TEHNIČKI GLASNIK

TEHNIČKI GLASNIK / TECHNICAL JOURNAL – GODIŠTE / VOLUME 18 – BROJ / ISSUE 3

RUJAN 2024 / SEPTEMBER 2024 – STRANICA / PAGES 337-506

/

SVEUČILIŠTE SJEVER / UNIVERSITY NORTH – CROATIA – EUROPE

ISSN 1846-6168 (PRINT) / ISSN 1848-5588 (ONLINE)

TECHNICAL JOURNAL

ISSN 1846-6168 (Print) ISSN 1848-5588 (Online)

TEHNIČKI GLASNIK - TECHNICAL JOURNAL

Scientific-professional journal of University North

Volume 18 Issue 3
Varaždin, September 2024 Pages 337-506

Editorial Office:
Sveučilište Sjever / University North – Tehnički glasnik / Technical journal
Sveučilišni centar Varaždin / University Center Varaždin
Jurja Križanića 31b, 42000 Varaždin, Croatia
Tel. ++385 42 493 338, Fax.++385 42 493 336
E-mail: tehnickiglasnik@unin.hr
https://tehnickiglasnik.unin.hr
https://www.unin.hr/djelatnost/izdavastvo/tehnicki-glasnik/
https://hrcak.srce.hr/tehnickiglasnik

Founder and Publisher:
Sveučilište Sjever / University North

Print:
Centar za digitalno nakladništvo, Sveučilište Sjever

All manuscripts published in journal have been reviewed.
Manuscripts are not returned.

The journal is free of charge and four issues per year are published
(In March, June, September and December)
Circulation: 100 copies

Journal is indexed and abstracted in:
Web of Science Core Collection (Emerging Sources Citation Index - ESCI), Scopus, EBSCOhost Academic Search Complete, EBSCOhost – One Belt, One Road Reference Source Product, ERIH PLUS, CITEFACTOR – Academic Scientific Journals, DOAJ – Directory of Open Access Journals, Hrčak – Portal znanstvenih časopisa RH

Registration of journal:
The journal "Tehnički glasnik" is listed in the HGK Register on the issuance and distribution of printed editions on the 18th October 2007 under number 825.

Preparation ended:	Published (online):	Published (print):
July 13, 2024	July 15, 2024	September 5, 2024

CONTENT **I**


ELSEVIER
Postillion Hotel,
Rotterdam, The Netherlands
25 - 28 May 2025
Transportation Research Symposium 2025
Elsevier's Transportation Research (TR) family of journals covers all scientific and practical aspects of transportation research. The Transportation Research Symposium (TRS) is an international conference organized by the TR editors and is equally comprehensive in scope. Other Elsevier transport journal editors will organize themed sessions within TRS.
TRS will give you an outstanding opportunity to interact with the editors of many of the most important journals in the transportation field, to learn about current research, and to present your own work to expert audiences.
Supporting Organization
WCTRS
Submit your abstract by: 15 November 2024
For further information please visit the website:
www.elsevier.com/events/conferences/all/transport-research-symposium

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20220426135403

Original scientific paper
Received: 2022-04-26, Accepted: 2023-08-21

Utilization of "Intersection" Methodology for Optimization with Many Objectives in Designed Experiments of Material Processing

Maosheng Zheng*, Jie Yu

Abstract: Material processing involves many factors, and evaluation of final quality of a product also relates to many indexes from different respects. Thus the optimization of qualities and processing of a product is an optimization problem with many objectives (MOO) inevitably. Although there are some approaches proposed to deal with optimization problem with many objectives in nowadays, the inherent shortcomings in these approaches make them puzzled, which include their missing standpoint and use of additive algorithm containing subjective factors. Currently, an "intersection" methodology for optimization with many objectives is proposed by initiating a novel idea of favorable probability to depict the favorite degree of an alternative in optimal option impersonally, which aims to characterize the concurrent optimization of many objectives in a system in spirits of probability theory and set theory. In this article, some regulations are put forward for performing optimal option of material processing parameters in designed experiments of response surface methodology, orthogonal experiment design and uniform experiment design by means of the total / global favorable probability. In the treatment, the total / global favorable probability of an alternative is the decisive indicator in the optimal option uniquely, which transfers the optimization problem with many objectives into a mono-objective one. The result indicates that the novel approach can be employed to deal with the optimal problem of designed experiments for material processing rationally.

Keywords: favourable probability; "intersection"; material processing; optimization of many objectives; scheme selection

1 INTRODUCTION

Material processing involves many factors, and the final evaluation of a product quality also relates to many indexes from different respects. Thus the optimization of processing parameters and qualities of a product is an optimization problem with many objectives (MOO), which aims to seek a set of appropriate processing and machining technique that guarantees the multi-indexes of quality reaching to compromised optimum simultaneously [1-4].

Material processing and option have a long history since the early work [1-4], many approaches have been proposed to optimize and analyze the processing factors and parameters involved in material processing process [5, 6]. Usually, parametric optimization in material processing could be conducted with designed experiments which contain many objectives [5-8], and regression analysis as well [9-13]. However, there exist inherent shortcomings in the previous MOO approaches due to their use of additive algorithm containing personal factors, in addition to their fatal normalization or scaling processes, such as those in MADM, MOORA, AHP, VIKOR and TOPSIS, etc., which make them puzzled and problematic [5, 6]. In fact, above mentioned approaches attribute to have algorithms only but without any standpoint on optimization with many objectives, which do not give the essential target of optimization with many objectives, i.e., the previous approaches of MOO could not tell us what is the essence and target of MOO [5, 6].

Currently, an "intersection" methodology for optimization with many objectives was proposed to treat the MOO problem by creating a novel idea of favorable probability to depict the favorite degree of an alternative in optimal option impersonally [5, 6]. In the novel methodology, all objectives are taken as "events" within a system individually, thus the concurrent optimization problem of the many objectives is the optimization of the system, which is the overall/global optimization of the system from the viewpoint of system theory. Furthermore, all utility indexes of performances of candidate are classified into beneficial or unbeneficial type in the optimal process, each utility index of the performance quantitatively proffers a partial favourable probability in linear manner, and the product of all partial favourable probabilities results in the total/global favourable probability of a candidate in the spirits of probability theory and set theory, which is the total/global decisive index in the optimal option process uniquely. The novel methodology is a promised methodology that could be utilized to deal with various optimization problems concerning many objectives.

In this paper, some regulations are proposed for conducting optimal option of material processing parameters in designed experiments by means of the total/global favourable probability in respect of the novel approach.

2 FUNDAMENT OF THE NOVEL OPTIMIZATION METHODOLOGY WITH MULTIPLE OBJECTIVES

The core points of the "intersection" algorithm of optimization with many objectives are as follows:

A. All objectives/attributes in the optimization problem with many objectives are taken as "events" within a system, which leads to the concurrent optimization of many objectives as the optimization of the system.

B. The "concurrent optimization" of many objectives is treated by multiplication algorithm, which reflects the feature of "joint probability" and "intersection" in spirits of probability theory and set theory.

C. A novel idea of favourable probability is proposed to depict the favourite degree of an alternative in optimization option.

D. The treatment for utility indexes of performance of both beneficial and unbeneficial types is equivalent and conformable.

3 REGULATIONS OF OPTIMAL OPTION IN DESIGNED EXPERIENTS BY MEANS OF THE NOVEL OPTIMIZATION METHODOLOGY WITH MULTIPLE OBJECTIVES

As the global/total favourable probability of an alternative is the decisive index in the optimal option process uniquely, the global/overall optimization of the system could be focused on this index. Therefore, regulations are proposed for conducting optimal option of material processing parameters in designed experiments as follows.

3.1 Orthogonal Experimental Design

As to orthogonal experimental design, range analysis could be undoubtedly conducted for the total/global favourable probability. Furthermore, the optimal configuration with appropriate experimental parameters could be obtained correspondingly, which is the optimal option that might be in accordance with the maximum total/global favourable probability.

3.2 Uniform Experiment Design and Response Surface Methodology

As to uniform experiment design and response surface methodology, the total/global favourable probabilities of all candidates of the designed experiment are employed to conduct a regression analysis first, a regressed formula of the total/global favourable probability *vs* input variables could be established. Furthermore, maximizing the total/global preferable probability is performed to obtain its maximum value at specific values of input variables. Subsequently, further regression for each response is conducted to get its regressed expression. Afterwards, a properly compromised result of each response is obtained by entering the input variables as the specific values in each regressed expression.

4 ULTILIZATION OF THE NOVEL OPTIMIZATION METHODOLOGY WITH MANY OBJECTIVES IN MATERIAL PROCESSING

4.1 Optimal Design of Machinability Characteristics of Electrical Discharge Machining by Using Response Surface Methodology

Hosni et al conducted design experiments with responses of machinability features of electrical discharge machining by means of response surface methodology, the design is shown in Tab. 1 [14]. The concentration of Span-20 surfactant C_s, and concentration of chromium powder C_p were taken as two independent variables (parameters) for the study. The objective responses in the study included wear rate of electrode *EWR* (mm^3/min), surface roughness *Ra* (mm) and removal rate of material *MRR* (mm^3/min) [14]. The experimental results are cited and shown in Tab. 2.

In applying the newly developed "Intersection" methodology, the removal rate of material *MRR* (mm^3/min) is classified into beneficial performance index, while the wear rate *EWR* of electrode (mm^3/min) and surface roughness *Ra* (mm) belong to unbeneficial performance index. Tab. 3 shows the partial components of favorable probability of each evaluation response and the total / global favorable probability of each experimental option, as well as comparative rank.

Tab. 3 shows that the scheme 2 exhibits the optimal result, which is followed by schemes 11, 9 and 6, et al.

Table 1 Parameters and levels of electrical discharge machining

Parameter	Level		
	1	2	3
C_s (g/L)	0	5	10
C_p (g/L)	0	2	4

Table 2 Experimental design and results of electrical discharge machining

No.	C_s (g/L)	C_p (g/L)	*EWR* (mm^3/min)	*MRR* (mm^3/min)	*Ra* (mm)
1	5.00	4.00	0.0103	41.0399	5.0300
2	10.00	2.00	0.0000	44.0563	4.4367
3	10.00	0.00	0.0186	37.5628	4.4600
4	0.00	2.00	0.0211	43.1316	5.5600
5	10.00	4.00	0.0086	39.7328	4.9167
6	5.00	2.00	0.0040	40.4615	5.0333
7	0.00	4.00	0.0203	40.8241	5.2500
8	0.00	0.00	0.0343	32.5657	6.2500
9	5.00	2.00	0.0041	41.1986	4.6767
10	5.00	0.00	0.0283	37.9382	4.9300
11	5.00	2.00	0.0021	42.6777	4.5300

Table 3 Favorable probability and comparative rank

No.	Favorable probability				Rank
	MRR	*EWR*	*Ra*	*Global* $P_i \times 10^3$	
1	0.0930	0.1063	0.0905	0.8953	6
2	0.0999	0.1520	0.1000	1.5183	1
3	0.0851	0.0696	0.0997	0.5903	7
4	0.0978	0.0587	0.0821	0.4710	8
5	0.0901	0.1140	0.0924	0.9477	5
6	0.0917	0.1343	0.0905	1.1147	4
7	0.0925	0.0621	0.0870	0.4998	9
8	0.0738	8.86E-07	0.0711	4.65E-06	11
9	0.0934	0.1338	0.0962	1.2023	3
10	0.0860	0.0266	0.0921	0.2108	10
11	0.0967	0.1426	0.0985	1.3596	2

4.2 Optimization Treatment with Many Objectives on Turning Process of Steel by Using Orthogonal Design Method

In general, the tool life will reduce with the increase of productivity in turning process [15]. Therefore, it is necessary to conduct optimal design of turning process parameters so as to ensure the coordination between goals in production process.

Trung [15] once used SKS3 steel to study the parametric optimization with orthogonal experimental design (OED) $L_9(3^4)$ in turning process. The tool used in the study was coated with TiN. The OED contains four input variables, each with three levels.

The test design and test results are cited in Tab. 4 [15]. Wherein, the input variables are: v_c cutting speed, f feeding speed, a_r cutting width, and a_p cutting depth; The objective responses include surface roughness of the sample *Ra*, and cutting removal rate *MRR*. *Ra* belongs to unbeneficial attribute index, while *MRR* belongs to beneficial attribute index. The evaluation results of PMOO are shown in Tab. 5.

Table 4 Results of the turning process of SKS3 steel by using OED [15]

No.	Input variable				Response	
	v_c (m/min)	f (mm/rev)	a_r (mm)	a_p (mm)	Ra (mm)	MRR (mm^3/min)
A1	80	0.05	4	0.1	0.970	25.465
A2	80	0.10	8	0.3	1.085	305.577
A3	80	0.15	12	0.5	2.032	1145.916
A4	100	0.05	8	0.5	0.746	318.310
A5	100	0.10	12	0.1	0.609	190.986
A6	100	0.15	4	0.3	1.001	286.479
A7	120	0.05	12	0.3	0.858	343.775
A8	120	0.10	4	0.5	0.326	381.972
A9	120	0.15	8	0.1	1.083	229.183

Tab. 5 shows the assessments of the partial and global favorable probabilities of *Ra* and *MRR* in the orthogonal test design.

Table 5 Evaluations of preferable probabilities under the orthogonal test design

No.	Favorable probability			Rank
	P_{Ra}	P_{MRR}	Global $P_i \times 10^2$	
A1	0.110934	0.00789	0.087522	9
A2	0.101742	0.094674	0.963239	4
A3	0.026055	0.355030	0.925029	6
A4	0.128836	0.098619	1.270575	3
A5	0.139786	0.059172	0.827135	7
A6	0.108456	0.088757	0.962626	5
A7	0.119885	0.106509	1.276882	2
A8	0.162404	0.118343	1.921942	1
A9	0.101902	0.071006	0.723565	8

From Tab. 5, the Test No. A8 is with the maximum value of the global favorable probability P_i, it could be directly chosen as the optimal scheme in orthogonal test design with many objectives at the first glance.

Furthermore, Tab. 6 shows the assessments of range analysis of the global favorable probabilities of the strengthened plate in drawing process by using OED.

Table 6 Evaluations of range analysis of the global preferable probabilities of turnin process

Level	v_c	f	a_r	a_p
Level 1	0.6586	0.8783	0.9907	0.5461
Level 2	1.0201	1.2374	0.9858	1.0676
Level 3	1.3075	0.8704	1.0097	1.3725
Range	0.6489	0.3670	0.0239	0.8264
Order	2	3	4	1

The range analysis in Tab. 6 shows that the impact order of input variables reduces from a_p, v_c, f to a_r. As a result, the optimal option is a_p3-v_c3-$f2$-a_r3, which only differs from experimental scheme A8 by the value of the weakest input parameter a_r, so the optimized configuration is really close to the experimental scheme A8 [15].

4.3 Multi-objective Uniform Experiment Design in Isoprene Rubber Formulation

Yao once studied the composition design of isoprene rubber [16]. In the test of isoprene rubber formulation, three input variables were investigated and their dosage ranges in parts are: x_1 (semi-reinforced carbon black): 20-40, x_2 (sulfur): 0.8-2.0, x_3 (TMTD): 0.8-2.2. The responses include tensile strength Y_1 (MPa), elongation at break Y_2 (%), tear strength Y_3 (kN/m) and permanent deformation at break Y_4 (%).The uniform design table $U_9(9^5)$ is chosen for uniform design of experiment (UDE) [16].

Tab. 7 shows the design and experimental results of responses, Y_1, Y_2, Y_3 and Y_4. In the optimal design, Y_1, Y_2 and Y_3 are beneficial performance utility indexes of the division in the new methodology [5, 6], while Y_4 is attributed to the unbeneficial performance utility index. The partial components of favorable probability of above utility indexes together with global favorable probabilities are shown in Tab. 8. The data in Tab. 8 shows that the test No. 4 indicates the maximum value of the global favorable probability P_{t1}, therefore it can be taken as the optimal option in the UDE with multi-response from the direct observation.

Table 7 Design and experimental results of optimization with the responses of Y_1, Y_2, Y_3 and Y_4

Variable				Response			
No.	x_1	x_2	x_3	Y_1 (MPa)	Y_2 (%)	Y_3 (kN/m)	Y_4 (%)
1	20	1.25	1.85	13.920	690.64	42.91	32.72
2	22	1.85	1.50	13.704	703.29	45.36	32.96
3	25	1.10	1.15	14.921	751.36	54.60	31.30
4	27	1.70	0.80	15.193	763.43	57.80	34.10
5	30	0.95	2.025	14.242	729.44	47.85	35.51
6	32	1.55	1.675	13.531	700.58	47.78	35.93
7	35	0.80	1.325	13.504	746.66	50.02	33.85
8	37	1.40	0.975	13.281	726.30	50.71	36.83
9	40	2.00	2.20	12.431	606.54	39.96	43.70

Table 8 Partial components of preferable probability of Y_1, Y_2, Y_3 and Y_4 and global preferable probabilities

Exp. run	P_{Y1}	P_{Y2}	P_{Y3}	P_{Y4}	P_{t1}×10^4
1	0.1116	0.1076	0.0982	0.1181	1.3927
2	0.1099	0.1096	0.1038	0.1174	1.4675
3	0.1196	0.1171	0.1249	0.1221	2.1363
4	0.1218	0.1189	0.1323	0.1142	2.1891
5	0.1142	0.1137	0.1095	0.1103	1.5670
6	0.1085	0.1092	0.1093	0.1091	1.4125
7	0.1087	0.1163	0.1145	0.1149	1.6570
8	0.1065	0.1132	0.1160	0.1066	1.4901
9	0.0997	0.0945	0.0914	0.0873	0.7518

Furthermore, regression analysis of total favorable probably *vs* input variables is conducted, which results in a regressed function,

$$P_{t1} \times 10^4 = 3.5448 + 0.1626x_1 - 4.5063x_2 - 0.0049x_1^2 + \\ +0.9701x_2^2 - 0.0961x_3^2 + 0.0696x_1x_2 - 0.3133x_2x_3, \quad (1) \\ R^2 = 0.999992.$$

The function P_{t1} reaches to its maximum value of $P_{t1} \times 10^4 = 2.7147$ at specific values of input variables $x_1 = 22.12$, $x_2 = 0.80$ and $x_3 = 0.8$, which is much higher than the values of the total favorable probability P_i in Tab. 8.

While, the regression function of response Y_1 is,

$$f_{Y_1} = 18.7356 + 0.2916x_1 - 9.6921x_2 - 0.0108x_1^2 + \\ +2.1053x_2^2 + 0.2473x_3^2 + 0.1839x_1x_2 - 1.268x_2x_3, \quad (2) \\ R^2 = 1.$$

The predicted value for f_{Y1} is 16.0807 at above specific values of input variables $x_1 = 22.12$, $x_2 = 0.80$ and $x_3 = 0.8$.

The regression function of response Y_2 is,

$$f_{Y_2} = 576.9396 + 18.972x_1 - 70.9024x_2 - 0.3377x_1^2 + \\ +36.7414x_2^2 + 2.2947x_3^2 - 0.1002x_1x_2 - 44.8255x_2x_3, \quad (3) \\ R^2 = 0.9994.$$

The predicted value for f_{Y2} is 769.1696 at the specific values of input variables $x_1 = 22.12$, $x_2 = 0.80$ and $x_3 = 0.8$.

The regression function of response Y_3 is,

$$f_{Y_3} = 53.8523 + 2.722x_1 - 48.2493x_2 - 0.0694x_1^2 + \\ +10.1200x_2^2 - 0.9884x_3^2 + 0.9111x_1x_2 - 5.4124x_2x_3, \quad (4) \\ R^2 = 1.$$

The predicted value for f_{Y3} is 60.0092 at the specific values of input variables $x_1 = 22.12$, $x_2 = 0.80$ and $x_3 = 0.8$.

The regression function of response Y_4 is,

$$f_{Y_4} = 34.05144 - 0.5809x_1 - 0.6335x_2 + 0.0120x_1^2 + \\ +1.0011x_2^2 + 2.0612x_3^2 + 0.1679x_1x_2 - 2.8228x_2x_3, \quad (5) \\ R^2 = 1.$$

The predicted value for f_{Y4} is 29.6786 at the specific values of input variables $x_1 = 22.12$, $x_2 = 0.80$ and $x_3 = 0.8$.

4.4 Multi-objective Optimization for Cleaning Device Design with Uniform Design Method

Sun et al. conducted cleaning device design and optimization with many objectives by means of uniform design, here it is reanalyzed by using the newly proposed method quantitatively.

The crank speed X_1, crank length X_2, airflow speed X_3, wind direction angle (installation angle of fan) X_4 and lower screen surface installation angle X_5, are chosen as the five input variables in the test design, and each variable takes ten levels. The sieve time Y_1 and the falling horizontal displacement Y_2 are used as the evaluation responses, which have the performance characters of the smaller the better for the optimization [17], so Y_1 and Y_2 are classified into the unbeneficial type of utility indexes of performance according to the novel method [5, 6]. The $U_{10}(10^{10})$ of the uniform design table is chosen for experiment design [17]. The experimental design and test results are cited in Tab. 9 [17]. The partial components of favorable probability of above performances together with global favorable probabilities are shown in Tab. 9. The data in Tab. 9 shows that the test option 7 indicates the maximum value of the global favorable probability P_{t2}, therefore it can be taken as one of the optimal combination in the UED with many objectives from the direct observation.

Similarly, the regression analysis of the total favorable probably *vs* input variables is conducted, which results in a regressed function,

$$P_{t_2} \times 10^2 = -28.5651 + 0.1998X_1 + 0.5245X_2 + \\ +0.1347X_3 - 0.0435X_4 - 0.0083X_5 - 0.0004X_1^2 - \\ -0.0219X_2^2 - 0.0149X_3^2 + 0.0003X_4^2, \quad (6) \\ R^2 = 1.$$

Table 9 Test design test results, partial components of favorable probability and global favorable probability

No.	X_1 (r/min)	X_2 (mm)	X_3 (m/s)	X_4 (°)	X_5 (°)	Y_1 (s)	Y_2 (mm)	P_{Y1}	P_{Y2}	$P_{t2} \times 10^2$
1	260	13.5	4	35	11.34	0.4258	579	0.1325	0.0911	1.2063
2	265	14.5	5.5	60	7.34	0.863	536.5	0.0862	0.0992	0.8552
3	270	15.5	7	30	14.34	0.637	542.6	0.1101	0.0980	1.0793
4	275	16.5	3	55	10.34	1.082	562.8	0.0631	0.0942	0.5941
5	280	17.5	4.5	25	6.34	0.799	497.2	0.0930	0.1066	0.9919
6	285	13	6	50	13.34	0.666	554.9	0.1071	0.0957	1.0242
7	290	14	7.5	20	9.34	0.382	482.9	0.1371	0.1094	1.4993
8	295	15	3.5	45	5.34	0.795	546	0.0934	0.0974	0.9096
9	300	16	5	15	12.34	0.354	540.9	0.1400	0.0983	1.3770
10	305	17	6.5	40	8.34	1.325	478.2	0.0374	0.1103	0.4126

The function P_{t2} reaches its maximum value of $P_{t2} \times 10^2 = 1.9737$ at specific values of input variables $X_1 = 277.5626$ r/min, $X_2 = 13.0$ mm, $X_3 = 4.6073$ m/s, $X_4 = 15°$ and $X_5 = 5.34°$, which is obviously higher than the values of the total preferable probability P_{t2} in Tab. 9.

The regression function of response Y_1 is,

$$f_{Y_1} = 48.1641 - 0.2717X_1 - 1.3349X_2 - 0.4256X_3 + \\ +0.0180X_4 - 0.1015X_5 + 0.0005X_1^2 + 0.0485X_2^2 + \\ +0.0453X_3^2 + 0.0049X_5^2, \quad (7) \\ R^2 = 1.$$

The predicted value for f_{Y1} is 0.0508 at the specific values of input variables $X_1 = 277.5626$ r/min, $X_2 = 13.0$ mm, $X_3 = 4.6073$ m/s, $X_4 = 15°$ and $X_5 = 5.34°$.

The regression function of response Y_2 is,

$$f_{Y_2} = -2076.83 + 9.5923X_1 + 164.4545X_2 + \\ +24.8636X_3 + 2.3195X_4 + 6.3X_5 - 0.0176X_1^2 - \\ -5.6364X_2^2 - 4X_3^2 - 0.027X_4^2, \quad (8) \\ R^2 = 1.$$

The predicted value for f_{Y2} is 510.9258 at the specific values of input variables X_1 = 277.5626r/min, X_2 = 13.0 mm, X_3 = 4.6073 m/s, X_4 = 15° and X_5 = 5.34°.

5 CONCLUSION

Through analyzing optimal scheme of material processing with designed experiments of response surface methodology, orthogonal experiment design and uniform experiment design in this paper, it can be seen that that the novel "intersection" methodology optimization with many objectives might be easily combined with designed experiments to treat optimal problem of designed experiments. The global favorable probability of an alternative is the decisive indicator in the optimal option uniquely, which could transfer the optimization problem with many objectives into an optimization problem with mono-objective. The subsequent analysis is focused on total favorable probability in the experimental designs.

