moderate	Serious adverse effect is expected: • Causing a significant degradation in the effectiveness of the system functions. • Losing customer trust. • Revealing information about system that can be used by competition (or attacker) to gain an advantage. • Unavailability of non-critical system components.
Low	Limited adverse effect is expected: • Minor damage to assets. • Financial loss.
Very low	Negligible adverse effects are expected which do not affect system performance.

Table 3
Likelihood level description

Likelihood	Description
Very high	An adversary is almost certain to initiate the threat event. This means that the system has serious security flaws that can be exploited, for example, if system is publicly available and anyone can use its functions or change data.
High	An adversary is highly likely to initiate the threat event. The system architecture has weak spots that an experienced adversary can leverage. Employee with higher privileges than necessary could be weak spot, either malicious (insider threat) or uneducated and tricked by an adversary.
Moderate	An adversary is somewhat likely to initiate the treat event. With great effort the attacker can obtain limited access to the system but still cannot endanger the system, so his motivation is low.
Low	An adversary is unlikely to initiate the threat event. Employees are loyal and well educated so the insider threats are not likely to happen. The system is protected on its perimeters, so if it comes to breach, attacker could not reach any sensitive information.
Very low	An adversary is highly unlikely to initiate the threat event. The system is protected from every aspect and could not be penetrated.

In the following subsections, analysis for each STRIDE threat (Spoofing, Tampering, Repudiation, Information disclosure, Denial of service, Elevation of privilege) is given.

5.1 Spoofing

In spoofing communication from an unknown source is masked so it seems like it is coming from a known source. Service Fabric cluster does not have public endpoints and is not accessible from the public internet, it is reachable only through point-to-site VPN, authorized workstations that are in the utilities control center. The weakest link are employees who an adversary can potentially take advantage of using phishing, getting them to install malicious software or reveal secrets (e.g., their password). Another potential breach is to spoof node in microservice cluster by using stolen certificates. Outcome is same, an adversary could degrade the integrity of the system or reveal insight in system's operations.

If an adversary gains control over the NDS for example, he can send commands to field and leave catastrophic consequences like outages and blackouts. In the Table 2, this is marked as High impact.

This threat could not be exploited without getting insider's help because the workstations from which the system could be accessed are physically isolated and protected. Even if an attacker steals a password, he would need to pass MFA. Another scenario for attacker is to deploy additional node in cluster that could act like real service and send malicious data to other services. Because of strong system deployment rules, the presence of more people with different privileges is required to add new node, meaning that one malicious insider could not perform this operation alone even with stolen certificates. Likelihood for this threat is Low, as it is unlikely that the adversary can initiate threat event (Table 3).

According to Risk matrix presented in Table 1, risk for this threat is ***Low***.

5.2 Tampering

This threat stands for changing of data (destruction, manipulation or alteration) through unauthorized channels. Communication between all system components is encrypted and protected so intercepting or tampering messages between services is worthless because without encryption key, those messages are meaningless to the adversary. Regarding databases, they are protected with RBAC, other files (e.g., log files) with ACL.

Impact is High because with this threat exploited attacker could gain insight in confidential client's information (Table 2). Compromising data from DMS or Historical database (e.g., attacker manipulates equipment's state) would result in wrong function calculation and would give false power grid status.

Likelihood for this threat exploit is Low according to Table 3 after applying proposed mitigation even if attacker is insider because communication and disks are encrypted and data is protected from unauthorized access. An adversary could not reach any sensitive information because each service implements appropriate authentication and authorization methods, encrypts messages and data.

Risk for this threat is ***Low***.

5.3 Repudiation

If the system lacks traceability, it is difficult to identify attacker who would perform malicious operations in the system. In the absence of proper auditing and logging controls, it would become impossible to implement any accountability in a system. There must be logging of all security events, user actions (e.g., successful and failed authentication), each service logs relevant activity. Constant monitoring of system is recommended through internal or external tools, like SIEM, which has automatic rules for incident reporting if suspicious system behavior or access occurred.

Impact for this threat is Very low because repudiation itself could not harm the system (Table 2). If this threat is exploited, tracing attacker is impossible and it is unknown if or when attack will be repeated.

Likelihood is Very low (Table 3). With implementation of logging and monitoring, set of events are defined whose processing can lead to instant responsible personnel informing if something is violated or the system is breached. Also, log files are protected from deletion and unauthorized access, so even insider could not make them inaccessible if analysis after the breach is needed.

Risk for this threat is ***Very low***.

5.4 Information Disclosure

Secrets can be any sensitive information, such as connection strings, passwords, state of equipment in the field, etc. Algorithm for DMS functions is also considered sensitive business logic.

Impact is High as attackers could perform sabotage to make clients lose trust in utility as their electricity distributor and to switch to the competitor (Table 2). Sabotage could involve equipment destruction, topology change, disclosure of client confidential information, blackmail, etc. DMS function algorithm is important from the system performance aspect. With better algorithm, topology and load flow analysis gives more accurate results faster. For example, load flow results can be used to predict energy consumption in order to optimize the consumption demand ratio so having a better DMS function algorithm gives a competitive advantage. It is considered confidential business logic because it is developed or improved by the company's power engineers.

Likelihood for this threat is Low (Table 3). As with appropriate authentication and authorization implemented, all secrets are protected with RBAC and encryption, including databases. Keys and certificates are stored in key vault which brings additional layer of protection. In order for client application to gain access to the system it must be on whitelist which is monitored and controlled on daily basis.

Risk for this threat is ***Low***.

5.5 Denial of Service

If adversary launches Distributed Denial of Service (DDoS) attack he could provoke shutting down a service or network so that it is inaccessible to other clients. Besides service unavailability, increased load on the system could result in vertical machine scaling which can significantly increase cost.

Impact is High because if system is unavailable, utility is cut off from its consumers and unable to react to any potential outage (Table 2). Major financial loss is also possible because of larger number of service instances engagement than needed.

Likelihood for this threat is Low (Table 3) because with SIEM any anomalous behavior coming from user or device is detected and requests coming from that client are no longer processed because that client is taken of the whitelist. On the other hand, the real threat is not malicious user, but endurance of the system. Employees can overload the system so that it becomes inaccessible. Similar situation can happen in case of storm during which a significant number of changes in the field occur in a short time. The system responds in these scenarios by increasing the number of instances so it will not become unusable.

Risk for this threat is ***Low***.

5.6 Elevation of Privilege

The attacker receives an account with fewer access rights and manages to obtain privileges with higher rights. Implementation of strong authentication methods relying on MFA, RBAC and ACL lowers the possibility for this threat to be exploited.

Reason for High impact (Table 2) is that if the attacker gets greater rights, he can create incidents in distribution network that will cause blackouts, gain insight to secret customer information, sabotage field crews, etc.

Likelihood is Very low (Table 3), because each service is protected with RBAC and user privileges are checked at each step. Users are required to authenticate with MFA before using any system component. List of users and user groups are visible only to administrators and any change that involves changing access to any resource requires administrator to re-authenticate.