Conflict Statement

There is not any conflict of interest in this article.

6 REFERENCES

[1] Ashby, M. F. (2000). Multi-objective optimization in material design and selection, *Acta Mater., 48*, 359-369. https://doi.org/10.1016/S1359-6454(99)00304-3

[2] Ashby, M. F., Bréchet, Y. J. M., Cebon, D. & Salvo, L. (2004). Selection strategies for materials and processes. *Mater. Des.*, 25, 51-67. https://doi.org/10.1016/S0261-3069(03)00159-6

[3] Deshmukh, D. & Angira, M. (2019). Investigation on switching structure materialselection for RF- MEMS shunt capacitive switches using Ashby, TOPSIS and VIKOR. *Transactions on Electrical and Electronic Materials*, 20, 181-188. https://doi.org/10.1007/s42341-018-00094-3

[4] Liu, W. & Yang, Y (2008). Multi-objective optimization of sheet metal forming process using Pareto-based genetic algorithm. *Journal of Materials Processing Technology*, 208, 499-506. https://doi.org/10.1016/j.jmatprotec.2008.01.014

[5] Zheng, M., Wang, Y. & Teng, H. (2021). A new "intersection" method for multi-objective optimization in material selection, *Tehnički glasnik, 15*(4), 562-568. https://doi.org/10.31803/tg-20210901142449

[6] Zheng, M. (2022). Application of probability-based multi-objective optimization in material engineering, *Vojnotehnički glasnik/Military Technical Courier*, 70, 1-12. https://doi.org/10.5937/vojtehg70-35366

[7] Abedi, E., Daneshmand, S., Neyestanak, A. A. L. & Monfared, V. (2014). Analysis and modeling of electro discharge machining input parameters of Nitinol shape memory alloy by de-ionized water and copper Tools. *Int. J. Electrochem. Sci., 9*, 2934-2943. https://doi.org/10.1016/S1452-3981(23)07980-4

[8] Daneshmand, S., Neyestanak, A. A. L. & Monfared, V. (2016). Modeling and investigating the effect of input parameters on surface roughness in electrical discharge machining of CK45. *Tehnički Vjesnik, 23*(3), 725-730. https://doi.org/10.17559/TV-20141024224809

[9] Yerel, S. & Ankara, H. (2011). Process control for a coal washing plant using a range control chart and multidimensional scaling analysis. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 33*(11), 1028-1034. https://doi.org/10.1080/15567030903096998

[10] Ankara, H. & Yerel, S. (2009). Determination of sampling errors in natural stone plates through single linkage cluster method. *Journal of Materials Processing Technology*, 209(5), 2483-2487. https://doi.org/10.1016/j.jmatprotec.2008.05.048

[11] Verel, S., Ankara, H., Konuk, A. & Ozdag, H. (2007). Preventive maintenance policy supported by a quality control chart and Kolmogorov-Smirnov tests: Emet Colemanite Mineral Processing Plant, Turkey. *Minerals & Metallurgical Processing, 24*(3), 152-156. https://doi.org/10.1007/BF03403210

[12] Yerel, S. & Ersen, T. (2013). Prediction of the calorific value of coal deposit using linear regression analysis. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 35(10), 976-980. https://doi.org/10.1080/15567036.2010.514595

[13] Yerel, S. (2016). Assessment of coal deposit using multivariate statistical analysis techniques. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 38*(7), 1002-1006. https://doi.org/10.1080/15567036.2010.540635

[14] Hosni, N. A. J. & Lajis, M. A. (2019). Multi-response optimization of the machining characteristics in electrical discharge machining (EDM) using span-20 surfactant and chromium (Cr) powder mixed. *Materialwiss. Werkstofftech., 50*, 329-335. https://doi.org/10.1002/mawe.201800204

[15] Trung, D. D. (2022) Multi-criteria decision making under the MARCOS method and the weighting methods: applied to milling, grinding and turning processes. *Manufacturing Rev., 9*(3), 1-13. https://doi.org/10.1051/mfreview/2022003

[16] Yao, Z. (1998) The Position the regression analysis has occupied in the data analysis of the uniform design, *Special Purpose Rubber Products, 19*, 186-194.

[17] Sun, S., Wang, J., Ke, J., Wang, Q. & Sun, Z. (2012). Cleaning device design and multi-objective optimization based on uniform design. *Machine Design and Research, 28*, 115-117.

Authors' contacts:

Maosheng Zheng
(Corresponding author)
School of Chemical Engineering, Northwest University,
No. 229, Taibai North Road, Xi'an, 710069, Shaanxi Province, China
E-mail: mszhengok@aliyun.com

Jie Yu
School of Life Science & Technology, Northwest University,
No. 229, Taibai North Road, Xi'an, 710127, Shaanxi Province, China

]ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20230601145322

Original scientific paper
Received: 2023-06-01, Accepted: 2023-08-21

Factors Affecting Public Transportation Use during Pandemic: An Integrated Approach of Technology Acceptance Model and Theory of Planned Behavior

Soheila Saeidi, Somayeh Nazari Enjedani*, Elmira Alvandi Behineh, Kian Tehranian, Seyedalireza Jazayerifar

Abstract: For preventing the spread of COVID-19, health authorities urgently turned their attention to urban public transportation. It is believed that virus transmission is more likely to occur in public transportation due to increased exposure to infected individuals in the closed and crowded spaces of public transport. This study aimed to model effective factors in the use of public transportation systems during a pandemic based on the technology acceptance model (TAM) and the theory of planned behaviour (TPB). The methodology used was structural equation modeling, with 358 Iranian passengers in Tehran participating and data collected through a questionnaire. The data underwent analysis by means of the partial least squares method with the assistance of SMARTPLS software. The results indicate that passenger satisfaction is affected positively and significantly by expectation and service quality. Behavioral control, subjective norm, attitude, perceived usefulness (PU), and perceived ease of use (PEU) each contribute positively and significantly to the formation of intention. Service quality, PU, and PEU positively and significantly affect attitude. Finally, expectation, intention, PU, and PEU positively and significantly affect the use of the public transportation system. Therefore, it can be inferred that amalgamating TPB and TAM can serve as a robust indicator of passengers' inclination towards using public transportation during pandemic situations, as well as their actual usage of it.

Keywords: pandemic conditions; TAM; TPB; use of public transportation system

1 INTRODUCTION

On December 31, 2019, it was reported that the first confirmed case of COVID-19 had emerged in Wuhan, China [1]. However, because of COVID-19's rapid transmission, the virus was estimated to have caused 4.75 million deaths and more than 232 million confirmed cases worldwide as of September 28, 2021. [2]. The transmission of the disease was immediately recognized as mainly occurring through respiratory droplets and contaminated surfaces within the first few days of the epidemic. There was evidence that respiratory droplets could reach nearby people if they were in close contact with an infected person (within one meter). In addition, preliminary evidence suggested that respiratory droplets and secretions could remain on surfaces and objects for prolonged periods. Therefore, it was also thought that the infection was transmitted indirectly when individuals touched contaminated surfaces and subsequently made contact with their mouth, nose, or eyes [3].

Efforts have been made by health officials to address the problem of urban public transportation with the aim of curbing the transmission of COVID-19. A high risk of virus transmission in public transportation can be attributed to the fact that in close and crowded spaces in public transportation, the chances of individuals being exposed to infected persons, insufficient ventilation, and recycled contaminated air are higher, as well as the efforts made to maintain basic hygiene standards [4]. Due to the increased risk of virus transmission in public transportation, there has been a significant decline in its usage, as noted by multiple observations worldwide, not only as a result of the reduced demand but also as a consequence of the regulations of health agencies to limit the number of passengers in public transportation [5, 6]. Now that people are familiar with self-care and coping methods in pandemic conditions and have experienced COVID-19, this study tends to predict in pandemic conditions. This paper utilizes two theories, namely the theory of planned behavior (TPB) and the technology acceptance model (TAM), in addition to service quality, to forecast the usage of public transportation.

1.1 TPB

It is believed commonly that the TPB is one of the most useful theories for predicting behavior. As a way of explaining human behavior, Ajzen [7] developed this theory. In line with this theory, people behave in a rational and motivated manner because of their motivations. In order to explain individual behavior, this theory focuses not only on voluntary control but also on involuntary control [8]. There are four major components of the TPB: attitudes, mental norms, perceived behavioral control, intention, and behavior. According to the TPB, the most important factor is individual intention, which provides the best indication of an individual's choice of actions [9-11]. A public transportation intention is defined as an intention to use public transportation in the study.

The TPB posits that an individual's behavior is a manifestation of their behavioral intentions, with the notion that one's behavior is influenced by their intentions. A mental norm is characterized as an individual's reaction to apparent social pressures, either in favor of or against engaging in a specific behavior, based on the perceived societal influence [12]. The perception of behavioral control reflects an individual's appraisal of the level of effort required to perform a particular behavior [9]. Because of perceived behavioral control, either one can determine behaviors indirectly, by influencing behavioral intentions, or directly, by determining precisely whether the behavior is controlled actually by the discussed behaviour [13]. It is important to understand that attitude towards behavior is a measure of how positive or negative one feels about one's behavior [14]. The factors described above, as well as attitudes toward behavior, subjective norms, and perceptions of behavioral control, all play a significant role in the formation of a behavioral intention, which is also an immediate predeterminant of the actual behaviour, which follows [15-17]. Previous studies [15, 18-21, 59, 61, 66, 67] have used the TPB in the context of public transportation use. Several studies have shown that the TPB plays an important role in explaining the reasons why people use public transportation daily. Therefore, considering the findings of the conducted studies, it is assumed that:

H_1: Perceived behavioral control is effective on intention.
H_2: Subjective norm is effective on intention.
H_3: Attitude is effective on intention.

1.2 TAM

The introduction of new urban transportation can be regarded as a technological advancement from the perspective of passengers [63, 64]. The TAM is a widely used theory in the study of technology acceptance [22]. The TAM has been widely recognized for its specificity and cost-effectiveness, as well as its high predictive power when it comes to technology usage. Empirical research suggests that the TAM is successful in predicting around 40% of system use [23]. Because of the model, it is suggested that PEU and perceived usefulness (PU) of technology play an important role in explaining recent technology usage. PEU and PU are two distinct concepts, with the latter pertaining to an individual's perception that utilizing a particular technology will improve their performance, while the former relates to the degree of complexity associated with its usage [22, 24]. The widespread applicability of the TAM has been explored in various technology-oriented contexts [25], and our study aims to investigate how PEU and PU impact the use of public transportation systems from the perspective of transportation technology acceptance. Prior research has highlighted the significance of these two factors in influencing the utilization of public transportation [17, 24, 26]. As Taylor and Todd [27] have noted, the TAM and the TPB can be used to explain individual behavior related to new technology adoption. Accordingly, we assume that these models will be applicable to our study on technology use in public transportation.

H_4: PEU is effective on PU.
H_5: PEU is effective on attitude.
H_6: PEU is effective on intention.
H_7: PEU is effective on the use of public transportation systems in pandemic conditions.
H_8: PU is effective on attitude.
H_9: PU is effective on intention.
H_{10}: PU is effective on the use of public transportation systems in pandemic conditions.
H_{11}: Intention is effective on the use of public transportation in pandemic conditions.

1.3 Satisfaction

Satisfaction is one's pleasant or unpleasant feelings, which result from mental performance in comparison with expectations [28]. Due to the competitive environment, the challenge of achieving customer satisfaction is more reflected. In order to satisfy customers, institutions and companies should consider implementing new marketing strategies [29]. It is psychologically and emotionally important to satisfy the needs of customers that customers experience When comparing their pre-consumption expectations to the actual service performance experienced post-consumption, customers can evaluate the quality of service received [30, 31, 56, 59, 65]. Customer satisfaction can be increased significantly if they perceive the service performance and when it exceeds their initial expectations, customers tend to have a positive emotional response and are more likely to be satisfied with the service. It has been shown that the reasons for passenger satisfaction are determined by both expectations of the passengers as well as the quality of perceived services by Chou et al. [32] and Koklic et al. [33]. In methods of transportation such as buses, subways, high-speed trains, and aviation, the degree of passenger satisfaction is one of the most important mediators that link passenger expectations and perceptions of the quality of service with the intention of passengers to engage in certain behaviors [31, 34]. Therefore, it is reasonable to assume that:

H_{12}: Satisfaction is effective on intention.
H_{13}: Satisfaction is effective on the use of public transportation in pandemic conditions.

1.4 Expectations

According to Heidari et al. [29], customer expectations refer to the psychological requirements that customers have for reliable or personalized services before making a purchase. Shen et al. [36] discovered that passenger expectations are a crucial factor in determining passenger satisfaction. Furthermore, research has shown that there is a positive correlation between passenger expectations and satisfaction [31, 35]. In this study, expectations are viewed as predictors of satisfaction. Hence, it can be assumed that:

H_{14}: Expectation is effective on satisfaction.

1.5 Service Quality

In the past thirty years, service quality and its outcomes have been researched extensively in the service marketing literature. Service quality is considered a distinguishing factor and a potent competitive tool [37]. Services are perceived as the outcome of customer interactions with service providers, including employees, service environment and facilities, and equipment. Service quality is defined as a sustainable mindset towards achieving service excellence [38], or alternatively, it is described as the disparity between perceptions and expectations [39]. This concept has also been widely discussed in various fields, including public transport services, where it has played an essential role in policy formation aimed at enhancing public transportation services [40]. Numerous empirical studies have established a positive correlation between the quality of service offered by public transportation providers and the behavior of passengers who are more inclined to utilize these services in the future [41]. Emami et al. [42] introduced a GIS-based multi-criteria decision analysis robust framework by considering, socioeconomic, transit ridership, and service frequency factors to improve the bus transit systems. According to Perez et al. [43], the concept of service quality can be divided into five components: tangible, reliable, adaptable, assured, and empathic. In another study conducted by Prasad et al. [44], it was found that the service quality of a railroad system is affected by eight key factors: assurance, empathy, reliability, responsiveness, tangibility, convenience, connectivity, and comfort. Chou et al. [45] discovered that passenger perception of service quality comprises four dimensions: tangibility, convenience, personnel, and reliability, which are also taken into consideration in this study. Several studies have shown in the past that the perceived service quality of public transportation services significantly impacts passenger satisfaction [46,47] as well as PEU and usefulness [48,49]. It is therefore reasonable to assume that the following statements are true:

H_{15}: Perceived service quality is effective on satisfaction.
H_{16}: Perceived service quality is effective on attitude.
H_{17}: Perceived service quality is effective on PEU.
H_{18}: Perceived service quality is effective on PU.

As highlighted in the theoretical literature, the role of TAM and TPB on behavioral intention and public transportation use has been studied extensively. However, a review of empirical research indicates that few studies have modeled the effective factors on the use of public transportation systems during pandemic conditions based on the TAM and TPB frameworks. In order to accomplish this primary objective, the focus of the study is to model the effective factors in the use of public transportation systems during pandemic conditions based on the TAM and TPB frameworks. The conceptual model of the study is developed based on the theoretical framework derived from the literature review, as illustrated in Figure 1. The independent variables in the model include expectations, perceived behavioral control, subjective norm, and perceived service quality. The mediating variables consist of satisfaction, attitude, PU, PEU, and intention, while the dependent variable is the use of public transportation systems during pandemic conditions.

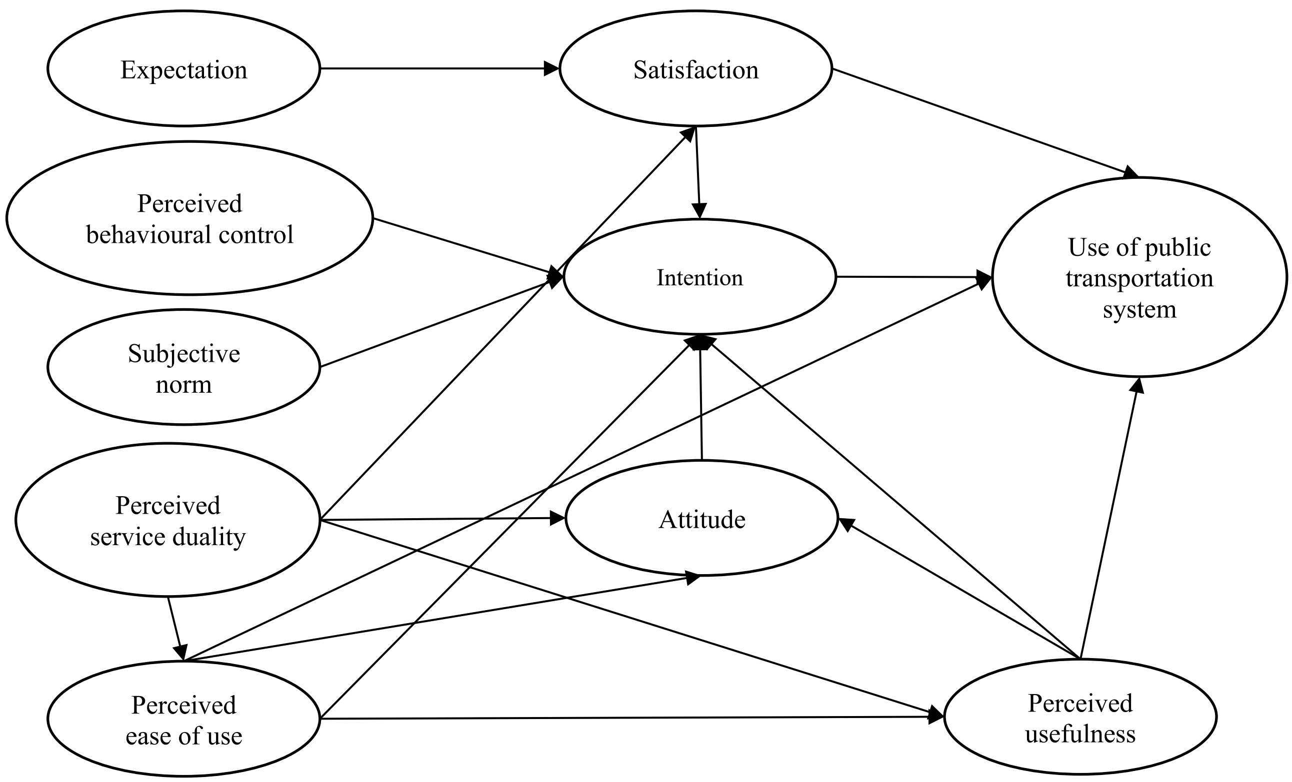


Figure 1 The conceptual model

2 RESEARCH METHODOLOGY

2.1 Statistical Population and Sample

The research was carried out in Iran. The data was collected in Tehran, which has high traffic and an active public transportation system. For this purpose, 500 questionnaires were distributed among citizens of Tehran, out of which 396 questionnaires were returned and 38 questionnaires were removed from analysis due to incomplete answers [57-58, 60]; finally, 358 questionnaires were analyzed. It was found that 69% of participants were males and 31% were females among the participants. It has been found that 34% of respondents were from the 20-30 age group, 31% from the 31-40 age group, 26% from the 41-50 age group, and 9% from the 50+ age group were interviewed.

2.2 Measures

The study employed a questionnaire consisting of 53 items to measure the variables. To measure expectations, the study utilized Yuan et al.'s [31] questionnaire, which consisted of 3 items. Perceived behavioral control was measured using Chen et al.'s [50] questionnaire, which had 4 items. The attitude was also measured using Chen et al.'s [50] questionnaire, which consisted of 6 items. Subjective norm was measured using Chen et al.'s [50] questionnaire that had 3 items. Moreover, Chou et al.'s [45] questionnaire with 18 items was used to measure perceived service quality, including tangibility (4 items), convenience (6 items), personnel (4 items), and reliability (4 items). PU was measured using Chen and Chao's [17] questionnaire with 5 items, while ease of use was measured using the same questionnaire with 5 items. The intention to use the public transportation system during pandemic conditions was measured using Chen et al.'s [50] questionnaire, consisting of 3 items. On a five-point Likert scale, each item in the survey was rated from one to five, ranging from complete disagreement (1) to complete agreement (5).

3 RESULTS

3.1 Measurement Model Testing

To ensure the accuracy of measurement models, it is important to examine the internal consistency (reliability) and discriminant validity of the constructs and instruments involved in the study. In order to gauge the reliability of constructs, three standards were employed: the reliability of every item, the composite reliability of each construct, and the average variance extracted (AVE). The reliability of each item can be ascertained by examining the factor load of each item within its respective construct through confirmatory factor analysis. A factor load of 0.6 or above and statistical significance at a level of 0.01 or better are both indicators of adequate reliability [51, 52]. To obtain the t-values for factor loadings, a bootstrap test using 500 subsamples was employed to determine their significance. The Dillon-Goldstein coefficient (ρ_c) was employed to verify the composite reliability of each construct. To fulfill the third criterion, which requires adequate discriminant validity, an AVE value of 0.50 or greater is required; this indicates that the construct can account for 50% or more of the variance in its markers [53, 54]. Tab. 1 displays the factor loadings, composite reliability, and AVE values for the variables, indicating that the constructs possess adequate and strong reliability.

In a research study, Chin [52] put forward two criteria for evaluating the validity and discriminant validity of constructs. To meet the initial criterion for establishing construct validity, items in a given construct should exhibit higher factor loadings for that construct than for other constructs, indicating that they are more closely related to their intended construct. Gefen and Straub [51] suggested that this difference in factor loadings should be at least 0.1. In order to satisfy the second criterion, it is necessary to confirm that the square root of a construct's AVE exceeds the correlation between that construct and other constructs studied. By meeting this criterion, it can be ensured that the construct is unique from other constructs and is precisely measuring the intended concept. In Tab. 2, details are given about the cross-sectional loads of items on constructs.

Table 1 Results of reliability

Variable	Item	Factor	α	CR	AVE
Expectation	1	0.742	0.781	0.868	0.688
	2	0.854			
	3	0.886			
Perceived behavioral control	1	0.857	0.873	0.913	0.724
	2	0.832			
	3	0.864			
	4	0.851			
Subjective norm	1	0.820	0.802	0.883	0.715
	2	0.847			
	3	0.868			
Tangibility	1	0.872	0.823	0.884	0.659
	2	0.712			
	3	0.90			
	4	0.747			
Convenience	1	0.874	0.896	0.921	0.661
	2	0.709			
	3	0.839			
	4	0.764			
	5	0.833			
	6	0.846			
Personnel	1	0.781	0.789	0.863	0.612
	2	0.833			
	3	0.744			
	4	0.770			
Reliability	1	0.763	0.798	0.868	0.623
	2	0.813			
	3	0.793			
	4	0.786			
Ease of use	1	0.826	0.877	0.910	0.669
	2	0.783			
	3	0.826			
	4	0.845			
	5	0.808			
PU	1	0.843	0.882	0.914	0.681
	2	0.844			
	3	0.844			
	4	0.842			
	5	0.747			
Attitude	1	0.878	0.933	0.947	0.750
	2	0.884			
	3	0.890			
	4	0.846			
	5	0.865			
	6	0.831			
Intention	1	0.917	0.874	0.923	0.799
	2	0.888			
	3	0.876			
Use of public transportation system in pandemic condition	1	0.847	0.850	0.909	0.769
	2	0.894			
	3	0.890			

In Tab. 2, we can see that the results indicate that each dimension has the highest factor loading on its intended construct, and there is at least a 0.1 difference between the factor loadings on their intended constructs compared to other constructs, indicating good validity of the constructs. Additionally, the second criterion for construct validity is presented in Tab. 3 together with the results from the correlation analysis.

AVE values of all variables in Tab. 3 exceed their correlations with other variables, confirming the satisfactory discriminant validity of the variables. Furthermore, the correlation matrix below the diagonal provides information about the relationships between the variables being studied.

Table 2 Cross-sectional factor loading

	Attitude	Intention	Expectation	Satisfaction	Behavior control	PEU	PU	Service quality	Subjective norm	Use of public transport in pandemic conditions
APT1	0.878	0.470	0.387	0.478	0.543	0.564	0.554	0.462	0.329	0.520
APT2	0.884	0.414	0.346	0.419	0.534	0.522	0.531	0.379	0.342	0.461
APT3	0.890	0.408	0.363	0.474	0.558	0.522	0.502	0.392	0.335	0.463
APT4	0.846	0.464	0.361	0.422	0.466	0.506	0.567	0.372	0.298	0.419
APT5	0.865	0.466	0.359	0.399	0.493	0.494	0.504	0.385	0.288	0.469
APT6	0.831	0.558	0.385	0.429	0.453	0.526	0.429	0.417	0.298	0.579
IPT1	0.583	0.917	0.362	0.499	0.484	0.505	0.489	0.411	0.414	0.426
IPT2	0.466	0.888	0.375	0.483	0.432	0.570	0.499	0.398	0.419	0.449
IPT3	0.453	0.876	0.356	0.508	0.434	0.567	0.538	0.351	0.420	0.544
PBC1	0.508	0.589	0.280	0.362	0.857	0.471	0.530	0.443	0.348	0.555
PBC2	0.447	0.565	0.324	0.388	0.832	0.440	0.500	0.397	0.338	0.431
PBC3	0.407	0.573	0.359	0.439	0.864	0.474	0.526	0.453	0.383	0.559
PBC4	0.420	0.621	0.373	0.467	0.851	0.520	0.486	0.439	0.307	0.452
PE1	0.271	0.280	0.742	0.338	0.257	0.352	0.170	0.239	0.140	0.329
PE2	0.302	0.270	0.854	0.405	0.295	0.238	0.196	0.244	0.197	0.346
PE3	0.438	0.425	0.886	0.613	0.394	0.351	0.355	0.355	0.408	0.453
PEU1	0.482	0.536	0.355	0.330	0.430	0.826	0.346	0.390	0.274	0.475
PEU2	0.409	0.423	0.297	0.288	0.343	0.783	0.284	0.356	0.263	0.402
PEU3	0.436	0.438	0.281	0.348	0.396	0.826	0.267	0.309	0.282	0.449
PEU4	0.575	0.581	0.302	0.392	0.544	0.845	0.410	0.349	0.337	0.528
PEU5	0.405	0.504	0.307	0.348	0.542	0.808	0.442	0.292	0.299	0.515
PS1	0.350	0.399	0.570	0.796	0.350	0.285	0.317	0.347	0.373	0.382
PS2	0.518	0.567	0.522	0.903	0.496	0.426	0.442	0.421	0.272	0.565
PS3	0.337	0.358	0.274	0.727	0.316	0.295	0.325	0.335	0.216	0.411
PU1	0.474	0.547	0.258	0.357	0.532	0.339	0.843	0.414	0.434	0.474
PU2	0.542	0.543	0.237	0.352	0.527	0.420	0.844	0.327	0.345	0.493
PU3	0.475	0.528	0.196	0.327	0.484	0.357	0.844	0.385	0.383	0.449
PU4	0.504	0.503	0.282	0.392	0.474	0.348	0.842	0.379	0.387	0.495
PU5	0.460	0.532	0.308	0.430	0.451	0.329	0.747	0.362	0.417	0.487
SN1	0.276	0.359	0.248	0.306	0.307	0.310	0.411	0.366	0.820	0.258
SN2	0.267	0.358	0.270	0.259	0.316	0.273	0.381	0.302	0.847	0.311
SN3	0.368	0.454	0.315	0.326	0.390	0.322	0.415	0.341	0.868	0.352
SQ1	0.219	0.218	0.164	0.234	0.234	0.134	0.309	0.872	0.298	0.247
SQ2	0.282	0.249	0.111	0.234	0.325	0.174	0.274	0.712	0.196	0.331
SQ3	0.205	0.175	0.213	0.256	0.218	0.125	0.282	0.900	0.266	0.215
SQ4	0.231	0.245	0.222	0.309	0.249	0.285	0.257	0.747	0.279	0.260
SQ5	0.174	0.169	0.272	0.313	0.251	0.208	0.196	0.709	0.201	0.234
SQ6	0.322	0.301	0.288	0.329	0.343	0.318	0.335	0.839	0.320	0.332
SQ7	0.315	0.323	0.278	0.382	0.368	0.352	0.314	0.764	0.268	0.359
SQ8	0.386	0.356	0.294	0.422	0.453	0.367	0.364	0.833	0.334	0.456
SQ9	0.285	0.305	0.307	0.347	0.351	0.309	0.283	0.846	0.311	0.351
SQ10	0.345	0.346	0.297	0.399	0.405	0.365	0.334	0.874	0.297	0.420
SQ11	0.305	0.290	0.242	0.301	0.346	0.312	0.293	0.781	0.283	0.331
SQ12	0.306	0.269	0.228	0.295	0.316	0.249	0.283	0.833	0.251	0.334
SQ13	0.350	0.278	0.255	0.231	0.350	0.268	0.319	0.744	0.248	0.340
SQ14	0.362	0.293	0.238	0.286	0.351	0.236	0.286	0.770	0.243	0.369
SQ15	0.476	0.457	0.294	0.328	0.511	0.454	0.381	0.763	0.303	0.492
SQ16	0.412	0.418	0.251	0.344	0.394	0.288	0.378	0.813	0.284	0.473
SQ17	0.317	0.278	0.120	0.275	0.361	0.325	0.344	0.793	0.219	0.374
SQ18	0.365	0.296	0.182	0.279	0.389	0.260	0.315	0.786	0.262	0.394
UPT1	0.657	0.587	0.378	0.439	0.636	0.501	0.519	0.443	0.275	0.847
UPT2	0.684	0.653	0.395	0.505	0.685	0.516	0.494	0.456	0.325	0.894
UPT3	0.689	0.641	0.446	0.540	0.685	0.521	0.520	0.463	0.361	0.890

3.2 Structural Model Modelling (SEM)

The study aimed to predict the use of public transportation systems under pandemic conditions through SEM. Because of partial least squares (PLS) calculations, the estimation of the model was carried out and the hypothesis was tested. It was determined that the path coefficients were significant by using the bootstrap method along with 500 sub-samples to test for significance. According to Fig. 2, there is a relationship between the variables that can be seen in the tested model, with numbers encapsulated in the circles representing the explained variance of each variable for the model as a whole. For each variable, Tab. 4 provides information regarding the path coefficients and the explanations of variance based on the path coefficients.

Table 3 Correlation matrix

	Attitude	Intention	Expectation	Satisfaction	Behavior control	PEU	Perceived usefulness	Service quality	Subjective norm	Use of public transport
Attitude	0.866									
Intention	0.510	0.894								
Expectation	0.423	0.408	0.829							
Satisfaction	0.505	0.556	0.574	0.812						
Perceived behavior control	0.518	0.691	0.393	0.488	0.851					
PEU	0.622	0.613	0.378	0.420	0.561	0.818				
PU	0.596	0.643	0.311	0.450	0.599	0.436	0.825			
Service Quality	0.464	0.433	0.349	0.455	0.509	0.414	0.452	0.685		
Subjective norm	0.364	0.467	0.331	0.353	0.403	0.358	0.476	0.397	0.845	
Use of Public Transport	0.572	0.616	0.464	0.565	0.463	0.584	0.582	0.518	0.367	0.877
Note: The diagonal elements in the correlation matrix is the square root of AVE for each construct										

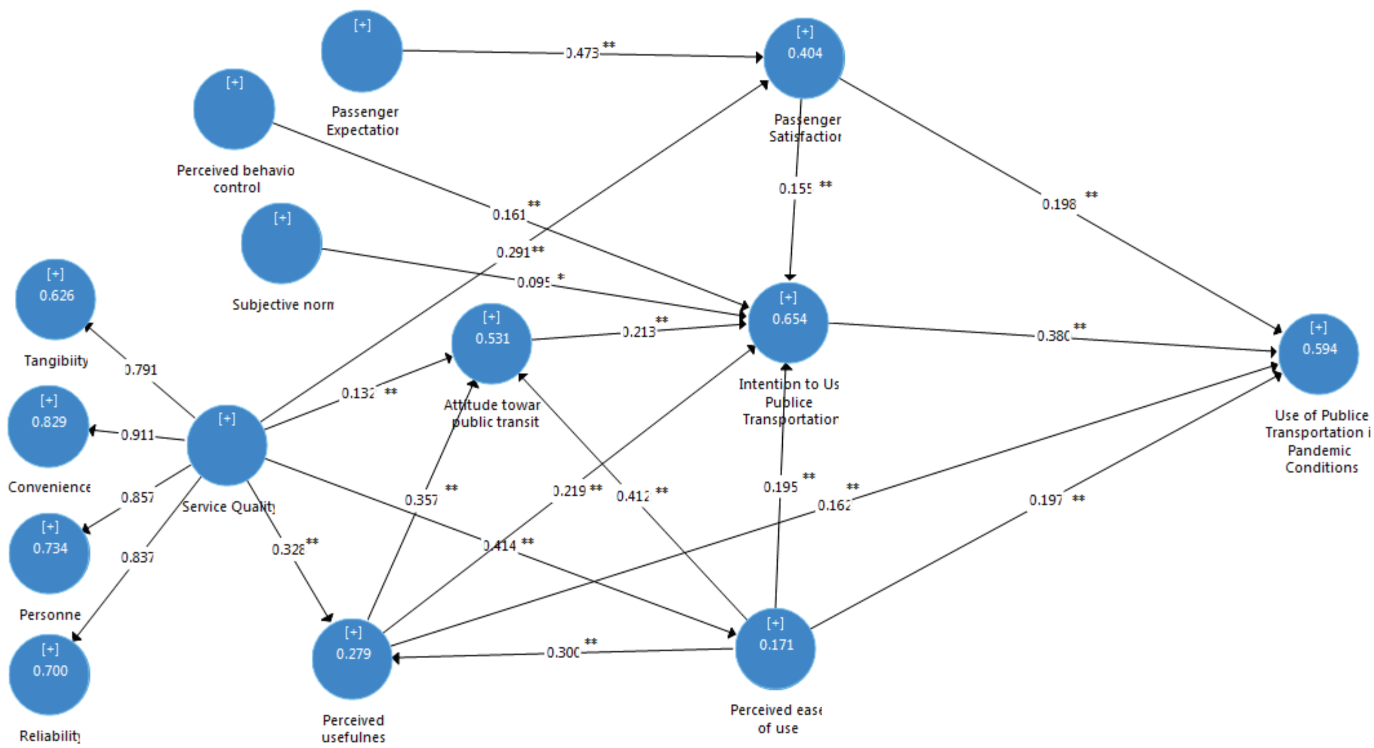


Figure 2 The tested model

Tab. 4 highlights a significant effect of expectations and service quality, indicating that both factors positively and significantly impact satisfaction. Perceptions of behavioral control, subjective norms, attitudes, PU, and PEU are also positively and significantly associated with intention. Additionally, a favorable perception of service quality, PU, and PEU has a positive and significant impact on attitude. Finally, expectations, intention, PU, and PEU significantly influence the effectiveness of the public transportation system. The proposed model accounts for 59% of the variation in public transportation usage during pandemics, 65% of the variance in intention, 53% of the variance in attitude, 40% of the variance in satisfaction, and 17% of the variance in PEU. Tab. 5 presents the indirect coefficients as well.

A mediation analysis is presented in Tab. 5, indicating that PEU and PU have a positive and significant mediating effect on the relationship between service quality and attitude, which is consistent with the results of the mediation analysis. It should be noted that PU has a positive and significant mediating effect on the effect of PEU on attitude in addition to PU. In addition to a positive and significant impact on the intention that service quality can have on attitude, satisfaction, PEU, and PU, these variables also play a positive and significant role in the mediating effect. Lastly, intention plays an important role in the perception of usefulness and PEU of public transportation during pandemic conditions as well as predicting attitude, satisfaction, PU, PEU, subjective norms, and perceived behavioral control. As shown in Tab. 6, the results of hypothesis testing have been reported.

Table 4 Path coefficients and explained variance

Variable	β	*t*-value	*P*-value	Explained variance
On the use of public transport system in pandemic conditions via:				
Intention	0.38**	6.697	0.001	0.594
Satisfaction	0.198**	4.294	0.001	
PU	0.162**	3.065	0.002	
PEU	197**	4.411	0.001	
On intention via:				
Perceived behavioral control	0.161**	3.106	0.002	
Subjective norm	0.095*	2.399	0.017	
Satisfaction	0.155**	3.250	0.001	0.654
PU	0.219**	4.504	0.001	
PEU	0.195**	3.726	0.001	
Attitude	0.213**	2.964	0.003	
On attitude via:				
PU	0.357**	6.893	0.001	0.531
PEU	0.412**	8.397	0.001	
Service quality	0.132**	2.688	0.007	
On satisfaction vie:				
Expectation	0.473**	9.717	0.001	0.404
Service quality	0.291**	5.471	0.001	
On PU vie:				
PEU	0.30**	5.472	0.001	0.171
Service quality	0.328**	5.823	0.001	

$^{**}p < 0.01$

As a way to assess the validity and quality of a PLS model, the goodness-of-fit index is often used as an indicator of the model's validity. As opposed to the GOF index, which determines whether the tested model is able to successfully predict the endogenous variables, the GOF index measures the overall predictive ability of the model. GOF value of 0.61 was achieved for the tested model in the current study, which indicates that the model is well fitted for the data. For a model to be considered adequate and acceptable for quality, the GOF value has to be higher than 0.36.

4 DISCUSSION

The objective of this study was to investigate the factors influencing the usage of public transportation amidst the COVID-19 pandemic, utilizing SEM and TAM as analytical tools. Our analysis revealed that the proposed model was a suitable fit for the data, accounting for a considerable proportion of the variance in public transportation use during the outbreak (59%). Moreover, it explained a significant portion of the variation in the intention to use public transportation (65%), attitude towards it (53%), satisfaction levels (40%), and PEU (17%).

Our study has revealed that perceived behavioral control, subjective norm, and attitude significantly and positively impact intention to use public transportation. These findings align with previous research conducted by Bamberg et al. [15], Gardner and Abraham [18], Erickson and Forward [19], Mahmoudabadi [55] and Zailani et al. [20]. Moreover, these results are consistent with the TPB, which suggests that behavior is directly influenced by behavioral intention, shaped by attitudes, subjective norms, and perceived behavioral control.

Table 5 Indirect coefficients

Indirect paths	Indirect effects	*t*-value	*P*-values
Service quality → PEU → Attitude	0.171	5.319	0.000
PEU → PU → Attitude	0.107	4.193	0.000
Service Quality → PU → Attitude	0.117	4.542	0.000
PEU → Attitude → Intention	0.088	2.782	0.006
PU → Attitude → Intention	0.076	2.794	0.005
Service quality → Attitude → Intention	0.028	1.999	0.046
Expectation → Satisfaction → Intention	0.073	2.852	0.005
Service quality → Satisfaction → Intention	0.045	2.829	0.005
Service quality → PEU → Intention	0.081	3.344	0.001
PEU → PU → Intention	0.066	3.544	0.000
Service quality → PU → Intention	0.072	3.604	0.000
Service quality → PEU → PU	0.124	4.360	0.000
Attitude → Intention → Use of public transport	0.081	2.514	0.012
Satisfaction → Intention → Use of public transport	0.059	2.962	0.003
Perceived behavior control → Intention → Use of public transport	0.061	2.689	0.007
PEU → Intention → Use of public transport	0.074	3.506	0.000
PU → Intention → Use of public transport	0.084	4.157	0.000
Subjective norm → Intention → Use of public transport	0.036	2.253	0.025
Expectation → Satisfaction → Use of public transport	0.094	4.267	0.000
Service quality → Satisfaction → Use of public transport	0.058	3.392	0.001
Service quality → PEU → Use of public transport	0.082	3.509	0.000
PEU → PU → Use of public transport	0.049	2.582	0.010
Service quality → PU → Use of public transportation	0.053	2.772	0.006

To explain the impact of perceived behavioral control on intention, it can be argued that people's intention to use public transportation will increase during a pandemic if they believe that they have the ability to use public transportation as intended. This includes perceiving themselves as capable of using public transportation, having access to adequate facilities and time to utilize this mode of transportation, possessing a strong inclination towards using public transportation, having numerous opportunities to do so, and feeling in control of their use of public transportation. If individuals have strong beliefs regarding the factors that facilitate a particular behavior, they are more likely to perceive themselves as having greater control over that behavior. Conversely, if their control beliefs are weak, they may perceive themselves as having less control and be less likely to engage in the behavior. This perception can be influenced by various factors such as past experiences, predictions of future events, and attitudes shaped by societal norms within their environment.

Table 6 Results of hypothesis testing

Hypothesis	Result
H1: Effect of perceived behavioral control on intention	Confirmed
H2: Effect of subjective norm on intention	Confirmed
H3: Effect of attitude on intention	Confirmed
H4: Effect of PEU on PU	Confirmed
H5: Effect of PEU on attitude	Confirmed
H6: Effect of PEU on intention	Confirmed
H7: Effect of PEU on the use of public transportation in pandemic conditions	Confirmed
H8: Effect of PU on attitude	Confirmed
H9: Effect of PU on intention	Confirmed
H10: Effect of PU on the use of public transportation in pandemic conditions	Confirmed
H11: Effect of intention on the use of public transportation in pandemic conditions	Confirmed
H12: Effect of satisfaction on intention	Confirmed
H13: Effect of satisfaction on the use of public transportation in pandemic conditions	Confirmed
H14: Effect of expectations on satisfaction	Confirmed
H15: Effect of perceived service quality on satisfaction	Confirmed
H16: Effect of perceived service quality on attitude	Confirmed
H17: Effect of perceived service quality on PEU	Confirmed
H18: Effect of perceived service quality on PU	Confirmed

To clarify the impact of subjective norms on intention, it can be argued that if societal expectations dictate that individuals use public transportation during a pandemic, there is a greater likelihood that they will engage in this behavior. Consequently, if people hold the belief that those around them possess positive attitudes towards using public transportation in pandemic conditions, their intention to utilize this mode of transportation will likely increase. As per Fishbein and Ajzen's theory [9], subjective norms reflect the social pressures that individuals perceive surrounding a specific behavior. Therefore, when specific reference groups or individuals hold expectations regarding public transportation use, these perceived expectations can motivate people to utilize this mode of transportation.

In order to explain the impact of attitude on intention, it can be argued that individuals will be more likely to utilize the public transportation system during a pandemic if they have a positive predisposition towards using this mode of transportation and typically rely on it while traveling. Additionally, factors such as short travel times to public transportation stations, ease of transitioning between different stations, acceptable transport schedules, clear timetables, and overall satisfaction with the environment of public transportation during travel may also contribute to increased inclination towards using this mode of transportation.

According to this study, one of the main findings was that people's perceptions of the usefulness and ease with which they can use public transportation during pandemic conditions were significantly correlated with their attitude, intention, and actual use of public transport during such conditions. It should be noted that this finding is consistent with earlier studies conducted by Chen and Chao [17], Muenrit et al. [26], and Ahn and Park [24], which also highlighted the importance of PU and PEU in determining the behavior of people. It can be concluded from these results that they are consistent with the findings of the TAM, which indicates that people who perceive public transport as convenient and easy to use during a pandemic are more likely to take advantage of it regularly.

The findings indicate that passengers' expectations influence significantly their satisfaction levels. There is good agreement between the results of this study and the previous studies conducted by Heidari et al. [35] and Yuan et al. [31]. One explanation for this finding is that when passengers have high expectations of the public transportation system, and these expectations are clearly communicated, it motivates the system to deliver better services. Consequently, when passengers receive the expected services, there is a smaller gap between their expectations and the actual service provided, leading to higher levels of passenger satisfaction during pandemic conditions.

One of the key findings of this study is that satisfaction influences significantly the intention to use public transportation during pandemic conditions. The results of this study are similar to those obtained in previous studies conducted by Shen et al. [36], Farooq et al. [34], and Yuan et al. [31]. To provide an explanation for this finding, it can be argued that if passengers are satisfied with their experiences while using the public transportation system during pandemic conditions, such as having a pleasant feeling and responsive services, they are more likely to use it. Satisfaction is typically associated with achieving desired goals and demands, which leads to a comfortable situation. When passengers achieve their desired goals and demands and feel satisfied, public transportation is more likely to be perceived positively by them and therefore, will have a higher tendency to use these services during pandemic conditions.

In addition, the study revealed that service quality significantly influenced passenger satisfaction, attitude, as well as perceived ease and usefulness of public transportation. In addition to these results, previous studies conducted by Wang et al. [46], Nguyen-Phuoc et al. [47], Yuan [31], and AL-Nawafleh et al. [48] have also found similar results. There are a few possible explanations for this finding, one of which is that passenger satisfaction, positive attitude toward using public transportation, PEU, and usefulness of public transportation will increase when the goal of public transportation companies is to provide proper services to passengers and operate on a service-oriented basis. They establish better interactions with passengers, try to solve passengers' problems, and help them while providing better services in pandemic conditions. Besides, they should be polite and respectful towards passengers. Their services should be available, and they should provide clean public transportation space. A sincere and friendly conversation with passengers can give them the right guidance. Therefore, service quality is a determining factor in the use of public transportation in pandemic conditions, which clearly influences the response of passengers.

5 MANAGERIAL IMPLICATIONS

When considering the variables of the TPB, it is suggested that the public transport service system provides high-quality services during pandemic conditions while also meeting the expectations of passengers and complying with health guidelines. This approach can help cultivate a positive attitude amongst passengers towards the use of public transport services during pandemic conditions. Additionally, friends and acquaintances can influence behavior as one of the significant factors, and therefore promoting the use of public transportation services in pandemic conditions can be encouraged through peer recommendations.

When considering the variables of the TAM, it is recommended that public transportation service systems prioritize easy access and usage for passengers during pandemic conditions. By promoting the role of public transportation in improving performance and providing quick access to desired destinations, passengers may be more willing to use these services.

Regarding passenger expectations and satisfaction when using public transportation service systems during a pandemic, it is suggested that service providers influence passenger behavior by offering fast, round-the-clock, and cost-effective services. Meeting customer expectations, complying with health guidelines, and offering health facilities can also encourage customers to use public transportation services during a pandemic.

Furthermore, service quality can play a significant role in determining the intention to use public transportation systems during a pandemic. It is recommended that service providers induce commitment from passengers by observing health guidelines and paying attention to their needs. Responding quickly to passenger problems, guaranteeing available and clean services, and providing a safe environment are also crucial factors in encouraging passengers to use public transportation services during a pandemic.

6 CONCLUSION

The study has demonstrated that both the TPB and TAM theories are robust predictors of public transportation usage during pandemic situations. Additionally, PEU and PU play a crucial role in mediating the impact of service quality on attitudes towards public transportation. Moreover, PU is instrumental in mediating the impact of PEU on positive attitudes towards public transportation. The effect of service quality on the intention to use public transportation is further mediated by attitudes towards public transportation, passenger satisfaction, PEU, and PU. Furthermore, public transportation intention serves as a critical mediator in determining how attitudes towards public transportation, subjective norms, perceived behavioral control, passenger satisfaction, PEU, and PU affect actual public transportation usage during pandemics. Therefore, to increase the intention to use public transportation systems during pandemics, it is vital to consider the variables of TAM, TPB, and service quality.

7 LIMITATIONS

It is important to take note of the fact that this research was limited to a sample of Tehran citizens, and therefore the generalization of findings may be restricted. Additionally, the results are based on self-report data, which may have some limitations. To gain a more comprehensive understanding of the effective factors on the intention to use public transportation systems during pandemic conditions, future studies should consider using qualitative, mixed research methods, and machine learning and artificial intelligence [68-73] as well.

8 REFERENCES

[1] Aknin, L. B., De Neve, J. E., Dunn, E. W., Fancourt, D. E., Goldberg, E., Helliwell, J. F., Jones, S. P., Karam, E., Layard, R., Lyubomirsky, S., Rzepa, A., Saxena, S., et al. (2021). The Lancet's COVID-19 Commission Mental Health Task Force. *Acta Neuropsychologica, 19*(3), 301-305. https://doi.org/10.5604/01.3001.0014.9953

[2] WHO. (2021). WHO Coronavirus (COVID-19) Dashboard, World Health Organization. https://covid19.who.int/

[3] Kamali, A., Ebadi, M., Vahidnezhad, H., Mohammadi, A., & Dalvandi, M. (2021). Clinical characteristics of coronavirus disease 2019 in a cohort of iranian patients: Olfactory dysfunction as a frequent clinical presentation. *Indian Journal of Forensic Medicine & Toxicology, 15*(2), 1768-1770. https://doi.org/10.37506/ijfmt.v15i2.14594

[4] Zhen, J., Chan, C., Schoonees, A., Apatu, E., Thabane, L., & Young, T. (2020). Transmission of respiratory viruses when using public ground transport: A rapid review to inform public health recommendations during the COVID-19 pandemic. *South African Medical Journal, 110*(6), 478-483.

[5] Murdoch, L. C., Germanovich, L. N., Slack, W. W., Carbajales-Dale, M., Knight, D., Moak, R., Laffaille, C., DeWolf, S., & Roudini, S. (2023). Shallow Geologic Storage of Carbon to Remove Atmospheric CO_2 and Reduce Flood Risk. *Environmental Science & Technology, 57*(23), 8536. https://doi.org/10.1021/acs.est.3c00600

[6] Gkiotsalitis, K., & Cats, O. (2021). Public transport planning adaption under the COVID-19 pandemic crisis: literature review of research needs and directions. *Transport Reviews, 41*(3), 374-392. https://doi.org/10.1080/01441647.2020.1857886

[7] Ajzen, I. (1985). From intentions to actions: A theory of planned behavior (pp. 11-39). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-69746-3_2

[8] Niu, C., Gutierrez, G., Sadeghi, S., Cossette, L., Portugal, M., Zeng, S., & Zhang, P. (2023). What Changes in Adult Online Learners' Motivation?: A Systematic Meta-Review of the Scholarly Landscape of Research on Adult Online Learning

Motivation. *Motivation and Momentum in Adult Online Education*, 1-17. https://doi.org/10.4018/978-1-6684-7712-0.ch001

[9] Ajzen, I. & Fishbein, M. (1980). Understanding Attitude and Predicting Social Behavior.