Risk for this threat is ***Low***.

Conclusions

Securing mission critical systems, has always been a challenge and in most cases, a showstopper for shifting them to the cloud environment. The utilization of microservice-based architectures introduces multiple benefits regarding operational technology (OT) system performance. In addition, when deploying such system in the cloud environment, upfront investment and maintenance costs are lowered. The goal of this research was to analyze a novel microservice-based architecture for OT systems from the security point of view and to propose mitigation strategy which will lower the likelihood of threat exploitation. During our research of zero-trust principle, we came to the conclusion that security should be embedded in system architecture in the system design phase. Keeping that in mind, we developed a threat model, investigated and grouped external and insider threats, and proposed a novel architecture with appropriate mitigations for each threat.

First, the services were divided into three groups: core services and data storage, services for analytic functions and perimeter services. The communication networks groups serving these groups of services need to be segregated, where inbound and outbound rules are defined and message security protection level set to encrypt and sign. All communication with outside systems should go through encrypted, point-to-site VPN tunnel and firewall. All data should be encrypted as well, including databases and data disks. For sensitive data storage, such as encryption keys or certificates, key vault is a suitable solution. A cloud-based identity provider needs to be deployed and used to authenticate users before granting system access. Anonymous access to the system is forbidden, strong password and multi-factor authentication (MFA) is necessary for all users. For node authentication, certificates are used. Role-based access control (RBAC) is implemented following defense in depth principles and authorization controls enabled. Activity logging and monitoring need to be planned carefully and appropriate alerts need to be defined to maximize the likelihood of early (cyber) attack detection.

Microsoft's STRIDE methodology was used to analysis the above-described smart grid OT architecture. The service-level risk analysis showed that the likelihood for threats exploitation is significantly reduced.

Future work will encompass comparing the proposed solution herein, with and without, the listed security controls and the calculation the exact overhead introduced, within different cloud environments.

References

[1] Macola, Ilaria Grasso: The five worst cyberattacks against the power industry since 2014, Power-Technology.com, April 2020

[2] D. Rosic, I. Lendak, and S. Vukmirovic, A Role-Based Access Control Supporting Regional Division in Smart Grid System, Acta Polytechnica Hungarica, Vol. 12, No. 7, pp. 237-250, 2015

[3] Dragoni, Nicola, Saverio Giallorenzo, Alberto Lluch Lafuente, Manuel Mazzara, Fabrizio Montesi, Ruslan Mustafin, and Larisa Safina., Microservices: yesterday, today, and tomorrow, Present and ulterior software engineering, pp. 195-216, 2017

[4] S. Stoja, S. Vukmirovic, N. Dalcekovic, D. Capko, B. Jelacic, Accelerating Performance in Critical Topology Analysis of Distribution Management System Process by Switching from Monolithic to Microservices, Rev. Roum 63, No. 3 (2018), pp: 338-343, 2018

[5] A. Pereira-Vale, E. B. Fernandez, R. Monge, H. Astudillo, G. Márquez, Security in microservice-based systems: A Multivocal literature review, Computers & Security, Vol. 103, April 2021

[6] N. Mateus-Coelho, M. Cruz-Cunha, L. Gonzaga Ferreira, Security in Microservices Architectures, Procedia Computer Science, Vol. 181, pp. 1225-1236, 2021

[7] Pereira-Vale A, Márquez G, Astudillo H, Fernandez EB, Security Mechanisms Used in Microservices-based Systems: A Systematic

Mapping, XLV Latin American Computing Conference (CLEI) IEEE, pp. 01-10, 2019

[8] Flora, José, Improving the security of microservice systems by detecting and tolerating intrusions, 2020 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW), pp. 131-134, IEEE, 2020

[9] de Almeida, Murilo Góes, and Edna Dias Canedo, Authentication and Authorization in Microservices Architecture: A Systematic Literature Review, Applied Sciences 12, No. 6 (2022): 3023, 2022

[10] T. Yarygina and A. H. Bagge, Overcoming Security Challenges in Microservice Architectures, 2018 IEEE Symposium on Service-Oriented System Engineering (SOSE), pp. 11-20, 2018

[11] Z. M. Yusop and J. H. Abawajy, Analysis of insiders attack mitigation strategies, Procedia-Social and Behavioral Sciences, Vol. 129, pp. 581-591, 2014

[12] Saxena, Neetesh, Emma Hayes, Elisa Bertino, Patrick Ojo, Kim-Kwang Raymond Choo, and Pete Burnap, Impact and key challenges of insider threats on organizations and critical businesses, Electronics 9, No. 9 (2020): 1460, 2020

[13] Nurse, Jason RC, Oliver Buckley, Philip A. Legg, Michael Goldsmith, Sadie Creese, Gordon RT Wright, and Monica Whitty, Understanding insider threat: A framework for characterising attacks, 2014 IEEE security and privacy workshops, pp. 214-228, IEEE, 2014

[14] Kim, Aram, Junhyoung Oh, Jinho Ryu, and Kyungho Lee, A review of insider threat detection approaches with IoT perspective, IEEE Access 8, pp. 78847-78867, 2020

[15] Agrafiotis, Ioannis, Jason RC Nurse, Oliver Buckley, Phil Legg, Sadie Creese, and Michael Goldsmith, Identifying attack patterns for insider threat detection, Computer Fraud & Security 2015, No. 7, pp. 9-17, 2015

[16] Ahmadvand, Mohsen, Alexander Pretschner, Keith Ball, and Daniel Eyring, Integrity protection against insiders in microservice-based infrastructures: From threats to a security framework, Federation of International Conferences on Software Technologies: Applications and Foundations, pp. 573-588, Springer, Cham, 2018

[17] Y. Sun, S. Nanda, and T. Jaeger, Security-as-a-service for microservices-based cloud applications, 2015 IEEE 7th International Conference on Cloud Computing Technology and Science (CloudCom), pp. 50-57, IEEE, 2015

[18] Buck, Christoph, Christian Olenberger, André Schweizer, Fabiane Völter, and Torsten Eymann, Never trust, always verify: A multivocal literature review on current knowledge and research gaps of zero-trust, Computers & Security 110 (2021): 102436, 2021

[19] de Weever, Catherine, and Marios Andreou, Zero trust network security model in containerized environments, University of Amsterdam: Amsterdam, The Netherlands, 2020

[20] S. Rodigari, D. O'Shea, P. McCarthy, M. McCarry, S. McSweeney, Performance Analysis of Zero-Trust multi-cloud, 2021 IEEE 14th International Conference on Cloud Computing (CLOUD), pp. 730-732, IEEE, 2021

[21] I. Ahmed, T. Nahar, S. Sultana Urmi, K. Abu Taher, Protection of sensitive data in zero trust model, Proceedings of the International Conference on Computing Advancements, pp. 1-5, 2020

[22] Mehraj, Saima, and M. Tariq Banday, Establishing a zero trust strategy in cloud computing environment, 2020 International Conference on Computer Communication and Informatics (ICCCI), pp. 1-6, IEEE, 2020

[23] C. Buck, C. Olenberger, A. Schweizer, F. Völter, and T. Eymann, Never trust, always verify: A multivocal literature review on current knowledge and research gaps of zero-trust, Computers & Security 110, 102436, 2021

[24] Microsoft MSDN documentation, the Threat Modeling tool https://docs.microsoft.com/en-us/azure/security/develop/threat-modeling-tool-mitigations

[25] Microsoft MSDN documentation, the STRIDE Threat Model. https://msdn.microsoft.com/en-us/library/ee823878(v=cs.20).aspx

[26] New Standards for Security Categorization of Federal Information and Information Systems, NIST Federal Information Processing Standard (FIPS) 199, Feb. 2015

[27] European Network and Information Security Agency (ENISA), Annex II, Security aspects of the Smart Grid, https://www.enisa.europa.eu/

Pneumatic Piston Control Modelling and Optimization

Tamás Szakács

Óbuda, University, Bánki Donát Faculty of Mechanical and Safety Engineering - BGK, MEI, 1034 Budapest, Bécsi út 96/b, szakacs.tamas@bgk.uni-obuda.hu

Abstract: Piston control can be used for position, speed, acceleration, and force control. The main difference between pneumatic, and hydraulic cylinders is that the hydraulic control uses fluid instead of gas, and fluids are practically incompressible. Therefore, hydraulic cylinders can be more accurately controlled for position and speed. Pneumatic cylinders are showing greater dependence on external loads. Pneumatic systems in the contrary are simpler as construction, cheaper than hydraulic ones, and more wide-spread in laboratories and shops, but usually they are used for end-to end motions without exact position control. This paper introduces a pneumatic piston control that decreases the drawbacks of the position control of pneumatic cylinders, making it allow for simple application.

Keywords: pneumatic piston; position; control; vehicle height control

1 Introduction

Modelling of a hydraulic cylinder is not an easy task it involves thermodynamic, fluidmechanic, and mechanical theories. Modelling of a pneumatic cylinder is even more difficult, because working medium in pneumatics is compressible, and because of that the piston position cannot be exactly controlled, piston position and behaviour is greatly depending on the changes on external loads. Based on this, pneumatic cylinders are not usually used for position control.

Recently a special interest can be seen on pneumatics due to development of many affected areas, like robotics, and pneumobiles [1], and conventional vehicles. Robotics requires accurate control, and sensorics, same as pneumobiles, where cylinder force, and motion are controlled. Haiato, et al. [2] is describing a discrete-valued model-predictive control (DVMPC) algorithm with promising results. Power & Motion Tech [3] describes an ElectroPneumatic Positioning System (EPPS) using Rexroth (Aventics) components. The article summarises advantages of EPPS compared to ElectroMechanical positioning (EMPS). It draws a conclusion of a frequently overlooked phenomenon: “With directional flow control, minor leaks cause the pneumatic cylinder piston to move slightly.

The positioning system's controller receives motion feedback from the sensor, and the controller responds by slightly adjusting flow to compensate for piston movement. This slight, but constant back-and-forth motion, known as dithering, negatively affects system accuracy and functionality." Their control schematics is described in Figure 1.

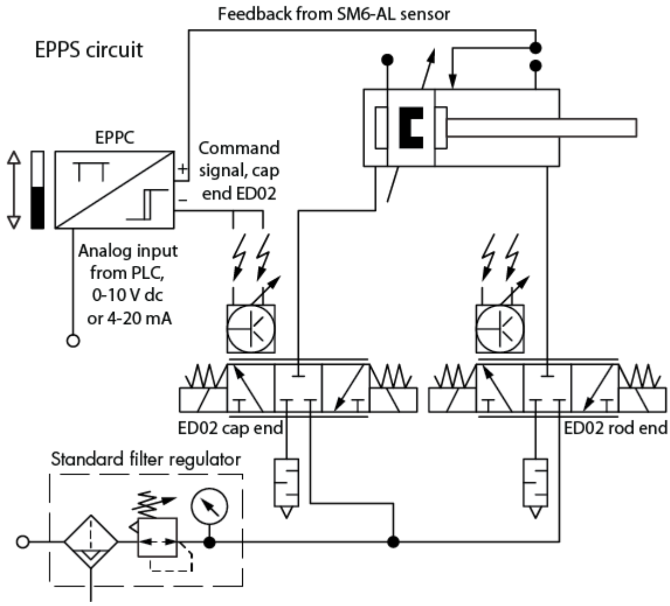


Figure 1
Typical EPPS system using a rod-type pneumatic cylinder [3]

Controlling pneumatic cylinders has drawn interest of many other researchers and skilled amateurs as well.

There are studies of pneumatic system control modelling using Graph model-based analysis of technical systems [4], and [5].

The density of the working medium in hydraulics is 750-800 kg/m^3, while in pneumatics is 1.2-1.3 kg/m^3 depending on the temperature, and pressure. The difference is in the range about three great-orders. In most of the cases low pressures are used, if the cylinder is controller for position, or speed. Gas density if a function of pressure. As the pressure increases, so does the density. Because the load dependency is related to the gas density, if the operating pressure could be increased, so would the load dependency be decreased.

If moving a piston would be realized not by filling one cylinder chamber, and releasing the opposite, but filling both sides with controlled pressure, the piston would move the same way with having significantly greater pressures in both chambers. In this way position dependency of external load would be significantly decreased as well.

A possible application of this control system can be sky-hook like purely pneumatic vehicle body control, or platform height control in industry, or home applications, or in robotics.

In this paper a pneumatic cylinder modelling and control for industrial applications will be introduced.

2 Descriptions

In the ACIPV conference series different pneumatic modelling theories have been introduced by a variety of authors, mostly for pneumobile pneumatic system modelling, and control. This model of an industrial control model is based on an earlier published model, which has been updated, and modified for the actual task.

Properties of the 2023 compared to earlier published models [6] are:

- The cylinder is a dual chamber type, to be able to modell, and simulate exhaust pressure drop, and pre-charging in both chamber sides.
- A controller block is added to simulate PLC, or any kind of microcontroller. In some case pats of program codes can be copy-pasted from PLC, or controller development environments.
- Directional valves are added, and updated to control gas flow, and direction.
- Mechanical load block has been re-worked to fulfill the special industrial task.

During the modell creation it was always kept in focus that the model should remain as simple as possible, and easy to understand for students, and colleagues.

Figure 2 shows the pneumatic schematics of the modelled system. Its detailed description can be found in earlier ACPIV publications [e.g.: 1].

Figures 3 and 4 show the difference between the top layers of the original and the advanced model. The most significant changes on the top layer are the controller block, and the air-line is added to the air-source beside the bottle.

The bottle, and reductor submodels have not changed because these parts are identical in both systems.

The bottle, body and reductor models are described in the ACIPV 2018 Conference paper. (Szakács, 2018 [6])

The 2019, two-chamber cylinder model is shown on Figure 5.