[10] Nazari-Shirkouhi, S., Badizadeh, A., Dashtpeyma, M., & Ghodsi, R. (2023). A model to improve user acceptance of e-services in healthcare systems based on technology acceptance model: an empirical study. Journal of Ambient Intelligence and Humanized Computing, 1-17. https://doi.org/10.1007/s12652-023-04601-0

[11] Rezvani, S., Heidari, S., Roustapisheh, N., & Dokhanian, S. (2022). The effectiveness of system quality, habit, and effort expectation on library application use intention: the mediating role of perceived usefulness, perceived ease of use, and user satisfaction. *International Journal of Business Information Systems*, 1-18. https://doi.org/10.1504/ijbis.2022.10049515

[12] Cuocci, S., Fattahi Marnani, P., Khan, I., & Roberts, S. (2023). A Meta-Synthesis of Technology-Supported Peer Feedback in ESL/EFL Writing Classes Research: A Replication of Chen's Study. *Languages, 8*(2), 114. https://doi.org/10.3390/languages8020114

[13] Fattahi Marnani, P. and Cuocci, S., 2022. Foreign Language Anxiety: A Review on Theories, Causes, Consequences and Implications for Educators. *Journal of English Learner Education, 14*(2), p. 2. https://stars.library.ucf.edu/jele/vol14/iss2/2

[14] La Barbera, F. & Ajzen, I. (2021). Moderating role of perceived behavioral control in the theory of planned behavior: A preregistered study. *Journal of Theoretical Social Psychology, 5*(1), 35-45. https://doi.org/10.1002/jts5.83

[15] Bamberg, S., Ajzen, I., & Schmidt, P. (2003a). Choice of travel mode in the theory of planned behavior: The roles of past behavior, habit, and reasoned action. *Basic and applied social psychology, 25*(3), 175-187. https://doi.org/10.1207/S15324834BASP2503_01

[16] Bamberg, S., Rölle, D., & Weber, C. (2003b). Does habitual car use not lead to more resistance to change of travel mode? *Transportation, 30*, 97-108. https://doi.org/10.1023/A:1021282523910

[17] Chen, C. F. & Chao, W. H. (2011). Habitual or reasoned? Using the theory of planned behavior, technology acceptance model, and habit to examine switching intentions toward public transit. *Transportation research part F: traffic psychology and behaviour, 14*(2), 128-137. https://doi.org/10.1016/j.trf.2010.11.006

[18] Gardner, B. & Abraham, C. (2008). Psychological correlates of car use: A meta-analysis. *Transportation Research Part F: Traffic Psychology and Behaviour, 11*(4), 300-311. https://doi.org/10.1016/j.trf.2008.01.004

[19] Zahmatkesh, S., Karimian, M., Pourhanasa, R., Ghodrati, I., Hajiaghaei-Keshteli, M., & Ismail, M. A. (2023). Wastewater treatment with algal based membrane bioreactor for the future: Removing emerging containments. *Chemosphere*, 139134. https://doi.org/10.1016/j.chemosphere.2023.139134

[20] Zailani, S., Iranmanesh, M., Masron, T. A., & Chan, T. H. (2016). Is the intention to use public transport for different travel purposes determined by different factors? *Transportation research part D: transport and environment, 49*, 18-24. https://doi.org/10.1016/j.trd.2016.08.038

[21] Murdoch, L. C., DeWolf, S., Germanovich, L. N., Moysey, S., Hanna, A., Roudini, S., & Moak, R. (2023). Using the Shallow Strain Tensor to Characterize Deep Geologic Reservoirs. *Water Resources Research, 59*(2), e2022WR032920. https://doi.org/10.1029/2022WR032920

[22] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340. https://doi.org/10.2307/249008

[23] Hu, P. J., Chau, P. Y., Sheng, O. R. L., & Tam, K. Y. (1999). Examining the technology acceptance model using physician acceptance of telemedicine technology. *Journal of management information systems, 16*(2), 91-112. https://doi.org/10.1080/07421222.1999.11518247

[24] Ahn, H. & Park, E. (2022). For sustainable development in the transportation sector: Determinants of acceptance of sustainable transportation using the innovation diffusion theory and technology acceptance model. *Sustainable Development*, *30*(5), 1169-1183. https://doi.org/10.1002/sd.2309

[25] Asl, B., Gill, C., Umstead, K., Mahtani, R., & Tully, K. (2022). A Collaborative Approach to Digital Storytelling in Healthcare Settings. In *DS 117: Proceedings of the 24th International Conference on Engineering and Product Design Education (E&PDE 2022)*, London South Bank University in London, UK. 8-9 September 2022.

[26] Muenrit, K., Satiennam, W., Satiennam, T., & Jaensirisak, S. (2017). The Effects of Psychological Factors On Travelers' Mode Choice Intention: A Case Study Of Light Rail Transit System (LRT) in Khon Kaen, Thailand. *Journal of the Eastern Asia Society for Transportation Studies, 12*, 1192-1200.

[27] Taylor, S. & Todd, P. A. (1995). Assessing IT usage: The role of prior experience. *MIS Quarterly, 19*(4), 561-570. https://doi.org/10.2307/249633

[28] Kotler, P., Keller, K. L., Manceau, D., & Hémonnet-Goujot, A. (2015). *Marketing management, 14*.

[29] Heidari, S., Zarei, M., Daneshfar, A., & Dokhanian, S. (2023). Increasing Sales through Social Media Marketing: The Role of Customer Brand Attachment, Brand Trust, and Brand Equity. *Marketing and Management of Innovations, 14*(1), 224-234. https://doi.org/10.21272/mmi.2023.1-19

[30] Khorsandi, H. & Khorsandi, R. (2022). Ranking the effective factors in creative marketing in Iran Novin insurance. *Journal of Positive School Psychology, 6*(5), 10009-10020.

[31] Nazari-Shirkouhi, S., Keramati, A., & Rezaie, K. (2015). Investigating the effects of customer relationship management and supplier relationship management on new product development. *Tehnički vjesnik, 22*(1), 191-200. https://doi.org/10.17559/TV-20140623130536

[32] Chou, J. S., Kim, C., Kuo, Y. C., & Ou, N. C. (2011). Deploying effective service strategy in the operations stage of high-speed rail. *Transportation Research Part E: Logistics and Transportation Review, 47*(4), 507-519. https://doi.org/10.1016/j.tre.2010.12.004

[33] Koklic, M. K., Kukar-Kinney, M., & Vegelj, S. (2017). An investigation of customer satisfaction with low-cost and full-service airline companies. *Journal of Business Research, 80*, 188-196. https://doi.org/10.1016/j.jbusres.2017.05.015

[34] Farooq, M. S., Salam, M., Fayolle, A., Jaafar, N., & Ayupp, K. (2018). Impact of service quality on customer satisfaction in Malaysia airlines: A PLS-SEM approach. *Journal of Air Transport Management, 67*, 169-180. https://doi.org/10.1016/j.jairtraman.2017.12.008

[35] Sasani, F., Mousa, R., Karkehabadi, A., Dehbashi, S., & Mohammadi, A. (2023). TM-vector: A Novel Forecasting Approach for Market stock movement with a Rich Representation of Twitter and Market data. *arXiv preprint arXiv*: 2304.02094.

[36] Shen, W., Xiao, W., & Wang, X. (2016). Passenger satisfaction evaluation model for urban rail transit: A structural equation modeling based on partial least squares. *Transport Policy, 46*, 20-31. https://doi.org/10.1016/j.tranpol.2015.10.006

[37] Khorsandi, H., & Bayat, M. (2022). Prioritizing operational strategies of saman bank. *International Journal of Health Sciences, 6*(S7), 1442-1453. https://doi.org/10.53730/ijhs.v6nS7.11548

[38] Ibrahim, A. N. H., Borhan, M. N., & Ismail, A. (2020). Rail-based public transport service quality and user satisfaction–a literature review. *Promet-Traffic&Transportation, 32*(3), 423-435. https://doi.org/10.7307/ptt.v32i3.3270

[39] Dam, S. M. & Dam, T. C. (2021). Relationships between service quality, brand image, customer satisfaction, and customer loyalty. *The Journal of Asian Finance, Economics and Business, 8*(3), 585-593.

[40] Ladhari, R. (2009). A review of twenty years of SERVQUAL research. *International journal of quality and service sciences, 1*(2), 172-198. https://doi.org/10.1108/17566690910971445

[41] Jen, W., Tu, R., & Lu, T. (2011). Managing passenger behavioral intention: an integrated framework for service quality, satisfaction, perceived value, and switching barriers. *Transportation, 38*, 321-342. https://doi.org/10.1007/s11116-010-9306-9

[42] Emami, B. D., Song, Y., & Khani, A. (2022). Prioritizing Bus Routes for Electrification: GIS-Based Multi-Criteria Analysis Considering Operational, Environmental, and Social Benefits and Costs. *Transportation Research Record, 2676*(8), 10-23. https://doi.org/10.1177/03611981221082565

[43] Pérez, M., Carlos Gázquez Abad, J., María Marín Carrillo, G., & Sánchez Fernández, R. (2007). Effects of service quality dimensions on behavioural purchase intentions: A study in public-sector transport. *Managing Service Quality: An International Journal, 17*(2), 134-151. https://doi.org/10.1108/09604520710735164

[44] Prasad, M. D. & Shekhar, B. R. (2010). Development of Railqual: A Service Quality Scale for Measuring Indian Railway Passenger Services. *Management Science and Engineering, 4*(3), 87-94. https:// doi.org/10.3968/j.mse.1913035X20100403.009

[45] Chou, P. F., Lu, C. S., & Chang, Y. H. (2014). Effects of service quality and customer satisfaction on customer loyalty in high-speed rail services in Taiwan. *Transportmetrica A: transport science, 10*(10), 917-945. https://doi.org/10.1080/23249935.2014.915247

[46] Wang, T. L., Tran, P. T. K., & Tran, V. T. (2017). Destination perceived quality, tourist satisfaction and word-of-mouth. *Tourism Review, 72*(4), 392-410. https://doi.org/10.1108/TR-06-2017-0103

[47] Nguyen-Phuoc, D. Q., Tran, A. T. P., Van Nguyen, T., Le, P. T., & Su, D. N. (2021). Investigating the complexity of perceived service quality and perceived safety and security in building loyalty among bus passengers in Vietnam–a PLS-SEM approach. *Transport Policy, 101*, 162-173. https://doi.org/10.1016/j.tranpol.2020.12.010

[48] AL-Nawafleh, E. A., ALSheikh, G. A. A., Abdulllah, A. A., & bin A. Tambi, A. M. (2019). Review of the impact of service quality and subjective norms in TAM among telecommunication customers in Jordan. *International Journal of Ethics and Systems, 35*(1), 148-158. https://doi.org/10.1108/IJOES-07-2018-0101

[49] Zaman, S., Khawaja, K. F., & Waqar, S. (2013). Impact of service Quality Perceived Risk, Perceived Privacy and Security and Perceived Trust on technology acceptance model: An empirical study based on online banking sector of Pakistan. *In the 2nd IEEE National Conference on Information Assurance (NCIA)*, 89-94. https://doi.org/10.1109/NCIA.2013.6725330

[50] Chen, H. S., Liang, C. H., Liao, S. Y., & Kuo, H. Y. (2020). Consumer attitudes and purchase intentions toward food delivery platform services. *Sustainability*, *12*(23), 10177. https://doi.org/10.3390/su122310177

[51] Noruzy, A., Dalfard, V. M., Azhdari, B., Nazari-Shirkouhi, S., & Rezazadeh, A. (2013). Relations between transformational leadership, organizational learning, knowledge management, organizational innovation, and organizational performance: an empirical investigation of manufacturing firms. *The International Journal of Advanced Manufacturing Technology, 64*, 1073-1085. https://doi.org/10.1007/s00170-012-4038-y

[52] Abbas Mofrad, H., Khalatbari, J., Malihi Zakerini, S., Mohammadi Shirmahalleh, F., & Shafti, V. (2021). Analysis of Structural Equations in the Relationship of Marital Conflicts and Affective Security with Perceived Stress and Pregnancy Worries and Biological Indexes with the Mediation of Psychological Wellbeing in Pregnant Women. *Women Studies, 12*(35), 97-127.

[53] Alipour, N., Nazari-Shirkouhi, S., Sangari, M. S., & Vandchali, H. R. (2022). Lean, agile, resilient, and green human resource management: the impact on organizational innovation and organizational performance. *Environmental Science and Pollution Research, 29*(55), 82812-82826. https://doi.org/10.1007/s11356-022-21576-1

[54] Abbasmofrad, H., Khalatbari, J., Zakerini, S. M., Mohammadi Shir Mahalla, F., & Shafti, V. (2022). Analysis of Structural Equations in Relation to Marital Conflicts and Emotional Security with Perceived Stress and Pregnancy Worries Mediated by Psychological Well-Being in Pregnant Women. *MEJDS, 12*, 140-149.

[55] Mahmoudabadi, N. S. (2020). The Study of Cable Behavior with Two Spring-Dampers and One Viscous Damper. *World, 9*(S1), 128-136.

[56] Dehghani, F. & Karbasian, H. (2023). House Price Prediction in US Metropolitan Areas. Available at SSRN 4494273. https://doi.org/10.2139/ssrn.4494273

[57] Moshtaghi Largani, S. & Lee, S. (2023). Efficient Sampling for Big Provenance. *In Companion Proceedings of the ACM Web Conference 2023*, 1508-1511. https://doi.org/10.1145/3543873.3587556

[58] Laube, A., Sasani, F. (2020). Analysis of Risk Measurement in Financial Companies. *Journal of Humanities Insights 4*(01), 27-28.

[59] Jafarzadeh, E., Kabiri-Samani, A., Boroomand, B., & Bohluly, A. (2023). Analytical modeling of flexible circular submerged mound motion in gravity waves. *Journal of Ocean Engineering and Marine Energy, 9*(1), 181-190. https://doi.org/10.1007/s40722-022-00248-9

[60] Ibrahim, H. & Sasani, F. (2021). Study of the Interactions of Human Resource Performance and Electronic Human Resource Management in Small *Companies. Journal of Humanities Insights, 5*(01), 24-32.

[61] Jafarzadeh, E., Bohluly, A., Kabiri-Samani, A., & Mansourzadeh, S. (2023). A study on the performance of circular and rectangular submerged breakwaters using nun-uniform FGVT method. *Coastal Engineering Journal, 65*(2), 234-255. https://doi.org/10.1080/21664250.2023.2170688

[62] Asl, S. B., Naeini, H. S., Ensaniat, L. S., Khorshidian, R., & Alipour, S. (2014). Injury Prevention among Construction Workers: A Case Study on Iranian Steel Bar Bending Workers. *International Journal of Industrial and Manufacturing Engineering, 8*(8), 467-470.

[63] Anbari, M., Arıkan Öztürk, E. B. R. U., & Ateş, H. (2020). Evaluation of sustainable transport strategies for Tehran with thetheir urbanization rate criterion based on the fuzzy ahp

method. *Journal of Xi'xxan University of Architecture Technology, 12*(7), 867-881.

[64] Anbari, M., Arıkan Öztürk, E. B. R. U., & Ateş, H. A. K. A. N. (2020). Urban Design and Upgrading Traffic and Urban Street Safety from the Perspective of Urban Users (Case Study: Tunalı Hilmi-Ankara Residential Commercial Street). *The Journal of International Social Research, 13*(69), 506-515. https://doi.org/10.17719/jisr.2020.3974

[65] Sadeghi, S., Marjani, T., Hassani, A., & Moreno, J. (2022). Development of Optimal Stock Portfolio Selection Model in the Tehran Stock Exchange by Employing Markowitz Mean-Semivariance Model. *Journal of Finance Issues, 20*(1), 47-71. https://doi.org/10.58886/jfi.v20i1.3061

[66] Dehghan, S. & Naghibi Iravani, S. (2022). Comparison of seismic behavior factors for reinforced concrete (RC) special moment resisting frames (SMRFs) in Iran in low-, mid-, and high-rise buildings based on Iranian seismic standard 2800 and ASCE, *Journal of Economics and Administrative Sciences, 5*(S1), 744-750.

[67] Iravani, S. N. & Dehghan, S. An Investigation to the Seismic Performance of Base Isolator-Equipped Moment Frame Steel Structures. *Economics and Administrative Sciences, 5*(S1), 751-759.

[68] Razavi, S. E., Moradi, M. A., Shamaghdari, S., & Menhaj, M. B. (2022). Adaptive optimal control of unknown discrete-time linear systems with guaranteed prescribed degree of stability using reinforcement learning. *International Journal of Dynamics and Control, 10*(3), 870-878. https://doi.org/10.1007/s40435-021-00836-x

[69] Aghakhani, S., Larijani, A., Sadeghi, F., Martín, D., & Shahrakht, A. A. (2023). A Novel Hybrid Artificial Bee Colony-Based Deep Convolutional Neural Network to Improve the Detection Performance of Backscatter Communication Systems. *Electronics*, *12*, 2263. https://doi.org/10.3390/electronics12102263

[70] Moradi, M., Weng, Y., & Lai, Y. C. (2022). Defending Smart Electrical Power Grids against Cyberattacks with Deep Q-Learning. *P R X Energy, 1*, 033005. https://doi.org/10.1103/PRXEnergy.1.033005

[71] Sadeghi, F., Larijani, A., Rostami, O., Martín, D., & Hajirahimi, P. (2023). A Novel Multi-Objective Binary Chimp Optimization Algorithm for Optimal Feature Selection: Application of Deep-Learning-Based Approaches for SAR Image Classification. *Sensors*, *23*, 1180. https://doi.org/10.3390/s23031180

[72] Zhai, Z. M., Moradi, M., Kong, L. W., & Lai, Y. C. (2023). Detecting Weak Physical Signal from Noise: A Machine-Learning Approach with Applications to Magnetic-Anomaly-Guided Navigation. *Physical Review Applied, 19*(3), 034030. https://doi.org/10.1103/PhysRevApplied.19.034030

[73] Amiri, M. K., Zaferani, S. P. G., Emami, M. R. S., Zahmatkesh, S., Pourhanasa, R., Namaghi, S. S., Klemeš, J. J., Bokhari, A., & Hajiaghaei-Keshteli, M. (2023). Multi-objective optimization of thermophysical properties GO powders-DW/EG Nf by RSM, NSGA-II, ANN, MLP and ML. Energy, 128176. https://doi.org/10.1016/j.energy.2023.128176

Authors' contacts:

Soheila Saeidi
Department of Civil, Environmental and Geomatics Engineering,
College of Engineering and Computer Science, Florida Atlantic University,
777 Glades Road, Boca Raton, FL 33431, USA

Somayeh Nazari Enjedani
(Corresponding Author)
Department of Civil Engineering, Boise State University,
1910 University Dr., Boise, ID 83725, USA
snazarienjedani@gmail.com

Elmira Alvandi Behineh
Department of Business Administration, University of the Cumberlands,
6191 College Station Dr, Williamsburg, KY 40769, USA

Kian Tehranian
Department of Economics, University of California - Los Angeles (UCLA),
315 Portola Plz Ste 8283, Los Angeles, CA 90095, USA

Seyedalireza Jazayerifar
Department of Marketing, Walton College of Business, University of Arkansas,
220 N McIlroy Ave Ste 302, Fayetteville, AR 72701, USA

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20230626220812

Original scientific paper
Received: 2023-06-26, Accepted: 2023-09-12

Artificial Neural Network Models for Solution Concentration Measurement during Cooling Crystallization of Ceritinib

Ivan Vrban*, Damir Šahnić, Nenad Bolf

Abstract: The development of a quantitative in-line UV spectroscopic method for monitoring of solute concentration during the crystallization process of the active pharmaceutical ingredient (API), ceritinib is described. The method is based on artificial neural networks (ANN). A seeded cooling crystallization process of ceritinib from tetrahydrofuran was studied as a model system. The model was constructed from collected ATR-UV spectra and temperature records within the metastable zone. The collected spectra were preprocessed with the first derivative using the Savitzky-Golay filter. ANN models with different architectures were created and the optimal architecture was chosen based on the root mean square error of prediction (RMSEP) criterion. In addition, ANN models were compared with the models obtained by the linear partial least squares regression (PLSR). Due to the nonlinear relationship in the data set, ANN models predict the solution concentration with higher accuracy compared to linear models. The developed models were successfully used in real-time solution concentration measurement during ceritinib crystallization along with a supersaturation control module developed in-house.

Keywords: artificial neural network; ATR-UV/Vis; ceritinib; concentration real-time measurement; cooling crystallization; supersaturation control

1 INTRODUCTION

Crystallization is a common key unit operation in the pharmaceutical industry used to separate and purify intermediate compounds and active pharmaceutical ingredients (APIs) [1]. It is also a crucial step in obtaining the desired polymorphic form and particle size distribution, parameters that are usually critical in later pharmaceutical development [2, 3]. To understand, develop, and control any given crystallization process, accurate and timely measurement of the solute concentration is a useful tool since it provides information about the supersaturation level at any given time during the crystallization process itself [4]. Using this information, kinetic modeling of crystallization systems [5] is possible, as well as the definition of a control strategy which relies on supersaturation control [6]. Since supersaturation is the driving force for crystallization, its control and the rate at which it is generated during the crystallization process is one of the key parameters affecting the outcome in terms of particle size distribution, polymorphic form, chemical purity and residual solvents [7].

Solute concentration measurement during a crystallization process is a multi-step complicated procedure that involves operations such as suspension sampling and filtration at elevated temperatures, followed by off-line solute concentration determination by techniques such as HPLC. These steps all contribute to the cumulative error inherent in these procedures [8]. On the other hand, developing a method that relies on process analytical technology tools (PAT) can enable in situ determination of the solute concentration during a crystallization process with high accuracy and without physical sampling [9]. To this end, various applications of chemometric models based on spectroscopic data and process temperature records have been developed [10, 11]. Spectroscopic techniques frequently used in the development of such models include attenuated total reflectance Fourier transform infrared spectroscopy (ATR-FTIR) [4, 12], Raman spectroscopy [4, 13, 14] and attenuated total reflectance ultraviolet/visible spectroscopy (ATR-UV/Vis) [15, 6, 16]. Probes functioning on the ATR principle can only record spectra of the liquid phase without detecting the signal of the solid phase which makes them suitable for solute concentration measurement in a suspension during a crystallization process. This is due to their low penetration depth (~2-3 μm) into the measurement field enabled by the probe design itself [17].

In order to develop models which are able to accurately calculate the solution concentration during a crystallization process from the recorded process temperature and spectrum, it is necessary to perform a calibration design over the entire concentration and temperature range of any given crystallization system. During the calibration experiments, UV/Vis spectra and process temperatures within the metastable zone should be collected. Models developed with such data are able to predict solution concentrations with higher accuracy because crystallization occurs in this part of the Ostwald-Miers binary diagram and measurements are not extrapolated outside the range of the developed models [12].

Regularly used chemometric methods for model building are common linear methods like partial least square regression (PLSR) and principal component regression (PCR) [8, 18, 19]. However, these methods are mainly suitable for model building on data sets where the relationship between the input and output data is linear [20]. If the algorithm does not adequately account for the nonlinearity of the data, the resulting model will have lower predictive accuracy. For data sets carrying non-linear relationships, nonlinear chemometric methods like ANN are able to correlate the relationship between input and output with higher predictive accuracy [21].

ANNs have attracted significant research interest due to their potential as powerful tool in various applications, from image processing to bioinformatics applications. In recent years, the development of PAT tools based on ANNs and spectroscopic data has been an important advancement in the chemometrics. ANNs showed higher accuracy for monitoring processes that traditional chemometric methods and even were used for process control.

ANNs are nonlinear methods inspired by the biological neural system and consist of an input layer, hidden layers, and an output layer of nodes [22]. ANNs are able to predict the complex relationship between the input and output data after training and fine-tuning of ANN architecture [23].

In many comparative studies, ANNs often exhibit superior predictive accuracy when compared to other nonlinear and linear methods, ensuring more reliable models for various tasks [24].

Development of applications for data acquisition from various PAT devices and modeling of the obtained data is possible in open source programming languages such as Python, which reduces the cost of developing analytical in-line tools. In contrast with commonly used GUI oriented softwares, coding programming languages offer more flexibility and mathematical models can be tailored specifically for every case.

In this work, a seeded cooling crystallization of the API ceritinib (5-chloro-2-*N*-(5-methyl-4-piperidin-4-yl-2-propan-2-yloxyphenyl)-4-*N*-(2-propan-2-ylsulfonylphenyl) pyrimidine-2,4-diamine, Fig. 1) in tetrahydrofuran as the solvent was chosen as the model system. Ceritinib is an anaplastic lymphoma kinase (ALK) inhibitor primarily used for the treatment of ALK positive metastatic NSCLC [25].

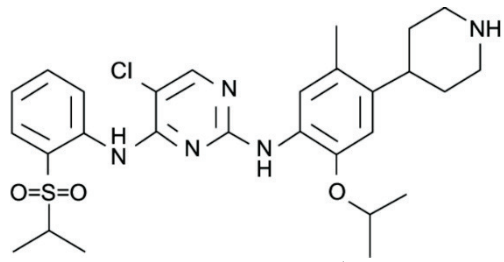


Figure 1 Molecular structure of ceritinib

The collected spectra were preprocessed using the Savitzky-Golay filter (first derivative) [26] to remove potential noise. Calibration models for solution concentration measurement were developed using linear chemometric methods (PCR and PLSR) and ANN (Fig. 2).

ANNs with different numbers of hidden layers and nodes per layer were created and compared. The optimal ANN architecture was selected based on validation of an independent dataset. Exponential linear unit (ELU) was used as activation function [27] in the hidden layers and the linear function was used in the output layer.

Developed ANN models predict the solution concentration with higher accuracy compared to linear models. In addition, ANN model with the lowest RMSEP was applied for solution concentration monitoring during cooling crystallization of ceritinib.

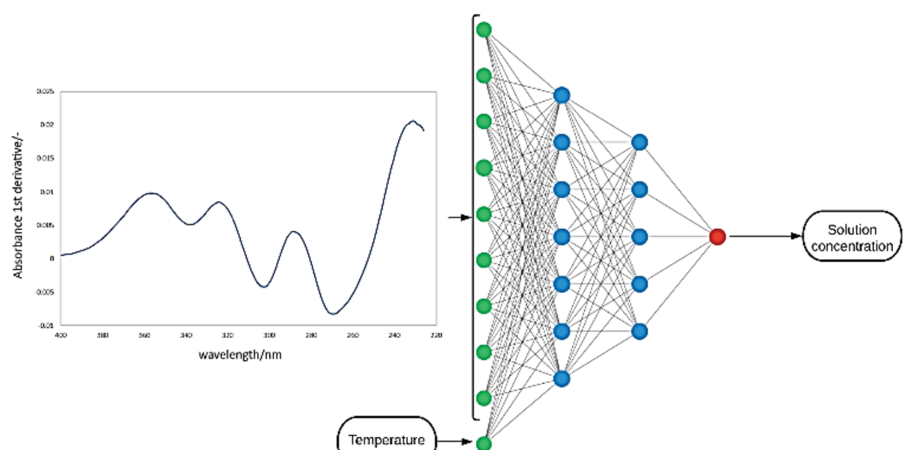


Figure 2 Schematic of ANN for solution concentration prediction

2 EXPERIMENTAL SECTION

2.1 Materials

Ceritinib is synthesized in-house and tetrahydrofuran (THF) was obtained from Kemika.

2.2 Solubility Determination of Ceritinib in Tetrahydrofuran

The solubility of ceritinib in THF was determined using the turbidity measurement embedded in the Blaze 900 process microscope in a 1 L jacketed glass reactor equipped with a PTFE four bladed pitched blade turbine mechanical agitator and a Pt-100 temperature sensor. A self-contained thermostat connected to the reactor jacket performed heating and cooling. A defined amount of ceritinib was added to the reactor containing 400 g of THF and heated at a constant rate (0.3 °C/min) until all crystals were dissolved, which was confirmed by turbidity measurement. Solubility points were determined for seven discrete concentrations (78.0, 116.4, 151.8, 187.8, 200.0, 223.8, 251.7 $g/kg_{solution}$) and these data points were used to regress the solubility equation parameters (Eq. (1) and Eq. (2)) in DynoChem Solubility Regress Design module. In Eq. (1) and Eq. (2), R is the ideal gas constant and equation parameters are lnA/-, B/kJmol^{-1} and C/(kJmol)^2: Solubility model with higher R^2 was used for planning of spectra collection within supersaturated zone.

$$C^* = \exp\left(\ln A - \frac{B}{R \cdot T}\right) \tag{1}$$

$$C^{*}=\exp\left(\ln A-\frac{B}{R\cdot T}-\frac{C}{(R\cdot T)^{2}}\right) \tag{2}$$

2.3 UV/Vis Spectra and Temperature Collection for Model Building

UV/Vis spectra in the 200 – 800 nm range were acquired using an immersible ATR fiber optic probe (Katana XP12 from Hellma Analitycs) connected to the UV/Vis spectrometer (Agilent Cary 60).

The calibration data set was acquired for nine different concentrations (99.8, 126.2, 142.7, 154.5, 170.8, 195.9, 227.3, 245.3, 273.5 $g/kg_{solution}$) within the metastable zone. The test data set was acquired for three different concentrations (109.9, 180.4, 214.3 $g/kg_{solution}$) also within the metastable zone. Data acquisition began with the lowest concentration by adding the appropriate amount of ceritinib together with 400 g of THF in the 1 L crystallizer equipped with a pitched blade turbine agitator and the monitoring system (Fig. 3).

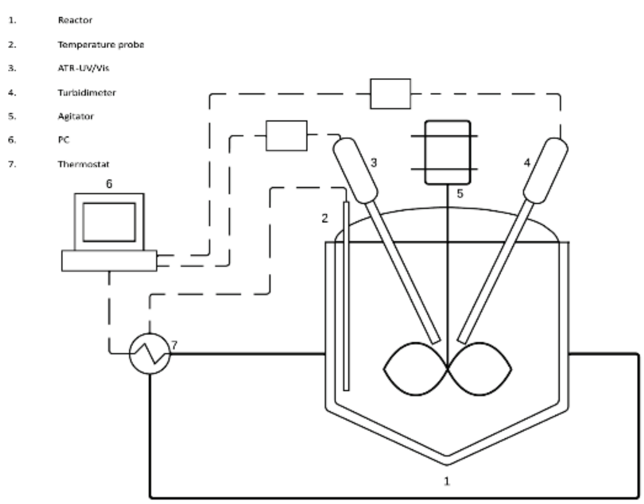


Figure 3 Schematic of equipment setup

Prior to data collection, the ceritinib suspension was heated up until complete dissolution of the crystals was observed. The undersaturated solution was cooled down at a cooling rate of 0.4 °C/min, and UV/Vis spectra and temperatures were collected during the cooling stage within the metastable zone. After data collection for a given concentration, the solution was reheated again into the undersaturated region and an additional amount of ceritinib was added into the reactor to achieve a higher concentration. The previously described procedure was repeated until data was collected for the highest concentration.

2.4 Data Preprocessing

The impact of different preprocessing methods on the validity of the calibration model were tested. Calibration models were developed for: 1) truncated spectra 2) smoothed and truncated 3) smoothed - derived spectra with Savitzky-Golay filter (second order polynomial, first derivative) and truncated.

The Savitzky-Golay filter is particularly useful in spectroscopy because it can effectively remove noise from a signal while preserving the shape of the underlying spectrum. The Savitzky-Golay filter uses a moving window of data points and fits a polynomial function to the data within the window [28]. The coefficients of the polynomial are then used to estimate the smoothed value at the center of the window. In addition, the derivative in each filter window can also be calculated with the given derivative order. When the filter is applied near the edges of the data, the polynomial fit can be distorted by missing data points. Because of these effects the filter is typically calculated using the entire spectra and then end values are excluded prior to modeling [29].

For that reason, the spectra were truncated after preprocessing to range from 226 to 400 nm where the significant peaks of ceritinib are located and the temperature was standardized.

Effect of preprocessing on model validity was compared for all three methods of preprocessing for models based on both, ANN and PLSR.

2.5 Model Building with ANN and PLSR

ANN models were developed using Python and Keras library. The feed-forward ANNs with back-propagation and multilayer perceptron architecture was used. The input variables of the ANNs were preprocessed truncated spectra along with standardized temperatures. The output variable was the concentration of ceritinib solution. The collected calibration samples were randomly divided into a training (90%) and a validation set (10%). The Adam optimization algorithm was used for ANN training with a batch size of 128 which is the number of training samples used in a single iteration of the training process. Adam optimizer is a stochastic gradient descent method that uses a combination of first and second-order moments to adapt the learning rate during training. The method is well-suited for problems with large datasets and high-dimensional parameter spaces, as it can efficiently compute the required statistics for updating the parameters. Additionally, Adam is generally less sensitive to the choice of hyperparameters than other optimization algorithms, making it easier to use in practice [30]. The ELU activation function was used in hidden layers since it can handle negative inputs and speeds up learning in ANN and linear activation function was used in output layer.[27] To assess the accuracy and validity of the established models based on different preprocessing methods and ANN architecture parameters, root mean square error of prediction (*RMSEP*) was used.

$$RMSEP=\sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i-y_i)^2}{n}} \tag{3}$$

In Eq. (3) $\hat{y}_i$ is the prediction value, y_i is the measured value and n shows the number of samples.

The ANN architecture parameters (number of hidden layers, number of neurons in the hidden layers) and number of epochs were altered, and the model with the lowest RMSEP was selected as optimal.

In addition, PLSR models were developed in Python using the calibration data set. The models were validated by test set validation method. The optimal number of PLSR factors was determined based on the lowest RMSEP.

The predictive capabilities of ANN and PLSR models were compared on an independent data set (model testing set) using the RMSEP as criterion.

2.6 Monitoring of the Seeded Cooling Crystallization Process

A seeded cooling batch crystallization was conducted in the same set-up as for the calibration experiments. ATR-UV/Vis probe was used for real-time monitoring of ceritinib solution concentration. Turbidity was measured in situ using the Blaze 900 system. Seeding point was at 250 $g/kg_{solution}$ and 37 °C. The amount of seeds used was 2 wt% relative to the initial mass of ceritinib. The seeds were prepared by micronization of crystallized ceritinib on an air-jet mill. The content of the reactor was cooled down to the final temperature of 2 °C over a period of 100 minutes using a cubic temperature cooling profile [3].

3 RESULTS AND DISCUSSION

3.1 Solubility Model

The solubility models of ceritinib form A in tetrahydrofuran at temperatures in the range of 0 – 60 °C was obtained by regression of parameters for Eq. (1) and Eq. (2) to the experimentally determined solubility points (Fig. 4).

Table 1 Solubility models of ceritinib in tetrahydrofuran

Solubility model 1	$C^* = \exp\left(12.06 - \frac{17.78}{R \cdot T}\right)$	$R^2 = 0.991$
Solubility model 2	$C^* = \exp\left(30.41 - \frac{109.02}{R \cdot T} - \frac{-113.00}{(R \cdot T)^2}\right)$	$R^2 = 0.994$

Values of fitted parameters are shown in Tab. 1. Solubility model 2 fitted experimental points with the higher correlation coefficient ($R^2 = 0.994$). Eq. (2) fits the solubility with higher accuracy when there is a very strong temperature-dependence of solubility. This could be interpreted as an increased dependence of solute activity coefficient and mole fraction on temperature [31].

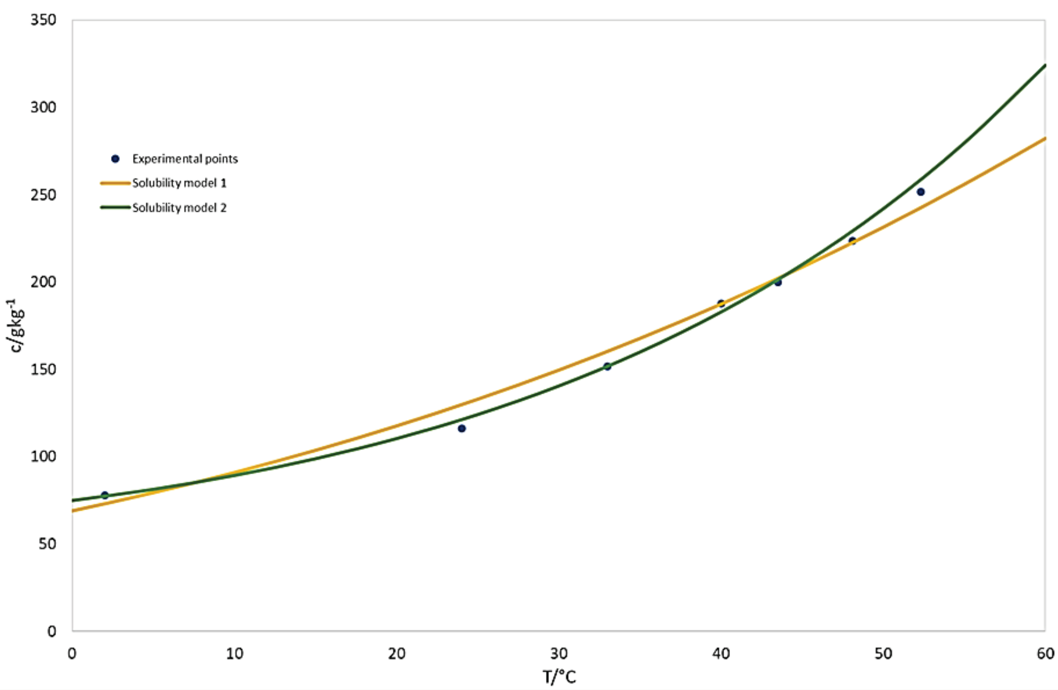


Figure 4 Ostwald-Miers diagram of ceritinib in tetrahydrofuran

3.2 UV/Vis Spectra and Temperature Collection for Model Building

Fig. 5 shows the collected data points for calibration and validation of the ANN and PLSR models for prediction of solution concentration of ceritinib. Data was collected in the temperature range 2 – 55 °C and in the concentration range of 99.8 – 273.5 $g/kg_{solution}$. Most of the data was collected within the metastable zone to ensure that the monitored variables during crystallization were within the range of the developed model. The metastable zone of ceritinib in tetrahydrofuran is wide and nucleation was not detected during data collection.

The collected data from undersaturated zone were excluded to reduce nonlinearities caused by the effect of temperature on the refractive index. As the temperature of a solution changes, its refractive index can change, causing shifts in the wavelengths at which light is absorbed [32]. This approach to developing a model to monitor crystallization can achieve higher accuracy. In contrast, models developed with combined data from both the undersatured zone and the metastable zone have lower accuracy but the dissolution process can be monitored. If both processes need to be monitored, it is better to develop two separate models: one to monitor crystallization and the other to monitor the dissolution process [12].

Figure 5 Collected data for ANN and PLSR model building and testing

3.3 Data Preprocessing

The spectra were preprocessed by calculating the first derivative and applying a Savitzky-Golay filter to remove noise and baseline shifts, factors that contribute to nonlinearities in the data set. Preprocessing with the first and second derivatives can be useful for fouling detection by revealing abnormal spectral changes [33]. After preprocessing, spectra were truncated from 226 to 400 nm, where characteristic ceritinib absorbance maxima are present. Models built with preprocessed and truncated spectra were found to be more accurate and reliable.

Fig. 6a) shows preprocessed and truncated spectra for a ceritinib concentration of 273.5 $g/kg_{solution}$ at different temperatures (34, 40, 56 °C). It can be seen that the absorption maxima change with temperature in such a way that the absorption intensifies with decreasing temperature. Models built without this temperature effect show an apparent increase in the concentration result during cooling of a clear solution. For this reason, it is very important to include data points for the same concentration at different temperatures when training models [16, 33].

Fig. 6b) shows preprocessed and truncated spectra for different concentrations of ceritinib at the same temperature (21 °C). It can be seen that solution concentration has a stronger effect on the spectra than temperature. This is an important feature because high sensitivity is important for measuring small changes in solution concentration during crystallization. Models developed with spectra where the sensitivity of the absorbance change to concentration changes is high, can be used in automated process control and data collection for kinetic crystallization models [12]. Additionally, a significant peak shift was found at 270 nm and 320 nm as the concentration increases, contributing to the nonlinearity.

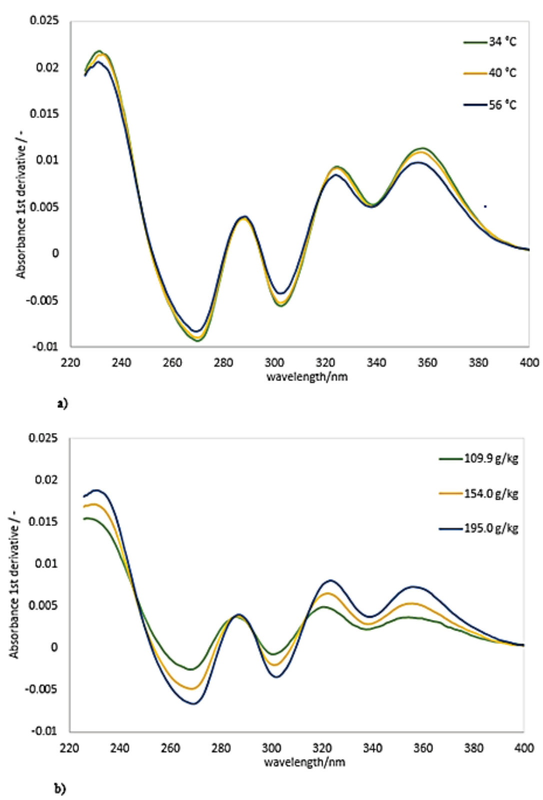


Figure 6 UV/Vis preproccessed spectra: a) for different temperatures at same ceritinib concentration; b) for different concentrations of ceritinib at same temeprature

3.4 Model Development

The ANN and PLSR models were developed using the same calibration experiments data set, shown in Figure 5. The concentration prediction ability of the developed models was compared using the model testing data set shown in Fig. 5.

During the development of ANN models, the calibration data set was randomly divided into training (90%) and validation set (10%) groups to minimize the possibility of overfitting. Any given ANN model is overfitted when it describes the training set with high accuracy, but on the other hand, does not accurately predict the test data set. Overfitting occurs for two reasons. First, the model is trained for too many epochs so the model starts to fit the noise within the data set, and second the architecture of ANN is too complex [35]. In contrast, underfitting occurs when the model is not trained for a sufficient number of epochs and is not able to predict relationships between input and output data with high accuracy [36]. In order to avoid overfitting and underfitting, ANNs with different architectures and numbers of epochs have been developed. Predictive performances of developed ANN models were tested on test data set and results are shown in Tab. 2. Furthermore, the impact of preprocessing on model accuracy can be seen when comparing models with the lowest (*RMSEP*) for truncated spectra (0.692 g/kg), smoothed/truncated (0.653 g/kg) and smoothed/derived/ truncated spectra (0.405 g/kg). Smoothing without derivative reduced noise in data set but minimal change in *RMSEP* is observed. While the combination of smoothing and derivative showed significant change in *RMSEP* due to partially elimination of baseline shifting, which tends to occur frequently during measurements.

Table 2 Comparison of RMSEP for ANN models

Neurons in each hidden layer	Hidden Layers	Epochs	*RMSEP*/ $gkg^{-1}_{solution}$		
			truncated	smoothed/truncated	smoothed/derived/truncated
4	1	4000	1.496	1.561	2.898
4	1	7000	1.723	1.252	1.520
4	1	10000	1.330	1.398	0.893
4	5	4000	1.256	1.888	0.756
4	5	7000	1.032	1.376	0.531
4	5	10000	1.131	0.886	0.486
4	9	4000	1.421	1.614	0.615
4	9	7000	1.138	1.077	0.523
4	9	10000	1.013	1.060	0.551
22	1	4000	1.734	0.809	2.150
22	1	7000	0.744	1.608	0.905
22	1	10000	0.878	**0.653**	0.651
22	5	4000	0.781	0.843	0.593
22	5	7000	0.788	1.720	0.464
22	5	10000	0.860	1.543	0.444
22	9	4000	0.935	0.850	**0.405**
22	9	7000	1.823	0.926	0.539
22	9	10000	2.270	1.570	0.62
40	1	4000	0.757	0.668	2.024
40	1	7000	1.000	0.848	0.750
40	1	10000	**0.692**	0.784	0.583
40	5	4000	1.444	0.696	0.509
40	5	7000	0.949	1.138	0.415
40	5	10000	1.559	1.427	0.445
40	9	4000	1.439	1.918	0.667
40	9	7000	1.252	2.568	1.250
40	9	10000	1.785	1.83	0.805

The ANN model developed on smoothed/derived/ truncated data which was trained for 4000 epochs and carrying 9 hidden layers and 22 neurons in the each hidden layer, resulted in the lowest *RMSEP* of 0.405 $g/kg_{solution}$. This model was used to monitor solution concentration during seeded cooling crystallization.

The optimal number of PLSR factors was chosen based on the *RMSEP*. PLSR model with four factors resulted in lowest *RMSEP* for all types of pretreatment used on spectra. Same effect of spectra pretreatment is observed for PLSR like for ANN models. PLSR model with four factors and smoothed/derived/truncated yield the lowest *RMSEP* of 1.295 $g/kg_{solution}$. A larger number of factors used in the model contributed to an increase in the *RMSEP* value due to overfitting, as noise was also modeled in the data set (Tab. 3). The model validation was performed using the test set validation method.

Table 3 Comparison of *RMSEP* for PLSR models

Number of factors/-	*RMSEP*/$gkg^{-1}_{solution}$		
	truncated	smoothed/truncated	smoothed/derived/truncated
1	6.360	6.388	22.858
2	5.505	5.550	2.704
3	2.125	2.164	2.114
4	**1.427**	**1.423**	**1.295**
5	1.530	1.493	1.497

Due to a wide concentration range (99.8 – 273.5 $kg/kg_{solution}$) used for model calibration, the absorbance maxima shift (Fig. 6a) and the change in solution density and consequently the change in the refractive index all contributed to nonlinearities in the data set.

The predictive performance of the ANN and PLSR models for the model testing data set is summarized in Tab. 2 and Tab. 3. The ANN models predicted solution concentration with a lower *RMSEP* compared to PLSR, with the exception of the undertrained network (trained for an insufficient number of epochs). This is due to the nonlinearities in the data set that the PLSR algorithm cannot model. It is expected that the difference in predictive ability will be even greater for data set with more nonlinearities (relationships).

3.5 Solution Concentration Monitoring During Cooling Crystallization

Solution concentration of ceritinib was monitored during a seeded cooling crystallization experiment with a nominal concentration of 250 $g/kg_{solution}$.

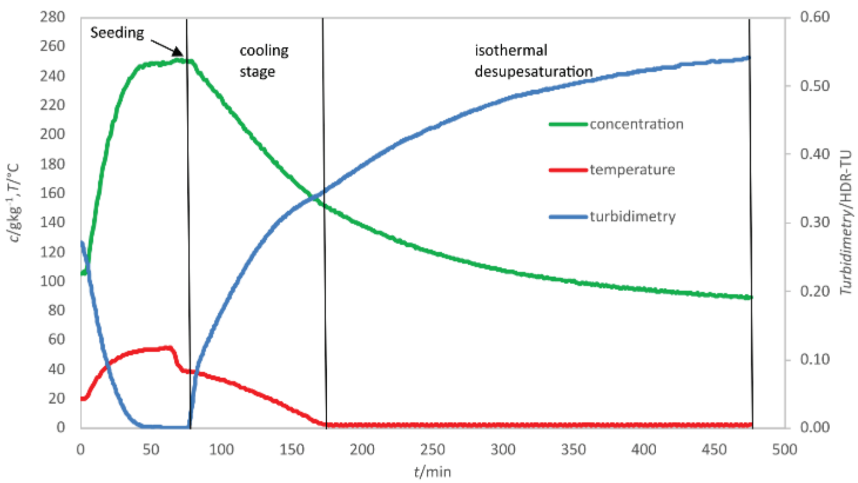


Figure 7 Solution concentration, temperature and turbidimetry during the crystallization experiment.

The experiment was initiated by heating the initial suspension of ceritinib in THF to 55 °C. The suspension was agitated until all suspended crystals were dissolved. After complete dissolution, confirmed by a turbidimetric probe, the monitoring system measured a solute concentration of 250 (±1) $g/kg_{solution}$, which was equal to the prepared nominal concentration. The solution was then cooled to 37 °C resulting in a supersaturation level of 80 $g/kg_{solution}$, followed by seed addition. After seed addition, the in-house developed supersaturation control system gradually cooled the suspension, keeping the supersaturation level constant during the cooling phase. Once the temperature inside the reactor reached 2 °C, the suspension was agitated for another five hours until complete desupersaturation was observed (Fig. 7).

4 CONCLUSION

This study has demonstrated that the solute concentration of ceritinib during a seeded cooling crystallization process can be successfully monitored in situ by in-line ATR UV/Vis spectroscopy in combination with a predictive ANN model. Comparison of models based on linear PLSR and non-linear ANN chemometric methods was performed. The ANN model predicted solution concentration with *RMSEP* of 0.405 $g/kg_{solution}$. In contrast, PLSR model with the lowest *RMSEP* of 1.295 $g/kg_{solution}$ was model with four factors. Based on the *RMSEP*, ANN models provided improved prediction capabilities for solution concentration monitoring of ceritinib. This was due to nonlinearities in the data set that the PLSR algorithm could not model. The difference in the prediction ability is expected to be even greater for more nonlinear data sets.

On the other hand, the advantage of models based on PLSR is data interpretation and shorter development time. The development of the ANN model was more time-consuming compared to PLSR because of the need to fine-tune the architecture of ANN and to determine the appropriate number of epochs for training ANN.