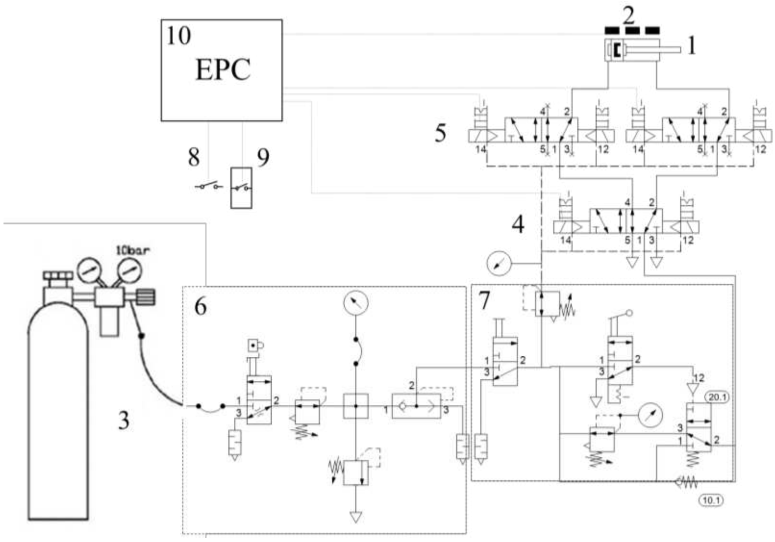


Figure 2
Schematics of the pneumatics

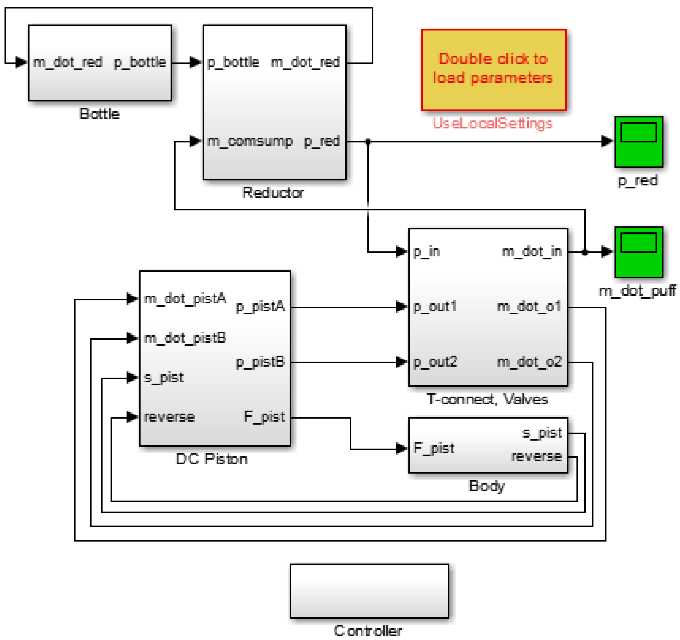


Figure 3
The 2019 Pneumobile pneumatic system modell

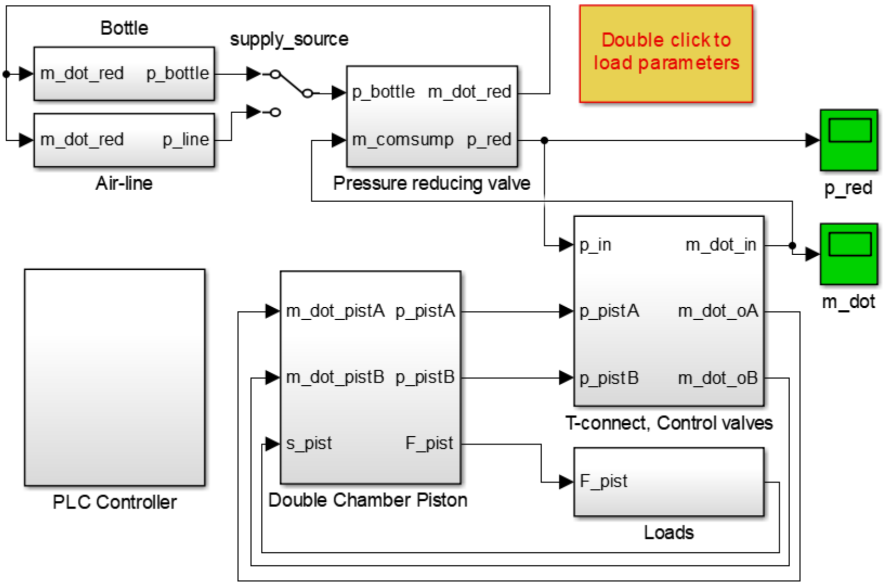


Figure 4

The current industrial piston control modell

The chambers usually are alternatively connected to the supply air, or the environment. One chamber is filled, the other is exhausting. In the new model the cylinder chambers can be independently connected to supply air, or the exhaust, or the inner air can be locked in. (See the schematics on Figure 1) To control which chamber is filled, locked or, which is exhausting the pressure the directional control valves model has been developed.

The 2019 controller model was representing a 5/3 type of directional valve to alternatively charge, and release the two chambers, and locking the air flow. This latest has been originally developed for ECO expansion control during long distance run in the pneumobile competition [1]. The piston direction control signal is calculated in the body submodel. At that state there was no pre-chare control on the engine. See on Figure 6.

The operation of the 5/3 directional valve is explained in Figures 6-8. Both channels fill, or release on the same time possible only in that modell, 5/2, 5/3 valves are not able to operate in that way.

Figure 9 shows a case when the chambers are alternately filled/released, as real valves do.

In the 2023 model the 5/3 directional valve has been replaced by a dual 3/3 one. By this mean both chambers of the cylinder can be independently filled, locked, or released not only alternately, but completely independently, also parallel if so needed.