Additionally, spectra preprocessing with Savitzky-Golay filter (second order polynomial and first order derivative) reduced noise and baseline shift, which resulted in lower *RMSEP* for PLSR and ANN models.

These highly accurate measurements of solution concentrations can be further used to control the crystallization process or to determine the optimal cooling profile to achieve the desired (or required) particle size distribution.

5 REFERENCES

[1] Barrett, M., McNamara, M., Hao, H. X., Barrett, P. & Glennon, B. (2010). Supersaturation tracking for the development, optimization and control of crystallization processes. *Chem. Eng. Res. Des*., 88(8), 1108-1119. https://doi.org/10.1016/j.cherd.2010.02.010

[2] Chen, J., Sarma, B., Evans, J. M. B. & Myerson, A. S. (2011). Pharmaceutical crystallization. *Cryst. Growth Des*., 11(4), 887-895. https://doi.org/10.1021/cg101556s

[3] Kim, S. et al. (2006). Control of the particle properties of a drug substance by crystallization engineering and the effect on drug product formulation. *AIChE Annu. Meet. Conf. Proc.*, 8.

[4] Alatalo, H. et al. (2008). In-line monitoring of reactive crystallization process based on ATR-FTIR and Raman spectroscopy. *J. Chemom., 22*(11-12), 644-652 https://doi.org/10.1002/cem.1161

[5] Joiner, D. E., Billeter, J., Mcnally, M. E. P., Hoffman, R. M. & Gemperline, P. J. (2014). Comprehensive kinetic model for the dissolution, reaction, and crystallization processes involved in the synthesis of aspirin. *J. Chemom., 28*(5), 420-428. https://doi.org/10.1002/cem.2605

[6] Saleemi, A. N., Rielly, C. D. & Nagy, Z. K. (2012). Comparative investigation of supersaturation and automated direct nucleation control of crystal size distributions using ATR-UV/vis spectroscopy and FBRM. *Cryst. Growth Des., 12*(4), 1792-1807. https://doi.org/10.1021/cg201269c

[7] Feng, L. & Berglund, K. A. (2002). ATR-FTIR for Determining Optimal Cooling Curves for Batch Crystallization of Succinic Acid. *Cryst. Growth Des*., 2(5), 449-452. https://doi.org/10.1021/cg025545e

[8] Togkalidou, T., Tung, H. H., Sun, Y., Andrews, A. & Braatz, R. D. (2002). Solution concentration prediction for pharmaceutical crystallization processes using robust chemometrics and ATR FTIR spectroscopy. *Org. Process Res. Dev*., 6(3), 317-322. https://doi.org/10.1021/op015516x

[9] Nagy, Z. K. & Braatz, R. D. (2012). Advances and new directions in crystallization control. *Annu. Rev. Chem. Biomol. Eng., 3*, 55-75. https://doi.org/10.1146/annurev-chembioeng-062011-081043

[10] Chen, Z. P. et al. (2009). On-line monitoring of batch cooling crystallization of organic compounds using ATR-FTIR spectroscopy coupled with an advanced calibration method. *Chemom. Intell. Lab. Syst*., 96(1), 49-58. https://doi.org/10.1016/j.chemolab.2008.11.002

[11] Togkalidou, T., Fujiwara, M., Patel, S. & Braatz, R. D. (2001). Solute concentration prediction using chemometrics and ATR-FTIR spectroscopy. *J. Cryst. Growth, 231*(4), 534-543. https://doi.org/10.1016/S0022-0248(01)01518-4

[12] Zhang, F., Liu, T., Wang, X. Z., Liu, J. & Jiang, X. (2017). Comparative study on ATR-FTIR calibration models for monitoring solution concentration in cooling crystallization. *J. Cryst. Growth, 459*, 50-55. https://doi.org/10.1016/j.jcrysgro.2016.11.064

[13] Qu, H. et al. (2009). Raman and ATR FTIR spectroscopy in reactive crystallization: Simultaneous monitoring of solute concentration and polymorphic state of the crystals. *J. Cryst. Growth, 311*(13), 3466-3475. https://doi.org/10.1016/j.jcrysgro.2009.04.018

[14] Lin, M., Wu, Y. & Rohani, S. (2020). Simultaneous Measurement of Solution Concentration and Slurry Density by Raman Spectroscopy with Artificial Neural Network. *Cryst. Growth Des., 20*(3), 1752-1759. https://doi.org/10.1021/acs.cgd.9b01482

[15] Zhang, Y., Jiang, Y., Zhang, D., Li, K. & Qian, Y. (2011). On-line concentration measurement for anti-solvent crystallization of β-artemether using UVvis fiber spectroscopy. *J. Cryst. Growth, 314*(1), 185-189. https://doi.org/10.1016/j.jcrysgro.2010.10.165

[16] Billot, P., Couty, M. & Hosek, P. (2010). Application of ATR-UV spectroscopy for monitoring the crystallisation of UV absorbing and nonabsorbing molecules. *Org. Process Res. Dev*., 14(3), 511-523. https://doi.org/10.1021/op900281m

[17] Herman, C., Haut, B., Douieb, S., Larcy, A., Vermylen, V. & Leyssens, T. (2012). Use of in situ Raman, FBRM, and ATR-FTIR probes for the understanding of the solvent-mediated polymorphic transformation of II-I etiracetam in methanol. *Org. Process Res. Dev., 16*(1), 49-56. https://doi.org/10.1021/op200168g

[18] Trampuž, M., Teslić, D. & Likozar, B. (2018). Batch Crystallization of Active Pharmaceutical Ingredient: Kinetics and Equilibrium Modelling. *Kem. Ind., 67*(13), 43-48. https://doi.org/10.15255/KUI.2018.014

[19] Saleemi, A. N., Steele, G., Pedge, N. I., Freeman, A. & Nagy, Z. K. (2012). Enhancing crystalline properties of a cardiovascular active pharmaceutical ingredient using a process analytical technology based crystallization feedback control strategy. *Int. J. Pharm., 430*(1-2), 56-64. https://doi.org/10.1016/j.ijpharm.2012.03.029

[20] Wold, S., Sjöström, M. & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems, 58*(2), 109-130. https://doi.org/10.1016/S0169-7439(01)00155-1

[21] Barmpalexis, P., Karagianni, A., Nikolakakis, I. & Kachrimanis, K. (2018). Artificial neural networks (ANNs) and partial least squares (PLS) regression in the quantitative analysis of cocrystal formulations by Raman and ATR-FTIR spectroscopy. *J. Pharm. Biomed. Anal., 158*, 214-224. https://doi.org/10.1016/j.jpba.2018.06.004

[22] Marini, F., Bucci, R., Magrì, A. L. & Magrì, A. D. (2008). Artificial neural networks in chemometrics: History, examples and perspectives. *Microchem. J., 88*(2), 178-185. https://doi.org/10.1016/j.microc.2007.11.008

[23] Passos, D. & Mishra, P. (2022). A tutorial on automatic hyperparameter tuning of deep spectral modelling for regression and classification tasks. *Chemom. Intell. Lab. Syst., 223*. https://doi.org/10.1016/j.chemolab.2022.104520

[24] Xiong, Y., Zhang, Z. & Chen, F. (2010). Comparison of Artificial Neural Network and Support Vector Machine Methods for Urban Land Use / Cover Classifications from Remote Sensing Images. in *ICCASM 2010*, 52-56.

[25] Khozin, S. et al. (2015). FDA approval: Ceritinib for the treatment of Metastatic Anaplastic lymphoma Kinase-positive non-small cell lung cancer. *Clin. Cancer Res., 21*(11), 2436-2439. https://doi.org/10.1158/1078-0432.CCR-14-3157

[26] Guo, Y., Liu, C., Ye, R. & Duan, Q. (2020). Advances on water quality detection by uv-vis spectroscopy. *Applied Sciences (MDP), 10*(19). 1-18. https://doi.org/10.3390/app10196874

[27] Clevert, D. A., Unterthiner, T. & Hochreiter, S. (2016). Fast and accurate deep network learning by exponential linear units (ELUs). *The 4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings*, 1-14. https://doi.org/10.48550/arXiv.1511.07289

[28] Schmid, M., Rath, D. & Diebold, U. (2022). Why and How Savitzky-Golay Filters Should Be Replaced. *ACS Meas. Sci. Au*, 2(2), 185-196. https://doi.org/10.1021/acsmeasuresciau.1c00054

[29] Eilers, P. H. C. (2003). A perfect smoother. *Anal. Chem.*, 75(14), 3631-3636. https://doi.org/10.1021/ac034173t

[30] Kingma, D. P. & Ba, J. L. (2015). Adam: A method for stochastic optimization. *The 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, 1-15.

[31] Daubies, A. (2014). Applicability evaluation of AspenONE and Dynochem for solubility modeling in the pharmaceutical industry.

[32] Saleemi, A. N., Rielly, C. D. & Nagy, Z. K. (2012). Monitoring of the combined cooling and antisolvent crystallisation of mixtures of aminobenzoic acid isomers using ATR-UV/vis spectroscopy and FBRM. *Chem. Eng. Sci., 77*, 122-129.

https://doi.org/10.1016/j.ces.2012.02.005
[33] Zhang, F., Du, K., Guo, L., Xu, Q. & Shan, B. (2021). Comparative Study of Preprocessing on an ATR-FTIR Calibration Model for In Situ Monitoring of Solution Concentration in Cooling Crystallization. *Chem. Eng. Technol., 44*(12), 2279-2289. https://doi.org/10.1002/ceat.202100104
[34] Chong, M. W. S. et al. (2022). Temperature Correction of Spectra to Improve Solute Concentration Monitoring by In Situ Ultraviolet and Mid-Infrared Spectrometries toward Isothermal Local Model Performance. *Org. Process Res. Dev., 26*(11), 3096-3105. https://doi.org/10.1021/acs.oprd.2c00238
[35] Belkin, M., Hsu, D., Ma, S. & Mandal, S. (2019). Reconciling modern machine-learning practice and the classical bias–variance trade-off. *Proc. Natl. Acad. Sci. U. S. A., 116*(32), 15849-15854. https://doi.org/10.1073/pnas.1903070116
[36] Pothuganti, Swathi. (2018). Review on over-fitting and under-fitting problems in Machine Learning and solutions. *Int. J. Adv. Res. Electr. Electron. Instrum. Eng.*, 7(9), 3692-3695. https://doi.org/10.15662/IJAREEIE.2018.0709015

Authors' contacts:

Ivan Vrban, mag. ing. cheming.
(Corresponding author)
Teva Pharmaceutical Industries Ltd.
Prilaz Baruna Filipovića 25,
10000 Zagreb, Croatia
ivan.vrban@pliva.com

Damir Šahnić, dr. sc.
Teva Pharmaceutical Industries Ltd.
Prilaz Baruna Filipovića 25.
10000 Zagreb, Croatia
damir.sahnic@pliva.com

Nenad Bolf, prof. dr. sc.
University of Zagreb, Faculty of Chemical Engineering and Technology,
10000 Zagreb, Croatia
bolf@fkit.hr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20230710223321

Original scientific paper
Received: 2023-07-10, Accepted: 2023-08-04

Analytical Description of the Electric Field inside a High Voltage Glass Insulator

Tomislav Barić, Hrvoje Glavaš, Željko Hederić, Mirko Karakašić*

Abstract: Research presented in this paper analyzes the accuracy and applicability of analytical expressions to describe the electric field of a high-voltage glass insulator. The metal cap and the metal pin of the glass insulator are represented by coaxial cylindrical electrodes. The insulating materials between them are modeled as coaxial hollow cylinders. Analytical expressions for the electric field in these insulating materials are derived. Using these expressions, the spatial distribution of the electric field in the insulating materials was determined. Suitable software, based on the finite element method (FEM), was used to determine the spatial distribution of the electric field in dielectrics in the same space. Both analyses were performed under quasi-static conditions, and the results of the analytical expressions were compared with the FEM results. Discrepancies between the results of the two mentioned analyses were observed. The reasons for the observed discrepancies and the applicability of the presented analytical expressions are commented on.

Keywords: analytical description; electric field; finite element method (FEM); glass insulator; high voltage; numerical calculation

1 INTRODUCTION

The significant development of the electric power grid began in the second half of the 19th century. Since then, the power grid has become more complex and the electric power which is carried from different sources to consumers is getting increasingly higher. To reduce transmission losses, voltage levels have become higher. Reliability and availability are also getting higher. All of the above has become possible with the simultaneous development of all components in the transmission and distribution of electricity. One of the important components in the transmission and distribution of electricity are insulators [1-3].

In this article, when researches talk about insulators, they refer to high-voltage (HV) insulators with the application in the transmission and distribution of electricity. As the mentioned development of the electric power grid took place all the time from applied electricity beginnings, also developments of HV isolators have been continuously improved.

Knowledge of the exact electric field (E-field) distribution is crucial in the process of designing and analysing HV components such as HV insulators. The calculation of E-fields requires the solution of the field formulation according to Maxwell's equations with meeting the boundary requirements [4]. This can be done by applying analytical or numerical methods [5-7]. Due to the geometric complexity of HV components, analytical expressions are rarely applicable [4]. To demonstrate, a short review of the selected examples will be provided. Analytical expressions in HV technology can be used to determine the electric field in the vicinity of HV overhead transmission lines [8, 9]. Also, analytical expressions can be used to determine the induced current densities and voltages in a two-layer soil by an HV transmission line [10]. Since the energy stored in an E-field must be equal to the energy stored in the equivalent capacitance, when the expression for the E-field of a particular geometry is known, the capacitance of this geometry is also known. This enables the application of analytical expressions for E-fields to determine the capacitance of overhead transmission lines [11]. The aforesaid also applies to HV cables [12]. With analytical expressions, it is possible to describe the E-field inside of HV direct current cable insulations, considering a varying conductivity in the vicinity of an interface, e.g. in cable joints, which usually consists of two different dielectrics [13]. Analytical expressions in HV technology can also be combined with the results obtained by numerical methods. This has proven to be particularly suitable for analysing and interpreting the results obtained by numerical methods. For example, by combining analytical expressions and simulation results obtained by the Finite Element Method (FEM), it is possible to analyse and interpret the intensification of the E-field due to the presence of water droplets on the surface of HV insulators [14].

A special advantage of numerical methods, in relation to the analytical ones, is that they enable the analysis of arbitrarily complex geometries and geometric configurations, while certain approximations of geometric shapes are always present in analytical methods. Due to the stated advantages of numerical methods, the use of analytical methods in describing the E-field in insulators has decreased, and has given way to numerical methods in electromagnetism. Today, the use of numerical methods in electromagnetism exclusively dominates the analysis of electromagnetic fields in insulators.

Nevertheless, it is still beneficial to have suitable analytical expressions, although approximate, to describe the E-fields in insulators. Analytical expressions simplify theoretical considerations, indicate functional relations between physical quantities and enable rapid qualitative considerations. At the same time, if they give reasonably (acceptable) accurate results, the results obtained by them can serve as an indicator of the validity of numerical calculation. That is, to check whether there is any error in modelling using numerical methods. This was a strong motivation to investigate whether it is possible, and if yes, how accurately the E-field in the space between the metal cap and the metal pin can be described.

In order to derive analytical expressions that can describe the electric field in the specified zone, an approximation of the geometry of the specified zone was performed. The metal cap and metal pin are represented by coaxial cylindrical electrodes. The insulating materials between the metal cap and the metal pin (cement and glass) are modelled as hollow coaxial cylinders. Using derived expressions on a concrete example of a disk insulator, the spatial distribution of the E-field in dielectrics was determined. In order to determine the accuracy of the obtained results, the same calculation was performed by the FEM. For this purpose, commercial software Ansys based on the FEM was used. Deviations of the results obtained by analytical expressions in relation to the accurate results obtained by applying the FEM were observed. The underlying reasons for the observed deviations and the applicability limits of the presented analytical expressions are commented on.

2 HIGH VOLTAGE GLASS INSULATORS

Significant progress on several levels has been made since HV insulators were first used. The technological process for producing insulating materials without microdefects and microinhomogeneities has been improved. New insulating materials, such as composite polymers, have been discovered [15, 16]. The insulator geometry has been improved and the surface has been treated with materials (coatings) that reduce their adhesion to the insulator surface [17-19]. This is especially important for porcelain insulators, where moisture and dirt deposits on the surface of porcelain insulators affect the E-field distribution, which can lead to partial discharges. The influence of non-uniform thermal properties of the materials making up the insulators, i.e. non-uniform expansions due to temperature changes causing material stresses and micro-fractures, are minimized by proper design.

Despite advances in the synthesis of polymeric composite insulators with suitable mechanical, thermal and electrical properties for the fabrication of HV insulators, traditionally used materials such as porcelain and glass still dominate in applications. One such glass insulator is described in this article. Common types of HV insulators used in transmission and distribution are pin insulators, shackle insulators, suspension insulators (disk or cord), strain or tension insulators. As for the insulating materials used, insulators for power transmission and distribution HV are made of glass, ceramic materials, such as porcelain, or polymers, also called polymer composites. Sometimes, the insulator name also includes the name of the insulator geometric shape, e.g. disk insulators.

In this article, the focus is on disk-shaped (bell-shaped) insulators. HV cap-and-pin (disk) insulators [2], regardless of which material is chosen as the primary insulator (glass or porcelain), have a similar shape. A typical HV disk insulator (porcelain or glass as the primary insulator material) has the shape of a bell (disk) (Fig. 1 [20, 21]). When the line voltage requires more insulation than a single disk insulator can provide, disk insulators are connected in series.

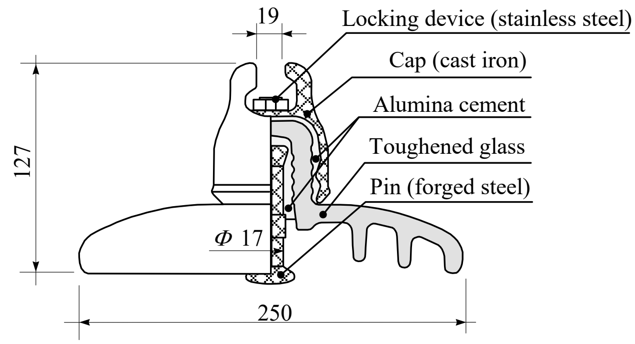


Figure 1 Structure of the analyzed glass disk insulator

3 THEORETICAL BACKGROUND (BOUNDARY CONDITIONS)

Solving electromagnetic problems requires knowledge of the boundary conditions at the boundary between two media with different electromagnetic properties [22-24]. In this section, a brief overview of the boundary conditions at the boundaries of the two media found in the case study analyzed in this article is given. The boundary conditions for an E-field on an interface separating two homogeneous isotropic materials can be written in a concise form [22, 23]:

$$\vec{n} \times (\vec{E}_1 - \vec{E}_2) = 0 , \tag{1}$$

where $\vec{n}$ is the normal vector to the boundary surface, $\vec{E}_1$ is electric field in medium no 1 next to the boundary surface and $\vec{E}_2$ is electric field in medium no 2 next to the boundary surface. Basically, the previous equation shows that the tangential component of the electric field remains unchanged at the boundary of the two media. That is, the tangential component of the electric field is continuous across the interface (Fig. 2) [23]:

$$\vec{E}_{1,t} = \vec{E}_{2,t} . \tag{2}$$

Boundary conditions for electric displacement field $\vec{D}$ on a boundary surface separating two homogeneous isotropic materials can be written in a concise form [23]:

$$\vec{n} \cdot (\vec{D}_1 - \vec{D}_2) = 0 , \tag{3}$$

where $\vec{n}$ is the normal vector to the boundary surface, $\vec{D}_1$ is electric displacement field in the medium denoted by 1 next to the boundary surface, $\vec{D}_2$ is electric displacement field in the medium denoted by 2 next to the boundary surface. Basically, the previous equation shows that the normal component of the electric displacement field remains unchanged at the boundary of the two media (Fig. 2). That is, the normal component of the electric displacement field $\vec{D}$ is continuous across the interface [23]:

$$\vec{D}_{1,n} = \vec{D}_{2,n} \ . \tag{4}$$

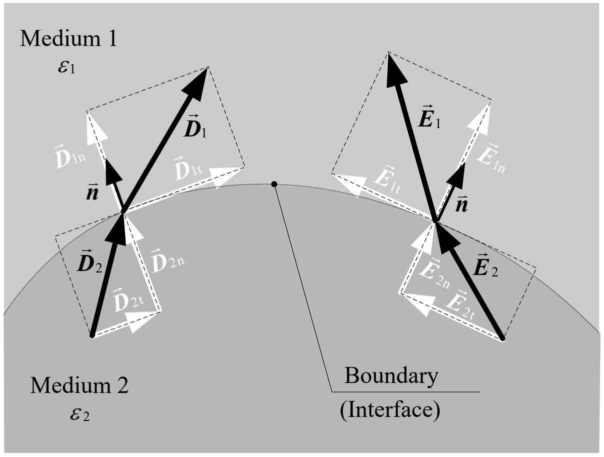


Figure 2 Graphical representation of boundary conditions for the electric field and electric displacement field at the interface of two media with different permitivities

4 MODELLING THE ELECTRIC FIELD OF THE GLASS INSULATOR

The space of interest for modelling is the space between the metal cap and the metal pin (Fig. 3 and 4). The inner cylinder represents the metal pin and the outer cylinder represents the metal cap. The space between these two metal cylinders is filled with a complex (heterogeneous) insulating material. That is, with the cement around the inner electrode, then with the glass, and again with the cement between the glass and the outer electrode. Each layer of the insulating material is represented by a hollow cylinder, and the cylinders are coaxial (Fig. 3 and 4).

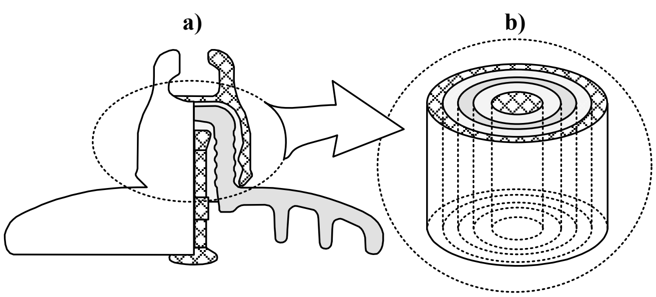


Figure 3 A sketch of the disc insulator and an equivalent system of coaxial electrodes with a complex dielectric between them

In the following text, are presented the key steps in deriving expressions for the E-field of a coaxial structure consisting of three dielectrics. The presented derivation is an extension of the typical derivation for E-fields of coaxial structures to the case of multiple dielectrics [22, 23]. The expressions presented for determining the electric field in the space between coaxial metal cylinders are derived assuming very long cylinders. However, when this is not the case, as in this article, the expressions provide an approximation of the electric field. Another limitation of the analytical approach is cement modelling. Electrical conductivity of cement (semi-conductive, Tab. 1) is high compared to insulating materials, which affects the field distribution.

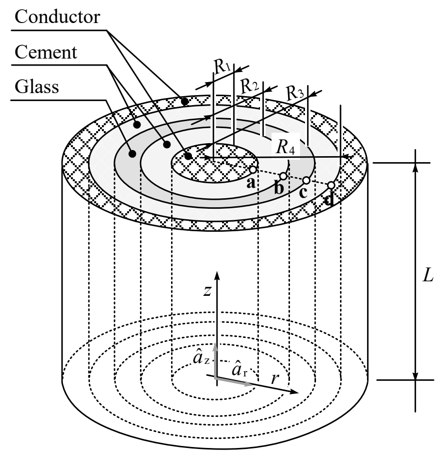


Figure 4 Detailed view of the equivalent system of coaxial electrodes with a complex dielectric between them

Since cylinders are initially assumed to be infinite in length, the electric field has only a radial component. The electric field is a function of the distance from the longitudinal axis of the coaxial cylinders and it also depends on the dielectric constant of the medium.

$$\vec{E}(r) = \begin{cases} E_1(r)\cdot\hat{a}_r \ \text{ for } R_1 \le r \le R_2, \\ E_2(r)\cdot\hat{a}_r \ \text{ for } R_2 \le r \le R_3, \\ E_3(r)\cdot\hat{a}_r \ \text{ for } R_3 \le r \le R_4, \end{cases} \tag{5}$$

where $\hat{a}_r$ is unit vector along the r axis. The electric field in the space around the inner electrode, a space filled with cement, is denoted by E_1. The electric field in the space filled with glass is denoted by E_2. The electric field in the space filled with cement (the space between the glass and the outer electrode) is denoted by E_3. The E-field in the space around a flat and very long cylindrical conductor is determined by the expression [22, 23]:

$$E_1(r) = \frac{\lambda}{2\pi\varepsilon_1}\cdot\frac{1}{r}, \tag{6}$$

where ε_1 is the permittivity of the medium in which the inner electrode is located (cement), r is radial distance from the longitudinal axis of the inner electrode and λ is line charge density. According to the boundary condition described by Eq. (4), the following applies:

$$D_1(R_2) = D_2(R_2), \tag{7}$$

$$D_2(R_3) = D_3(R_3). \tag{8}$$

Applying constitutive equation to expression (7) and (8), and taking into account (6), gives:

$$\vec{E}(r)=\begin{cases}\dfrac{\lambda}{2\pi\varepsilon_1}\cdot\dfrac{1}{r}\cdot\hat{a}_\mathrm{r}\ \text{ for }\ R_1\le r\le R_2,\\ \dfrac{\lambda}{2\pi\varepsilon_2}\cdot\dfrac{1}{r}\cdot\hat{a}_\mathrm{r}\ \text{ for }\ R_2\le r\le R_3,,\\ \dfrac{\lambda}{2\pi\varepsilon_1}\cdot\dfrac{1}{r}\cdot\hat{a}_\mathrm{r}\ \text{ for }\ R_3\le r\le R_4.\end{cases} \tag{9}$$

For the practical application of the previous expression, it is necessary to determine the line charge density λ. One way this can be done is presented below. The relationship between electric potential and electric field is described by the expression [22, 23]:

$$\vec{E}=-\mathrm{grad}\varphi\,. \tag{10}$$

By integrating the previous expression along the integration curve lying on the radial axis, the expression for the voltage between the outer and inner metal cylinders is obtained:

$$U_{da}=\varphi_d-\varphi_a=-\int_a^d \vec{E}\mathrm{d}\vec{s}\,. \tag{11}$$

Taking into account Eq. (9), the previous line integral with integration limits from a to d, needs to be broken down into three line integrals:

$$U_{da}=-\int_a^b \vec{E}_1\mathrm{d}\vec{s}-\int_b^c \vec{E}_2\mathrm{d}\vec{s}-\int_c^d \vec{E}_3\mathrm{d}\vec{s}\,. \tag{12}$$

Inserting Eq. (9) into the previous expression and after performing the integration gives:

$$U_{ad}=\frac{\lambda}{2\pi\varepsilon_1}\ln\frac{R_2}{R_1}+\frac{\lambda}{2\pi\varepsilon_2}\ln\frac{R_3}{R_2}+\frac{\lambda}{2\pi\varepsilon_1}\ln\frac{R_4}{R_3}\,. \tag{13}$$

The charge on the electrodes is determined by the expressions:

$$Q=C\cdot U_{ad}\,, \tag{14}$$

$$Q=\lambda\cdot L\,, \tag{15}$$

where λ is line charge density, C is capacitance, U_{ad} is voltage between the electrodes and L is cylinder length. Combining the previous two expressions gives:

$$\lambda=\frac{C\cdot U_{ad}}{L}\,, \tag{16}$$

$$C=\frac{2\pi L}{\dfrac{1}{\varepsilon_1}\ln\dfrac{R_2}{R_1}+\dfrac{1}{\varepsilon_2}\ln\dfrac{R_3}{R_2}+\dfrac{1}{\varepsilon_1}\ln\dfrac{R_4}{R_3}}\,. \tag{17}$$

Inserting of Eq. (17) into Eq. (16) gives:

$$\lambda=\frac{2\pi U_{ad}}{\dfrac{1}{\varepsilon_1}\ln\dfrac{R_2}{R_1}+\dfrac{1}{\varepsilon_2}\ln\dfrac{R_3}{R_2}+\dfrac{1}{\varepsilon_1}\ln\dfrac{R_4}{R_3}}\,. \tag{18}$$

Eq. (18), together with Eq. (9), allows the determination of the electric field in the space between the inner and outer electrodes. Typically, at 10 kV/400 V distribution tower, on the 10 kV side, two glass insulators are used per phase (Fig. 1). Voltage distribution, along the insulator string (glass or porcelain), is not linear due to stray capacitances. Therefore, the voltage in a two-unit string is not equally shared by each unit. Nevertheless, given the comparison of the results obtained by the analytical and numerical approach, it is sufficient to choose a reasonable value of voltage as input given for both calculation types. Therefore, voltage of 5 kV was adopted for the field calculation in the insulator.

The properties of insulating materials relevant for the analysis are summarized in Tab. 1 [25-27]. The field calculation was performed for values U_{ad} = 5 kV, ε_{r1} = 15 (cement), ε_{r2} = 5 (glass), ε_{r3} = ε_{r1} = 15 (cement), ε_0 = 8.854×10^{-12} F/m. According to the previously derived expressions, the calculation of the electric field in the area of interest was performed using software package Mathcad 14 [28]. Software package Mathcad was used to automate the process of calculations and obtain graphical representations. Other programs, such as Matlab, can be used for this purpose [29].

The results of the calculation using analytical expressions are shown in Fig. 9 and Fig. 11.

Table 1 Properties of materials relevant for insulator modelling

Material	Relative permitivity ε_r (-)	Electric conductivity σ (S/m)
Air	1,0	10^{-13} - 10^{-18}
Glass	4 - 6,5	10^{-11} - 10^{-15}
Cement	2 - 30	10^{-3} - 10^{-4}
Steel (cast)	1,0	5×10^{6}

5 FINITE ELEMENT ANALYSIS

In the calculation of the electric field with the finite element method, the same material properties were used as in the calculation with analytical expressions. For this purpose, commercial software Ansys based on the FEM was used [30]. Within the Ansys software package, the Maxwell module and the electrostatic solver were selected. The segmentation of the computational domain was done by limiting the number of elements to 10,000 elements. Next, selected is an adaptive solver setup with a maximum number of pass equal to 20 and a percentage error of 0.1%. The 2D Ansys model of the disc insulator with two paths, along which the electric field is calculated in particular detail, is shown in Fig. 5. The results of the calculation with FEM are shown in Figs. 6-8 and Fig. 10.

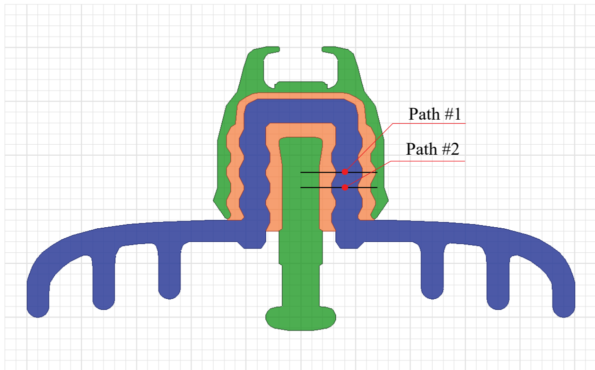


Figure 5 2D Ansys model of the disc insulator with two paths along which the electric field is calculated in particular detail

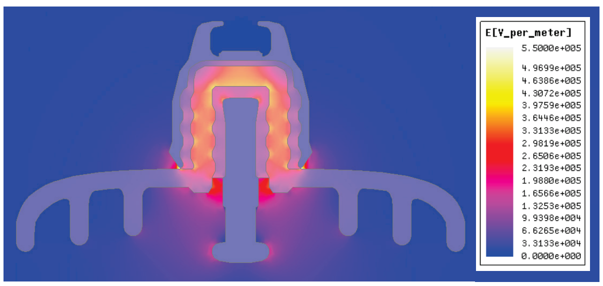


Figure 6 Distribution of the electric field inside and outside the disk insulator

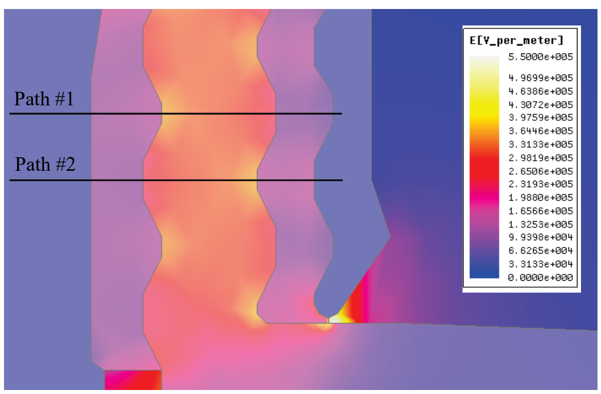


Figure 7 Distribution of the electric field inside and outside the disk insulator. Enlarged detail on the area between the metal pin and the metal cap

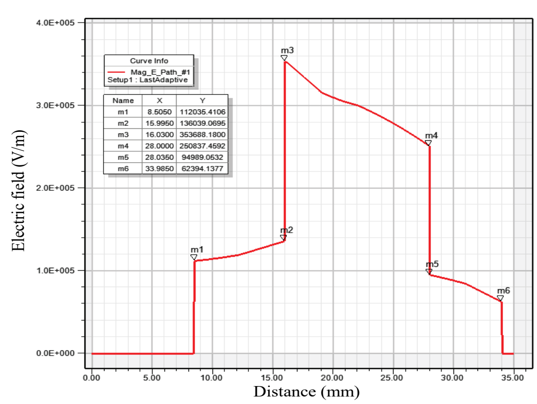


Figure 8 Electric field along path # 1 calculated by the finite element method

The electric field strengths at the boundary points between two media (insulating materials) determined by the finite element method and the analytical expressions are summarized in Tab. 2 and Tab. 3.

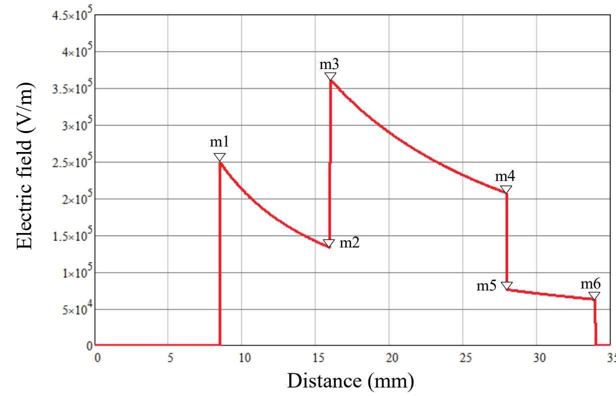


Figure 9 Electric field along path # 1 calculated by the analytical expressions

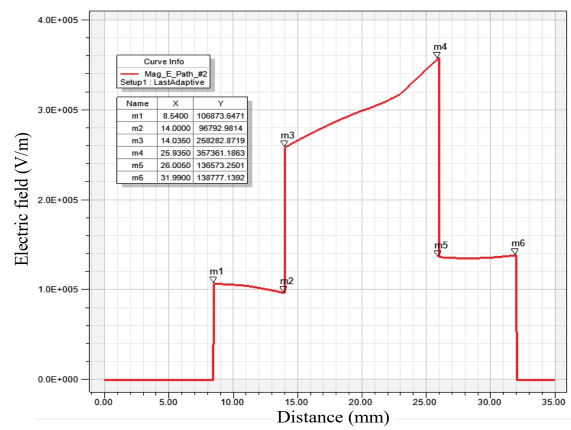


Figure 10 Electric field along path # 2 calculated by the finite element method

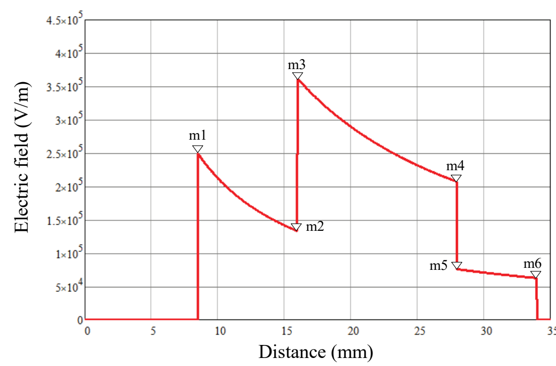


Figure 11 Electric field along path # 2 calculated by the analytical expressions

Table 2 Electric field along path # 1 at points of discontinuity *, **

Radial distance	Marker	Electric field (kV/m)	
		FEM	Analytic
R_{1+}	m1	112	250
R_{2-}	m2	136	134
R_{2+}	m3	354	362
R_{3-}	m4	251	208
R_{3+}	m5	95	76
R_{4-}	m6	62	63

* $R_1 = 8{,}5$ mm, $R_2 = 16$ mm, $R_3 = 28$ mm, $R_4 = 34$ mm

** $R_+ = \lim_{\Delta R \to 0}(R + \Delta R)$, $R_- = \lim_{\Delta R \to 0}(R - \Delta R)$

Table 3 Electric field along path # 2 at points of discontinuity *, **

Radial distance	Marker	Electric field (kV/m)	
		FEM	Analytic
R_{1+}	m1	107	246
R_{2-}	m2	97	150
R_{2+}	m3	258	407
R_{3-}	m4	357	220
R_{3+}	m5	137	80
R_{4-}	m6	139	65

* R_1 = 8,5 mm, R_2 = 14 mm, R_3 = 26 mm, R_4 = 32 mm

** $R_+ = \lim_{\Delta R \to 0}(R + \Delta R)$, $R_- = \lim_{\Delta R \to 0}(R - \Delta R)$

6 ANALYSIS OF RESULTS

When analyzing the results obtained by the analytical expressions and FEM, the results obtained by FEM are considered accurate and the results obtained by the analytical expressions are considered approximate. Theoretically, there are several physical phenomena that affect the distribution of the electric field in the space between the metal cap and the metal pin of the glass pane insulator. It is known from electrostatics that on the surface of conductive bodies the charge concentration is higher in areas with smaller radius of curvature such as edges, spikes, ends of elongated bodies, etc. In the case of disk insulators, the increased charge concentrations are found at both ends of the metal pin and at the opening of the metal cap (Fig. 6 and Fig. 7). For this reason, the analytical expressions presented are not valid (applicable) for calculating the electric field in the immediate vicinity of the end of the metal pin inside the metal cap, nor are they valid (applicable) for calculating the electric field in the immediate vicinity of the opening of the metal cap. In addition, there is another phenomenon that significantly affects the distribution of the electric field in the considered part of the space. In order to mechanically strengthen the connection of the metal cap with the rest of the insulator, the inner surface of the metal cap has a corrugated (fluted) cross-section. Therefore, the insulating materials (glass and cement) inside the cap have the same wavy pattern. Due to the change in the radius of curvature on the inside of the metal cap, the charge concentration is non-uniform along the inner surface of the metal cap. In convex areas, the charge concentration is higher than in concave areas. Therefore, the electric field strength along the inner surface of the metal cap is higher in convex regions than in concave regions.

According to the analytical expressions, the E-field may exhibit sudden (sharp) jumps at dielectric boundaries, but the E-field always decreases monotonically as the radial distance increases. Such a global trend of E-field behavior is disturbed by the local influence of non-uniform charge concentration on convex and concave zones of the inner surface of the metal cap. If the path, along which the E-field is analyzed, coincides with the concave zone of the inner surface of the metal cap, as in path #1, the waveforms of the E-field obtained by the finite element method and analytical expressions are similar (Fig. 8 and Fig. 9). When the path, along which the E-field is analyzed, coincides with the convex zone of the inner surface of the metal cap, as in path # 2, the waveforms of the E-field obtained by the finite element method and analytical expressions differ drastically (Fig. 10 and Fig. 11).

The numerical values of the E-field (Figs. 8 - 11, Tab. 2 and Tab. 3) show that the numerical values obtained with FEM and the analytical expressions are of the same order of magnitude. Although the analytical expressions provide approximate values, these values are physically meaningful. As such, they are useful for theoretical considerations and can be used in the initial stages of insulator design. However, for more detailed analysis and design, it is inevitable to use FEM.

7 RECAPITULATION ANNOTATION

Investigating the possibility of describing E-fields with simple analytical expressions, such as those used in coaxial capacitors to describe E-fields in HV glass disk insulators, can be done relatively easily at the analytical level. The space between the metal cap and the metal pin is represented as a complex dielectric. Analytical expressions for the electric field in these dielectrics are derived. The applicability of the presented expressions was verified by numerical calculations using the finite element method. In this work, the significant influence of convex and concave zones on the inner surface of the metal cap on the electric field distribution in the space between the metal pin and the metal cap was demonstrated. The work showed that it is possible to determine the approximate distribution of the electric field in the space between the metal pin and the metal cap of the glass insulator using the analytical expressions presented. However, the article also pointed to the limitations of the analytical expressions. Therefore, the presented analytical expressions have their theoretical meaning and can be useful in the initial stage of glass insulator design. The finite element method once again proved its superiority and proved to be an indispensable tool for the analysis and design of glass insulators.

Acknowledgments

This research is supported by the European Union from the European Regional Development Fund within the Operational Program "Competitiveness and Cohesion 2014-2020", under the project entitled Research and development of systems for monitoring high-voltage routes, and especially the discharge of high-voltage insulators using IoT technologies, under the project number KK.01.2.1.02.0056.

8 REFERENCES

[1] James, R. E. & Su, Q. (2008). *Condition Assessment of High Voltage Insulation in Power System Equipment*. The Institution of Engineering and Technology. https://doi.org/10.1049/PBPO053E

[2] Macey, R. E., Vosloo, W. L. & De Tourreil, C. (2004). *The practical guide to outdoor high voltage insulators.* Crown Publications.

[3] Kind, D. & Kärner, H. (1985). *High-Voltage Insulation Technology*. Springer Fachmedien Wiesbaden GmbH. https://doi.org/10.1007/978-3-663-14090-0

[4] Khan, M. J. (1980). *Computation of electric fields in and around high voltage insulators*. Master Thesis, Department of Electrical and Computer Engineering, University of Windsor.

[5] Rienen van U. (2001). *Numerical Methods in Computational Electrodynamics.* Springer, 1st edition. https://doi.org/10.1007/978-3-642-56802-2

[6] Schilders, W. H. A. & Maten, E. J. W. (2005). *Numerical Methods in Electromagnetics*. North Holland, 1st edition.

[7] Sadiku, M. N. O. (2000). *Numerical Techniques in Electromagnetics*. CRC Press, 2nd edition. https://doi.org/10.1201/9781420058277

[8] Mujezinović, A., Čarsimamović, A., Čarsimamović, S., Muharemović, A. & Turković, I. (2014). Electric field calculation around of overhead transmission lines in Bosnia and Herzegovina. *2014 International Symposium on Electromagnetic Compatibility*, Gothenburg, Sweden, 1001-1006. https://doi.org/10.1109/EMCEurope.2014.6931049

[9] Virbalis, J. A., Marciulionis, P., Lukocius, R., Deltuva, R. & Nedzinskaite, G. (2015). Analytical Expression of Electric Field Strength under Three–Phase Line. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, *4*(8), 7176-7182. https://doi.org/10.15662/ijareeie.2015.0408087

[10] Micu, D. D., Ceclan, A., Darabant, L. & Stet, D. (2009). Analytical and Numerical Development of the Electromagnetic Interference between a High-Voltage Power Line and a Metallic Underground Pipeline. *2009 8th International Symposium on Advanced Electromechanical Motion Systems & Electric Drives Joint Symposium*, Lillie, France, 1-6, https://doi.org/10.1109/ELECTROMOTION.2009.5259090

[11] El-Hawary, M. E. (2008). *Introduction to Electrical Power Systems.* Wiley-IEEE Press. https://doi.org/10.1002/9780470411377

[12] Malik, N. H., Al-Arainy, A. A., Qureshi, M. I. & Pazheri, F. R. (2011). Calculation of Electric Field Distribution at High Voltage Cable Terminations. *Journal of Energy Technologies and Policy*, *1*(2), 1-10. ISSN 2225-0573

[13] Christoph, J. & Clemens, M. (2019). Electric Field Model at Interfaces in High Voltage Cable Systems. *19th International Symposium on Electromagnetic Fields in Mechatronics, Electrical and Electronic Engineering (ISEF)*. https://doi.org/10.1109/ISEF45929.2019.9097007

[14] Soares, D. M., Mendonça, S., Neto, E. T. W. & Martinez, M. L. B. (2016). Electrical field on non-ceramic insulators and its relation to contact angles for constant volume droplets. *Journal of Electrostatics*, *84*, 97-105. https://doi.org/10.1016/j.elstat.2016.10.001

[15] Saleem, M. Z. & Akbar M. (2022). Review of the Performance of High-Voltage Composite Insulators. *Polymers, 14*(3), 431. https://doi.org/10.3390/polym14030431

[16] Pleşa, I., Noţingher, P., Schlögl, S., Sumereder, C. & Muhr, M. (2016). Properties of Polymer Composites Used in High-Voltage Applications. *Polymers, 8*(5), 173. https://doi.org/10.3390/polym8050173

[17] Sundararajan, R. & Gorur, R. S. (1994). Effect of insulator profiles on DC flashover voltage under polluted conditions. A study using a dynamic arc model. *IEEE Transactions on Dielectrics and Electrical Insulation, 1*(1), 124-132. https://doi.org/10.1109/94.300239

[18] Ramos, G. N., Campillo, M. T. R. & Naito, K. (1993). A study on the characteristics of various conductive contaminants accumulated on high voltage insulators. *IEEE Transactions on Power Delivery*, *8*(4) 1842-1850. https://doi.org/10.1109/61.248293

[19] Mizuno, Y., Kusada, H. & Naito, K. (1997). Effect of climatic conditions on contamination flashover voltage of insulators. *IEEE Transactions on Dielectrics and Electrical Insulation, 4*(3), 286-289. https://doi.org/10.1109/94.598284

[20] Sediver. (2021). Toughened glass insulators for HVAC applications. https://www.sediver.com

[21] JSC U.M.E.K. (2023). Toughened glass suspension insulator. https://umek.pro/products/catalogues

[22] Hayt, W. H. & Buck, J. A. (2001). *Engineering Electromagnetics*. McGraw-Hill Science/Engineering/Math, 6th edition.

[23] Haznadar, Z. & Štih, Ž. (2000). *Electromagnetic Fields, Waves, and Numerical Methods*. IOS Press.

[24] Sadiku, M. N. O. (2018). *Elements of Electromagnetics*. Oxford University Press, 7th edition.

[25] Uckol, H. I., Karaca, B. & Ilhan, S. (2020). DC and AC electric field analysis and experimental verification of a silicone rubber insulator. *Electrical Engineering, 102*, 503-514. https://doi.org/10.1007/s00202-020-00954-3

[26] Matmatch. (2023). Matmatch is a materials search platform. https://matmatch.com

[27] Totalmateria. (2023). Totalmateria is comprehensive materials Database. https://www.totalmateria.com

[28] Mathcad. (2023). Math Software for Engineering Calculations. https://www.mathcad.com/en

[29] MathWorks, Matlab. (2023). Matlab is a programming and numeric computing platform. https://www.mathworks.com/products/matlab.html

[30] Ansys. (2023). Engineering simulation and 3D design software. https://www.ansys.com

Authors' contacts:

Tomislav Barić, Full professor, PhD
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, University of Osijek,
Kneza Trpimira 2B, 31000 Osijek, Croatia
E-mail: tomislav.baric@ferit.hr

Hrvoje Glavaš, Full professor, PhD
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, University of Osijek,
Kneza Trpimira 2B, 31000 Osijek, Croatia
E-mail: hrvoje.glavas@ferit.hr

Željko Hederić, Full professor, PhD
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, University of Osijek,
Kneza Trpimira 2B, 31000 Osijek, Croatia
E-mail: zeljko.hederic@ferit.hr

Mirko Karakašić, Full professor, PhD
(Corresponding author)
University of Slavonski Brod, Mechanical Engineering Faculty in Slavonski Brod,
Trg Ivane Brlić-Mažuranić 2, 35000 Slavonski Brod, Croatia
E-mail: mirko.karakasic@unisb.hr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20230712131540

Original scientific paper
Received: 2023-07-12, Accepted: 2023-10-18

Probabilistic Analysis of Window Frame Junctions' Thermal Reliability Indicators to Precast Concrete and Brick Walls

Zeljko Kos*, Victor Pashynskyi, Mykola Pashynskyi, Stanislav Dzhyrma, Iryna Grynyova

Abstract: The side junctions of metal-plastic block frame windows to the walls of residential buildings at different positions of window frames in the wall cavity are analysed. The duration of thermal failure state by the condensate formation criterion was used as a probabilistic indicator of reliability. The units of uninsulated walls, erected in the second half of the 20th century, have a clearly insufficient thermal reliability level. Thermal insulation of walls according to current design standards in most cases ensures an acceptable of the thermal failure state duration.

Keywords: condensate; thermal failure state duration; window junctions

1 INTRODUCTION

A necessary condition for ensuring the comfort of buildings is a sufficient thermal reliability level of enclosing structures. A significant number of studies and regulations are devoted to solving this problem, which establish general criteria for ensuring thermal reliability [1, 2], performed climatic zoning of territories [3, 4], investigated the thermal characteristics of insulating materials and enclosing structures [5, 6], the optimal characteristics of thermal insulation are calculated [7-9] and heat losses are analysed in order to minimize them [10-12].

In [1, 13] it is shown that one of the dangerous types of thermal failures is the condensate formation in areas of high thermal conductivity. Condensation can form if the temperature of the inner surface falls below the dew point temperature [1, 2, 13]. This phenomenon can be prevented by sufficient insulation and the choice of a rational design scheme that eliminates areas of high thermal conductivity or reduce their impact.