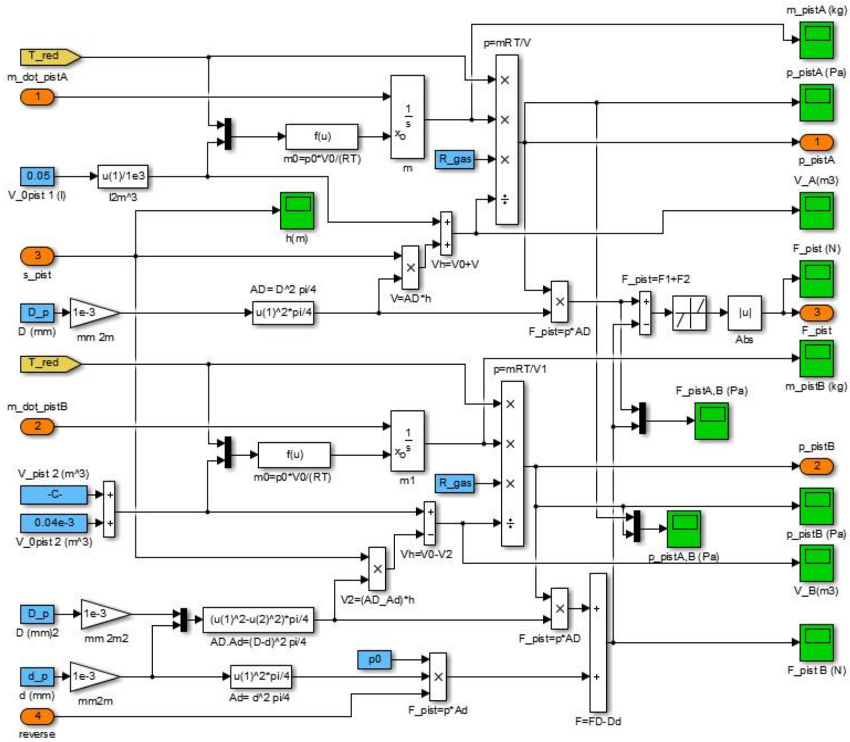


Figure 5

The 2019, dual chamber piston modell

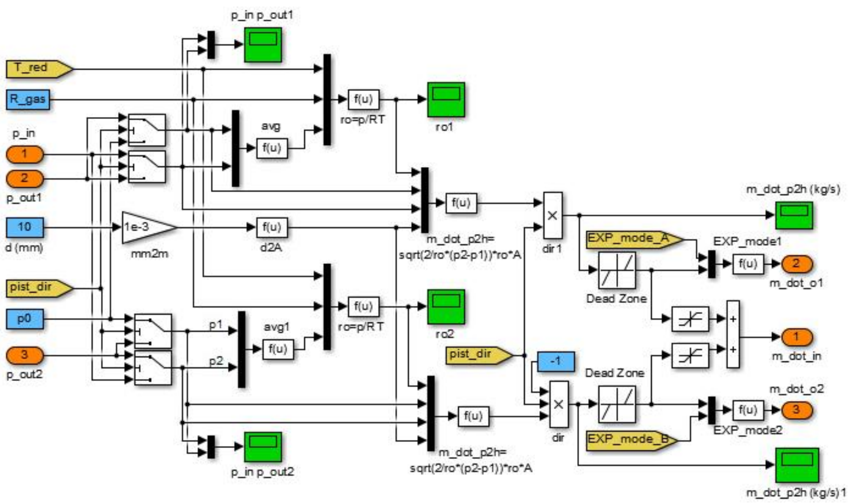


Figure 6

5/3 valves, and T-junction modell

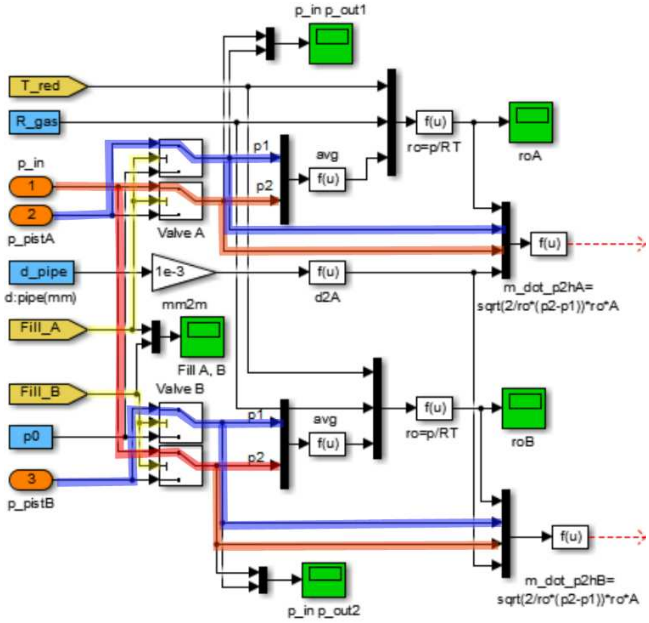


Figure 7
Operation of 5/3 valves: both chamber fill

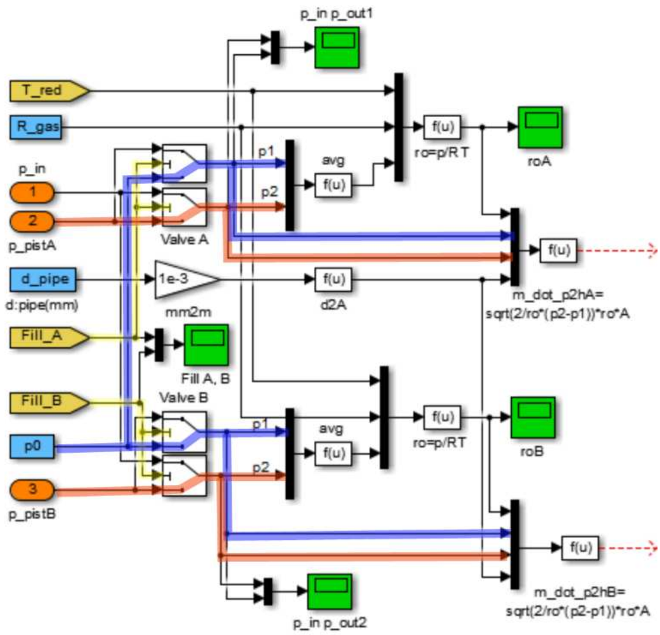


Figure 8
Operation of 5/3 valves: both chamber release

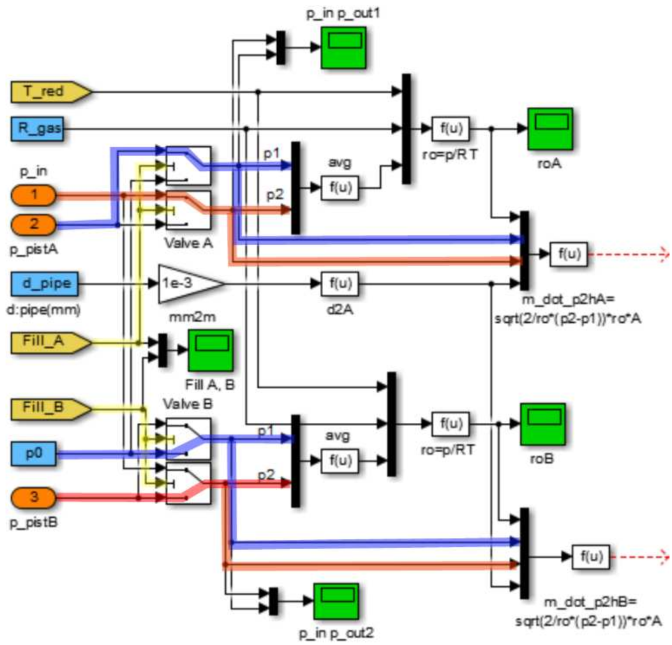


Figure 9
Operation of 5/3 valves: Chamber A fill, B release

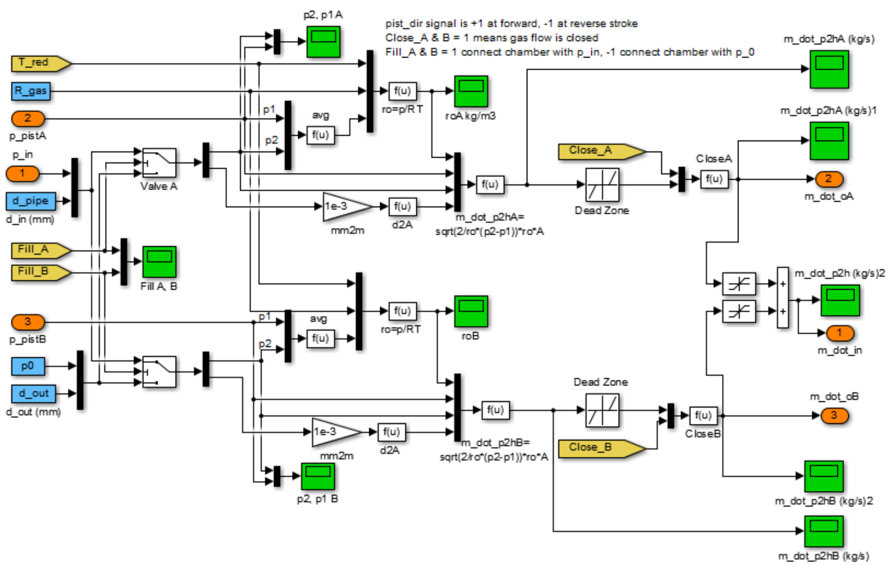


Figure 10
Dual 3/3 valves with differential intake and exhaust

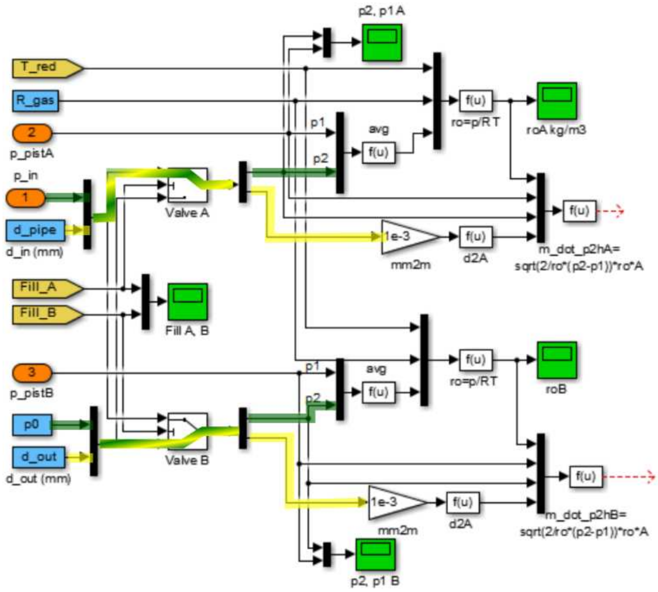


Figure 11

Dual 3/3 valves with differential intake and exhaust

Figure 11 shows the mechanical, and load model of the cylinder. In this modell block the mechanical motion of the piston, and the connected load is modelled. Acceleration, speed, and piston-rod position are calculated, as well as friction force, and damping. Piston end positions are detected, which limits piston motion, and further accelerating forces.

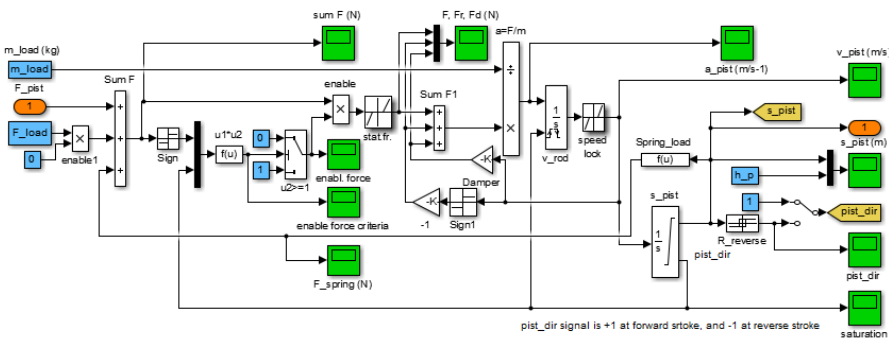


Figure 12

The mechanical motion, and load modell

Figure 13 shows the external loads, spring force (can be disabled in the modell) friction force, and variables used in the modelling.

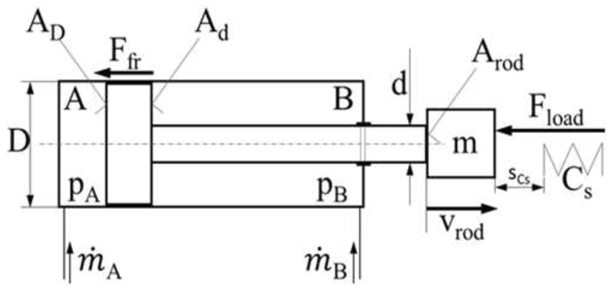


Figure 13
Annotation of the piston used in the modell

$$F = p_A \cdot A_D - p_B \cdot A_d - F_{fr} - p_0 \cdot A_{rod} \quad (1)$$

Equation (1) explains the force balance (including spring load, but excluding damping force).

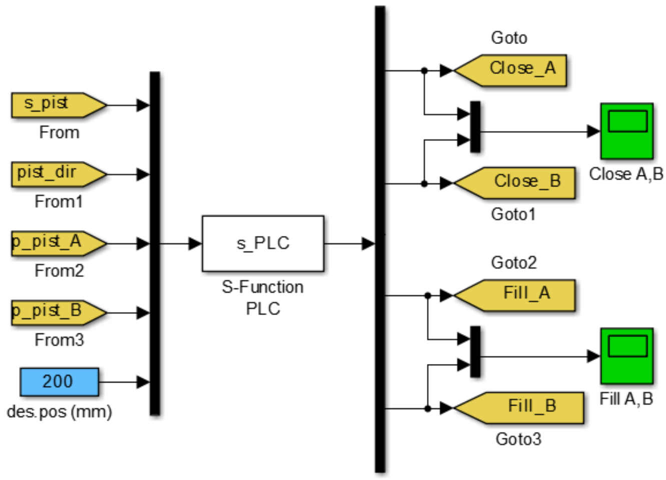


Figure 14
The controller modell

The controller is not connected to the other blocks of the model by data line, which represent that the electric control does not mix pneumatic, or mechanic lines in the modell. Figure 14 shows an example of the PWM position control code.

In part "A: Action detection" the fill action is detected, and the action trigger time is recorded based on trigger occurrence timepoint and PWM period length Fill action is pulled high, so in this period no more action detection is done till the action end.

```
% PWM Position Control
% ****************************
% ******** <PWM Fill> ********
% ****************************
% Part A: Action detection
if (des_pos+1)>s_pist*1000
    if fill_action == 0
        fill_end = t+0.1;% PWM period time 0.1s
        fill_action=1;
    elseif fill_end<t
        fill_action=0;
    end
end
% Part B: PWM ratio determination
if fill_action
delta_s=des_pos-s_pist*1000;
    if delta_s>100 delta_t=0.1; %    100% PWM
    elseif delta_s<10 delta_t=0.001;%  1% PWM
    elseif delta_s<20 delta_t=0.01;%  10% PWM
    elseif delta_s<60 delta_t=0.03;%  30% PWM
    else              delta_t=0.05;%  50% PWM
    end
%   delta_t=0.01;% fixed 10% PWM
% Part C: PWM On-OFF realization
    if t<(fill_end-(0.1-delta_t)) % PWM ON
        Fill_A=1;
        Fill_B=-1;
        Close_A=0 ;
        Close_B=0 ;
    elseif t < fill_end            % PWM OFF
        Fill_A=0;
        Fill_B=-1;
        Close_A=1 ;
        Close_B=0 ;
    end
end
% ******** </PWM Fill> ********
```

Figure 15

Example code for PWM cylinder chamber fill

In part “B: PWM ratio determination” the PWM ON period is calculated based on the difference between desired, and actual piston position. PWM ON period is greater when difference is higher, and getting less when desired position is being achieved.

In part “C: PWM ON-OFF realization” as the name indicates the fill, and the hold states are switched. For the PWM ON period fill is ON, and hold (close) is OFF. For the PWM OFF period fill is ON, and hold (close) is also ON, so there is no actual fill for this period.

This modell code does not include directional control valve switch-on (tF), and switch-off (tE) times. In case of the Aventics 3/2-directional valve, Series CD12 (Part No.: 5724500920) tF=34, tE=90 ms [7].

The control code can control directional valve states either by time, or by event.

Event control can be based on piston position, chamber pressure, piston force, etc. These variables can be passed through the PLC code as global variables, or through the input of the s-function. In Figure 14 variables passed through the input are piston position, piston motion direction, chamber pressure A, and B, and the desired position.

Examples for time control can be seen in the results chapter Simulation 1 part.

3 Results

The first simulations have been run to present the differences between released, and filled counter-chamber (chamber B)

Simulation 1:

In this pair of simulation simple piston motion will be presented once regular fill, and then fill against pressurized opposite chamber.

p_{set}=4 bar, d_{in}=3 mm, d_{out}=10 m, m_{load}=500 kg, h_p=400 mm, D_p=80 mm, d_p=36 mm

Set 1.1:

```
if and (t>=0.5,t<1)
    Fill_A=1;
    Fill_B=-1;
    Close_A=0;
    Close_B=0;
end
```

Figure 16
Piston rod position (Scope s_pist (m) in Loads modell) set1.1

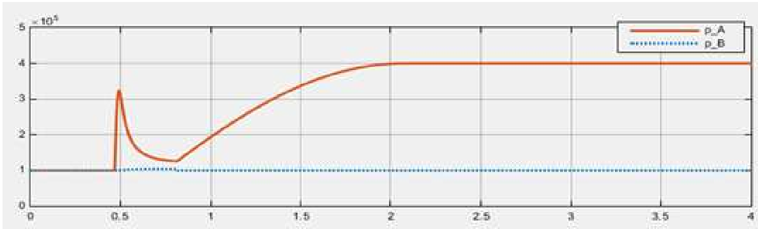


Figure 17
Chamber A, and B pressures (p_pistA,B (Pa) scope in Double Chamber Piston modell), set1.1

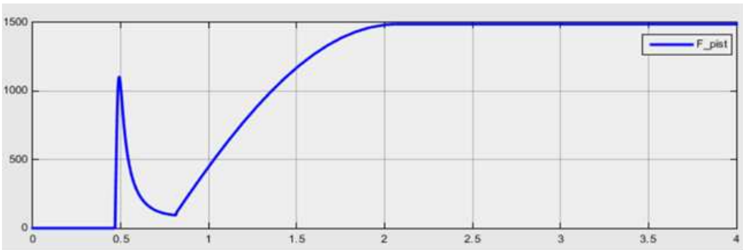


Figure 18
Piston force (F_pist (N) scope in Double Chamber Piston modell), set1.1

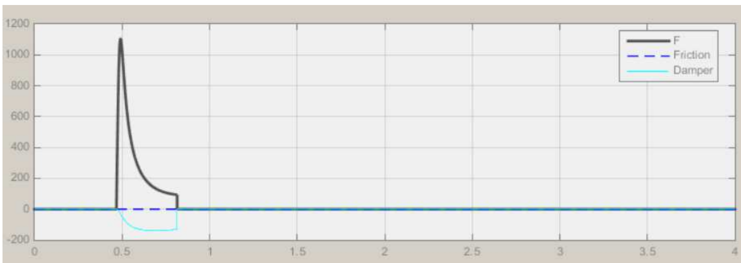


Figure 19
Acting force, and resistances (Scope F, Fr, Fd (N) in Loads modell) set1.1

Set 1.2:

```
if and (t>=0.5,t<1)
    Fill_A=1;
    Fill_B=1;
    Close_A=0;
    Close_B=0;
end
```

Figure 20
Piston rod position (Scope s_pist (m) in Loads modell) set1.2

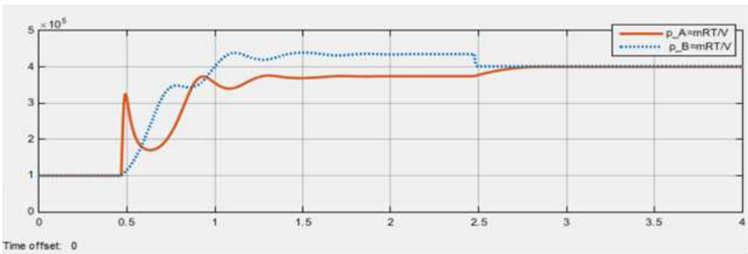


Figure 21
Chamber A, and B pressures (p_pistA,B (Pa) scope in Double Chamber Piston modell), set1.2

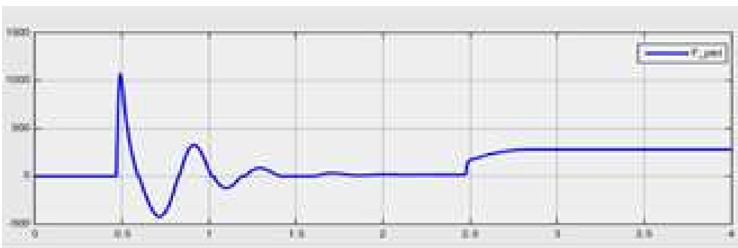


Figure 22
Piston force (F_pist (N) scope in Double Chamber Piston modell), set1.2

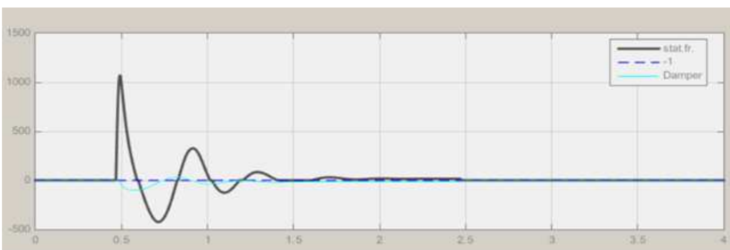


Figure 23
Acting force, and resistances (Scope F, Fr, Fd (N) in Loads modell) set1.2

Evaluation of simulation 1, sets 1.1, and 1.2:

The main difference between sets 1.1, and 1.2 can be seen is figure 17 and 21 regarding the chamber pressures, and figure 16 and 20 regarding piston position. The piston travelled at lower speed of course, but the chamber pressures are significantly greater on both sides. This results higher gas density, thus less dependency on external loads.

Simulation 2:

In this simulation the PWM piston position control will be demonstrated without opposite chamber pressuration.

p_{set}=6 bar, d_{in}=2 mm, d_{out}=10 mm, m_{load}=20 kg, h_p=400 mm, D_p=80 mm, d_p=36 mm

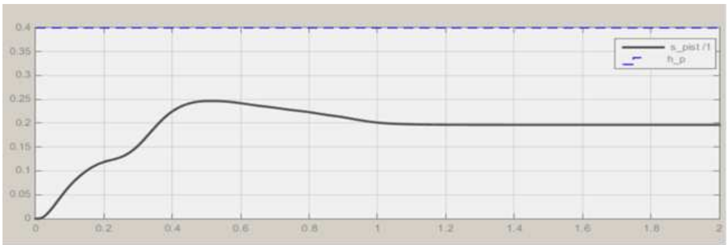


Figure 24
Piston rod position (Scope s_pist (m) in Loads modell) set2

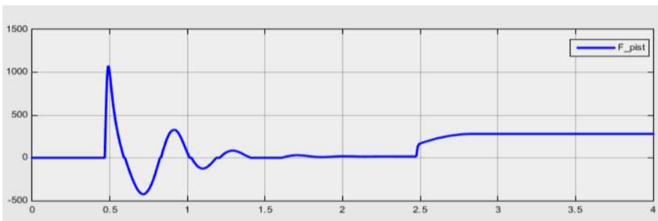


Figure 25
Chamber A, and B pressures (p_pistA,B (Pa) scope in Double Chamber Piston modell), set2

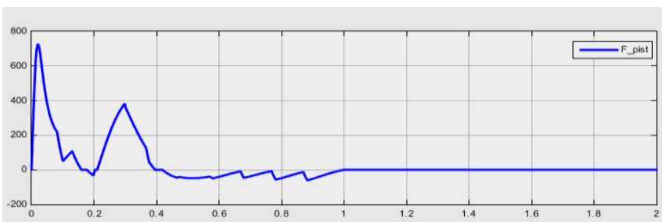


Figure 26
Piston force (F_pist (N) scope in Double Chamber Piston modell), set2

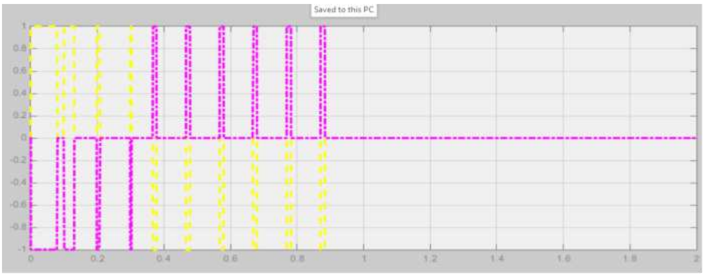


Figure 27
A and B chamber Fill (Scope Fill A,B in PLC Control modell) set2

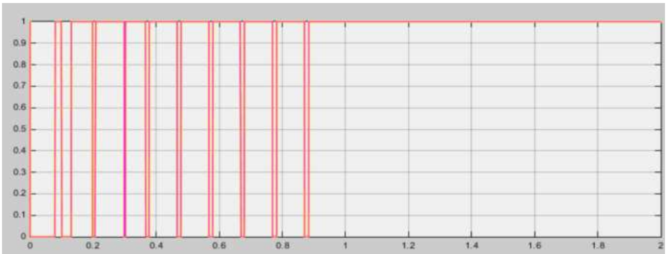


Figure 28
A and B chamber Close (Scope Close A,B in PLC Control modell) set2

4 Outlook

The further development of this modell will include a Graphical User Interface (GUI), and a graphical representation of the cylinder including piston motion, and chamber pressure indication by colors to help better usage even for those who are not expert in coding in m-code, and programming in Simulink.

Fine-tuning of the controller logic is needed, and more simulation must be performed for PWM-control with pressurized opposite chambers, to modell load changing reactions, and comparison of position control with, and without pressurized opposite chamber in varying external load conditions.

Switch-on and switch-off delays of pneumatic directional valves has to be included in the modell, and control code must be corrected accordingly.

Air leakage have to be included in the modell to simulate dithering.

Conclusions

This research has proved that position controlling of pneumatic cylinders using PWM position control is possible. Also, that positioning the piston with a controller that pressurizes both cylinder chambers decrease external load dependency with a factor of 5-10.

Expectations during modell, and controller development indicated that vibrations may have an issue for control. Usage of a vibration damper in the controller system may be necessary.

References

[1] Emerson's AVENTICS™ Pneumobile homepage. https://en.pneumobil.hu/archives/pneumobile_2021

[2] Haitao Qi, Gary M. Bone, and Yile Zhang: Position Control of Pneumatic Actuators Using Three-Mode Discrete-Valued Model Predictive Control. Actuators 2019, 8, 56; doi:10.3390/act8030056 www.mdpi.com/journal/actuators

[3] Steve Vincent, John Bridges: Positioning with Air. Power & Motion Tech April 14, 2015, https://www.powermotiontech.com/pneumatics/pneumatic-valves/article/21884761/positioning-with-air

[4] Pokorádi, László: Graph model-based analysis of technical systems Pokorádi, László, IOP Conference Series: Materials Science and Engineering 1757-8981 1757-899X 393 (1) Paper: 012007, 8 p. 2018

[5] Pokorádi, László, Logical Tree of Mathematical Modeling, Theory and Applications of Mathematics & Computer Science 2067-2764 2247-6202 5 (1) pp. 20-28 2015

[6] Szakács, T: Pneumatic modelling of a pneumobil: In: Pokorádi, László (editor.) Proceedings of the 2nd Agria Conference on Innovative Pneumatic Vehicles ACIPV 2018 Eger, Hungary: Óbudai Egyetem, (2018) pp. 25-30, 6 p.

[7] Aventics catalogue. Directional valves, Electrically operated, Series CD12 https://boschrexroth.africa/en/series-cd12-electrical-valves#