A typical example of such zones of high thermal conductivity are the junctions of the windows to the walls. The study of temperature regime and heat loss depending on the location of windows in the cavity wall was performed in [13-18] by constructing a mathematical model of two-dimensional temperature fields. In these works, it is shown that shifting the window inside can significantly increase the internal jamb temperature and thus reduce the risk of condensation. This applies both to windows in the walls of modern structures [13, 14] and to old buildings that are subject to thermal modernization. [17, 18]. In the study [13], based on the data of temperature fields in seven junctions of window frames to walls of different construction, it is shown that high-quality wall insulation drastically increases the temperature of window jamb. The block frame window position in a cavity of the sufficiently insulated wall affects the temperature in critical areas of the nodes slightly.

The main disadvantage of the studies above is the use of fixed values of the design parameters without considering their random nature. To adequately evaluate the possibility of thermal failure based on the criterion of condensation formation, it is worth to consider random nature of the both outside and the inside air temperatures, temperatures of inner surface and the dew point. Probabilistic methods are widely used for evaluating the reliability level by the criterion of loss strength of load-bearing structures [19-22].

A much smaller number of studies are devoted to the probabilistic approach of assessing the thermal reliability level of enclosing structures. In particular, in [1] the general requirements for thermal reliability are stated. Article [22] is devoted to the assessment of the thermal failures probability of enclosing structures that are made of lightweight expanded polystyrene concrete with a frame made of bent steel profiles. In article [23] it is proposed to use the probability level of failure and the possible absolute (in units of time) or relative (fraction of the service life) thermal failure state duration as a thermal reliability indicator. The principles of estimating the thermal failure durations according to the comfort criteria and condensate formation in areas of high thermal conductivity are substantiated there.

This study aim is to develop a practical method that allows to estimate the thermal failure duration by the condensate formation criterion and to analyse the thermal reliability of the window frame nodes to walls of residential and public buildings taking into account real climatic conditions and variable position of this windows in a wall cavity.

2 RESEARCH OBJECTS

The side junctions of metal-plastic windows of different structures are analysed, the schemes of which are shown in Fig. 1 according to the article [13]. Units of "a" and "c" types correspond to typical brick and panel walls of residential buildings of the last century, in which worn-out wooden windows are replaced by modern metal-plastic block frame windows with 70 mm thick elements and double-glazed frames. Units of "b" and "d" types are formed as a result of thermal modernization of such buildings by performing facade insulation. Units of types "e", "f" and "g" are typical for bearing and self-bearing walls of modern residential buildings that are insulated according with the requirements of the DBN [2]. Sealing of the wall junctions is performed according with [24]. Thermal insulation of walls is made with mineral wool or expanded polystyrene insulation slabs, which have very close coefficients of thermal conductivity. The thickness of the thermal insulation layer provides a heat transfer resistance of not less than 3.3 $m^2 \cdot K/W$, established by the standards [2] for the first temperature zone of Ukraine. Thermophysical characteristics of all materials that are shown in Fig. 1 are taken according to the standard [25].

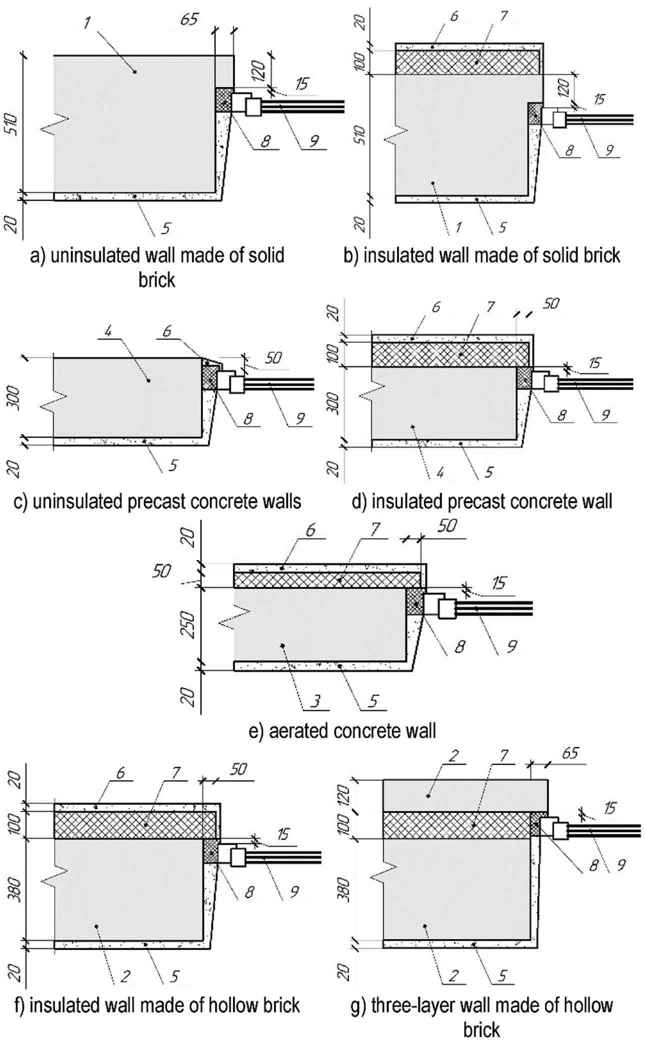


Figure 1 Schemes of block frame windows side junctions to the walls
1 – solid brick; 2 - hollow brick; 3 - aerated concrete blocks; 4 - expanded clay concrete wall; 5 - lime-sand mortar plaster; 6 - cement-based plaster; 7 - mineral wool insulation; 8 - polyurethane foam; 9 - triple-glazed window.

Fig. 1 shows the diagrams of the block frame windows junctions to walls along their thickness. In order to find a rational position of the windows, each of these junctions is analysed by different shift of window inside the building in relation to their normal position. The cavities formed on external jambs as a result of such shift are filled with mineral wool or expanded polystyrene which are protected from atmospheric influences by external plaster based on cement mortar along a reinforcing mesh.

3 TEMPERATURE REGIME OF NODES

The temperature regime of the units shown in Fig. 1 was investigated in our work [13]. The critical zones of all nodes are the junction points of the internal jamb to the block frame window, where the inside surface of the wall has the lowest temperature. Temperature fields of nodes at different positions of windows are constructed by computer simulation using the THERM program. The dimensions of the simulated nodes correspond to Fig. 1, and the thermophysical properties of the materials were taken as in [25]. The temperature of external atmospheric air is considered as –24 °C and the temperature of air indoors as +20 °C.

Table 1 The results of estimating the duration of thermal failures

Wall type	x, mm	ϑ_{cr}, °C	R_{ef} $m^2 \cdot K/W$	Duration of thermal failure (h/year)		
				Kropyvnytskyi	Krivoy Rog	Semenivka
a) uninsulated wall made of solid brick	0	7,7	0,411	334,3	289,8	483,9
	30	8,4	0,436	239,7	206,0	355,6
	60	8,9	0,456	183,8	157,1	277,5
	90	9,4	0,477	137,4	116,8	211,1
	120	9,7	0,491	113,9	96,5	176,8
	150	10,3	0,521	75,9	63,9	120,1
b) insulated wall made of solid brick	0	10,7	0,544	56,5	47,4	90,5
	30	11,2	0,575	37,9	31,7	61,6
	60	11,4	0,588	32,1	26,8	52,3
	90	11,8	0,617	22,5	18,7	37,0
	120	11,9	0,624	20,5	17,1	33,8
	150	12,3	0,657	14,0	11,7	23,2
c) uninsulated precast concrete wall	0	8,4	0,436	239,7	206,0	355,6
	25	8,9	0,456	183,8	157,1	277,5
	50	9,5	0,482	129,2	109,7	199,2
	75	9,5	0,482	129,2	109,7	199,2
	100	9,6	0,486	121,4	102,9	187,8
e) aerated concrete wall	0	12,8	0,702	8,42	7,00	14,01
	25	12,7	0,693	9,35	7,77	15,55
	50	12,7	0,693	9,35	7,77	15,55
	75	13,2	0,744	5,45	4,53	9,08
	100	13,1	0,733	6,09	5,06	10,15
f) insulated wall made of hollow brick	0	13,1	0,733	6,09	5,06	10,15
	25	13,0	0,722	6,80	5,65	11,32
	50	13,0	0,722	6,80	5,65	11,32
	75	13,1	0,733	6,09	5,06	10,15
	100	13,0	0,722	6,80	5,65	11,32
	125	13,2	0,744	5,45	4,53	9,08
	150	13,1	0,733	6,09	5,06	10,15
g) three-layer wall made of hollow brick	0	13,4	0,766	4,35	3,62	7,24
	25	13,4	0,766	4,35	3,62	7,24
	50	13,5	0,778	3,88	3,23	6,45
	75	13,5	0,778	3,88	3,23	6,45
	100	13,4	0,766	4,35	3,62	7,24
	125	13,4	0,766	4,35	3,62	7,24
	150	13,5	0,778	3,88	3,23	6,45

According to calculations performed in [13] the critical zone ϑ_{cr} temperatures are obtained for all nodes at different values of displacement x of the window from the normal position, shown in Fig. 1. The results of the study [13] are shown in Tab. 1. Taking into account the architectural and ergonomic requirements, as well as the design of the wall, the shift of the window varied within x = 0-150 mm. Tab. 1 shows that for walls of types "a", "b" and "c", ie walls made of solid brick and uninsulated precast concrete walls, the shift of the block window inside the building results in a considerable increase in the critical area temperature of the node. This effect is little expressed or even practically absent for nodes of other types.

4 METHOD OF THE PROBABILISTIC ESTIMATION OF THE THERMAL RELIABILITY LEVEL

The method is based on the representation of outdoor and indoor air temperature in the form of sequences of random variables with a normal distribution law for 12 months of the year. It also based on introduced by the authors concept of heat transfer resistance [23], which connects outdoor air temperature and a specified point of the enclosure. Statistical characteristics of outside air temperature can be established according to the data of works [4, 26], or by processing the results of meteorological observations.

As a numerical indicator of the thermal reliability, it is proposed in [23] to use the probable thermal failure state duration. This value is a physically understandable indicator that allows to compare different design solutions of junctions and to make conclusions about their suitability for operation. Principles and general methods for estimating the probable thermal failure state duration of enclosing structures using various criteria are proposed in the work of the authors [23]. In paper, these methods are adapted for estimation of the probable the thermal failure state duration by the condensate formation criterion in the zone of increased thermal conductivity. The basic thermal characteristic of the window side junction to the wall is the conditional value of heat transfer resistance R_{ef} for the suggested critical zone of the inside window jamb surface proposed in [23]

$$R_{ef} = \frac{\vartheta_{in} - \vartheta_{out}}{\alpha_{in}\left(\vartheta_{in} - \vartheta_{cr}\right)} \tag{1}$$

where ϑ_{out} and ϑ_{in} – outdoor and indoor air temperatures; α_{in} – internal heat transfer coefficient.

Random variable has expected value and the value of standard deviation of the critical zone temperature of the node in the i^{th} month that is calculated by the formulas obtained in [23]. These formulas are based on the method of linearization of functions of random variables [27]:

$$\begin{aligned} M_{cr,\,i} &= \frac{1}{R_{ef}\cdot\alpha_{in}}\left[M_{in}\left(\alpha_{in}\cdot R_{ef}-1\right)+M_{out,\,i}\right] \\ S_{cr,\,i} &= \frac{1}{R_{ef}\cdot\alpha_{in}}\sqrt{S_{out,\,i}^{2}+S_{in}^{2}\left(\alpha_{in}\cdot R_{ef}-1\right)^{2}} \end{aligned} \tag{2}$$

where R_{ef} – conditional heat transfer resistance of the node; M_{in} and S_{in} – expected value and the standard deviation of a random variable of indoor air temperature; $M_{out,i}$ and $S_{out,i}$ – expected value and the standard deviation of a random variable of outdoor air temperature in the i^{th} month.

Failure by the condensate formation criterion occurs due to temperature drop of the inner surface ϑ_{cr} below the dew point temperature ϑ_{dew}, so the reliability condition can be written in the form of an inequality

$$\vartheta_{cr} - \vartheta_{dew} \geq 0 \tag{3}$$

Considering the random nature of both temperatures, the expected value and the standard deviation in the i^{th} month in (3) are as follows:

$$M_i = M_{cr,\,i} - M_{dew}\,;\; S_i = \sqrt{S_{cr,\,i}^2 + S_{dew}^2} \tag{4}$$

where M_{cr} and S_{cr} – expected value and the standard deviation of random value of internal surface temperature of the wall according to (2); M_{dew} and S_{dew} – expected value and the standard deviation of random dew point temperature.

The presence of statistical characteristics (4) allows us to estimate the thermal failure state duration due to the probability of a random value of the thermal reliability reserve (3) below zero. The absolute and relative thermal failure state duration for one year are equal

$$Q_{abs} = \sum_{i=1}^{12}\left[\tau_i\; F_i(0)\right];\; Q_{rel} = \frac{1}{12}\sum_{i=1}^{12} F_i(0) \tag{5}$$

where $F_i(0)$ – normal distribution function of thermal reliability reserve for the i^{th} month; τ_i – duration of the i^{th} month.

The absolute duration of the thermal failure state during the year Q_{abs} is expressed in units of time, in which to the formula (5) durations of months τ_i are substituted. The relative thermal failure state duration Q_{rel} is a part of the enclosing structure service life, during which the thermal failure of the unit is occurred. Essentially, the values Q_{abs} and Q_{rel} determine the time during which the formation of condensate on the internal window jamb surface may occur.

5 THE RESULTS OF ASSESSING THE THERMAL RELIABILITY LEVEL OF THE NODES

The side junctions of block frame windows to the walls of different structures, which are shown in Fig. 1 and listed in Tab. 1, were analysed. For each node, the table shows the conditional heat transfer resistance for the critical zone of the node Ref by Eq. (1) and the absolute thermal failure state duration by the condensate formation criterion Q_{abs} in hours per year. These values are calculated by Eq. (5) for different positions of the windows, with the value of the shift x. Climatic conditions of the cities of Kropyvnytskyi, Kryvyi Rih and Semenivka of Chernihiv region were taken into account. The cities of Kryvyi Rih and Semenivka are selected as areas with the highest and lowest average annual air temperature within the first temperature zone of Ukraine according to the DBN map [2].

Expected values $M_{out,i}$ and standard deviations $S_{out,i}$ of the atmospheric air temperature in each month for the three cities of Ukraine are taken according to [4, 26]. The indoor air temperature is also a random variable, its statistical characteristics are established by the results of field observations in several dwellings during the heat season. The generalization of these data performed in [23], taking into account the regulatory requirements [2] for the temperature in dwellings allowed to take M_{in} = 20 °C, S_{in} = 0,6 °C.

The dew point temperature ϑ_{dew} at deterministic values of air temperature and relative humidity is determined by psychometric tables. Random variability of microclimate parameters in the room determines the random nature of the dew point temperature. Taking into account the possible variability of these parameters, the expected value and the standard deviation of the random dew point temperature in residential premises were established in [23] as M_{dew} = 10,6 °C, S_{dew} = 1,7 °C, which were taken into account when estimating the thermal failures duration of the considered nodes.

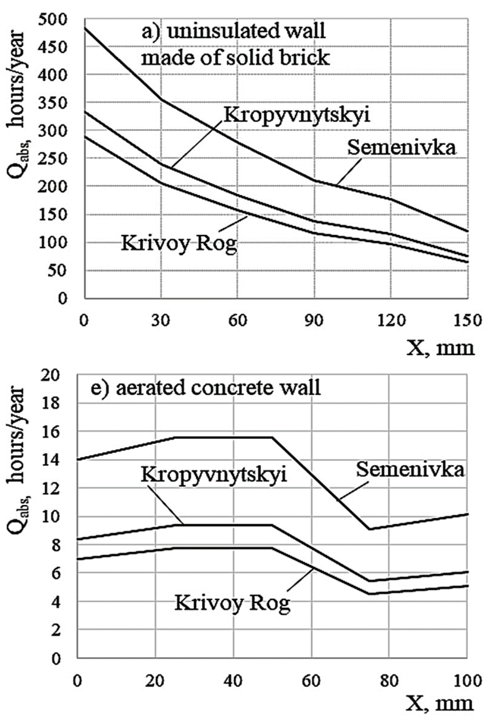


Figure 2 Dependences of the thermal failure state duration of the nodes on the position of the window frames.

As examples according to table 1 the dependences of the absolute thermal failure state duration on the amount of displacement of the window unit for nodes of type "a" and "e" are shown in Fig. 2. The figure shows that the the thermal failure state duration of nodes "a" in the initial state Q_{abs} = 290-484 hours/year is too high. Shifting the block frame window inside the building reduces the duration of the thermal failure state to about Q_{abs} = 100-150 hours per year, but it is also too high and will adversely affect the operation state of junction. The use of modern walls from aerated concrete with front insulation (node "e") reduces the duration of a thermal failure state to Q_{abs} = 4-16 h/year. The position of the window has less effect on the reliability of the node, although from Fig. 2 it is possible to determine the optimal shift value x = 80-100 mm.

The thermal failure state duration of other nodes in the first temperature zone [2] are shown in table 1. Similarly with the temperature of critical areas changes, nodes can be divided into two types. The reliability level of nodes "a", "b", "c" essentially increases with the shift of the window inside the wall cavity. As for nodes of other types, the displacement of the window has a small and unsystematic effect on the thermal failure state duration by the condensate formation criterion. Table 1 shows a significant variation in the thermal failure duration within the first temperature zone. During the transition from the city of Kryvyi Rih with the warmest climate to the city of Semenivka with the lowest air temperatures in the first zone, the duration of thermal failure increases by 1.7-2.0 times.

In the works of the authors based on the thermal reliability level analysis of the walls that meet the requirements of current regulations [2] on the inadmissibility of condensate on the inner surface, it is proposed to set the maximum allowable value of the relative duration of thermal failures equal to Q_{rel} = 0,005 for walls of residential and public buildings with a class of responsibility CC2 and Q_{rel} = 0,001 – for buildings with a class of responsibility CC3. The corresponding values of the absolute thermal failure state duration Q_{abs} = 43,8 hours/year and Q_{abs} = 8,8 hours/year. The table 1 data shows that in uninsulated brick and precast concrete walls of types "a" and "c" the duration of the thermal failure state always exceeds the recommended limits. Facade insulation of a brick wall (type "b" node) combined with the window shift inside the building ensures the implementation of these recommendations for buildings with the class of responsibility CC2. Nodes of type "d" in the insulated precast concrete walls always correspond to the given recommendations for buildings of a class of responsibility CC2, and at some positions of the block frame windows even correspond to the requirements of a class of responsibility CC3. The modern walls with nodes of types "e", "f" and "g" mainly fulfil the requirements for buildings of the responsibility class CC3. The exception is the city of Semenivka with a colder climate.

Recommendations on the maximum permitted value of the relative length of thermal failures based on the condensate formation criterion in the heat-conducting zones are obtained in accordance with [2] based on the analysis of enclosing structures with design values of air temperature –22 °C and –19 °C in the first and second temperature zones of Ukraine. The duration of the possible condensate formation period equal to 43.8 h/year is too long to ensure the normal operation of buildings. This encourages a possible revision of regulatory requirements, in particular the specified design values of air temperature.

6 CONCLUSIONS

1) The absolute and relative thermal failure state duration is a very convenient, scientifically proven and practically useful indicator of thermal reliability, which allows to compare different design solutions of enclosing

structures junctions by the condensate formation criterion.

2) The absence of facade insulation leads to the fact that the nodes of the side junction of windows to the brick and precast concrete walls are characterized by too large values of the thermal failure state duration, which makes it impossible for normal operation of buildings. n the vast majority of cases, insulation of existing or new walls in compliance with current design standards in Ukraine reduces the thermal failure state to 3... 30 hours per year.
3) The thermal failure state duration of the considered nodes by the condensate formation criterion may differ twice within the first climatic zone of Ukraine.
4) The existing recommendations on the maximum allowable value of the thermal failure state duration according to the condensate formation criterion in the heat-conducting zones should be revised and changed in the direction of reduction.

7 REFERENCES

[1] Farenyuk, G. (2019). The determination of the thermal reliability criterion for building envelope structures. *Tehnički glasnik, 13*(2), 129-133. https://doi.org/10.31803/tg-20181123111226

[2] DBN B.2.6-31: 2016: Thermal insulation of buildings. *Kyiv: Ministry of Construction of Ukraine, 2016,* 31 p. (in Ukrainian)

[3] Tsikaloudaki, K., Laskos, K. & Bikas, D. (2012). On the Establishment of Climatic Zones in Europe with Regard to the Energy Performance of Buildings. *Energies, 5*(1), 32-44. https://doi.org/10.3390/en5010032

[4] Pashynskyi, V., Pashynskyi, M., Pushkar, N. & Skrynnik, I. (2019). Method of administrative-territorial zoning of the design parameters of air temperature. *Electronic Journal of the Faculty of Civil Engineering Osijek-e-GFOS, 19*, 50-57, https://doi.org/10.13167/2019.19.5

[5] Son, H.-J., Ko, A.-R. & Lee, W.-K. (2015). Characteristics of Thermal Conductivity for Insulating Brick Made from Converter Slag. *Journal of Korea Society of Waste Management. 32*, 378-382. https://doi.org/10.9786/kswm.2015.32.4.378

[6] Košir, M., Pajek, L., Hudobivnik, B., Dovjak, M., Iglič, N., Bozicek, D. & Kunič, R. (2017). Non-Stationary Thermal Performance Evaluation of External Façade Walls under Central European Summer Conditions. *ISES Conference Proceedings*, 1-10. https://doi.org/10.18086/swc.2017.15.03

[7] Barrau, J., Ibanez, M. & Badia, F. (2014). Impact of the insulation materials' features on the determination of optimum insulation thickness. *International Journal of Energy and Environmental Engineering, 5*(2-3), 79. https://doi.org/10.1007/s40095-014-0079-3

[8] Fertelli, A. (2013). Determination of optimum insulation thickness for different building walls in Turkey. *Transactions of FAMENA, 37*(2), 103-113.

[9] Kaynakli, Ö. & Kaynakli, F. (2016). Determination of Optimum Thermal Insulation Thicknesses for External Walls Considering the Heating, Cooling and Annual Energy Requirements. *Uludağ University Journal of the Faculty of Engineering, 21*, 227-241. https://doi.org/10.17482/uujfe.27323

[10] Scartezzini, J.-L., Duman, Ö., Koca, A., Acet, R., Çetin, M. & Gemici, Z. (2015). A study on optimum insulation thickness in walls and energy savings based on degree day approach for 3 different demo-sites in Europe. *Future Buildings and Districts Sustainability from Nano to Urban Scale*, 155-160. https://doi.org/10.5075/epfl-cisbat2015-155-160

[11] Ekström, T. & Blomsterberg, E. (2016). Renovation of Swedish single-family houses to passive house standard - Analyses of energy savings potential. *Energy Procedia, 96*, 134-145. https://doi.org/10.1016/j.egypro.2016.09.115

[12] Ahmed, K., Carlier, M., Feldmann, C. & Kurnitski, J. (2018). A New Method for Contrasting Energy Performance and Near-Zero Energy Building Requirements in Different Climates and Countries. *Energies, MDPI, Open Access Journal, 11*(6), 1-22. https://doi.org/10.3390/en11061334

[13] Pashynskyi, V. A., Dzhyrma, S. O. & Pashynskyi, M. V. (2020). Thermal characteristics of window junctions to brick and reinforced concrete walls of civil buildings in the Kirovohrad region. *Central Ukrainian Scientific Bulletin. Technical sciences, 3*(34), 200-209. (in Ukrainian) https://doi.org/10.32515/2664-262X.2020.3(34).200-209

[14] Stolarska, A., Strzałkowski, J., & Garbalińska, H. (2018). Using CFD software for the evaluation of hygrothermal conditions at wall-window perimeters. *IOP Conf. Series: Materials Science and Engineering, Vol. 415*. https://doi.org/10.1088/1757-899X/415/1/012046

[15] Sierra, F., Gething, B., Bai, J. & Maksoud, T. (2017). Impact of the position of the window in the reveal of a cavity wall on the heat loss and the internal surface temperature of the head of an opening with a steel lintel. *Energy and Buildings, 142*, 23-30. https://doi.org/10.1016/j.enbuild.2017.02.037

16. Azmy, N. Y. & Ashmawy, R. E. (2018). Effect of the Window Position in the Building Envelope on Energy Consumption. *International Journal of Engineering & Technology, 7*(3), 1861-1868. https://doi.org/10.14419/ijet.v7i3.11174

[17] Kariuk, A., Rubel, V., Pashynskyi, V. & Dzhyrma, S. (2020). Improvement of Residential Buildings Walls Operation Thermal Mode. *Proceedings of the 2nd International Conference on Building Innovations. ICBI 2019. Lecture Notes in Civil Engineering, vol. 73,* 75-81. https://doi.org/10.1007/978-3-030-42939-3_9

[18] Pashynskyi, M., Dzhyrma, S., Pashynskyi, V. & Nastoyashchiy, V. (2020). Providing the thermal reliability of window junctions during the thermal modernization of civil buildings. *Electronic Journal of the Faculty of Civil Engineering Osijek-e-GFOS, 21*, 45-54. https://doi.org/10.13167/2020.21

[19] Pichugin, S. F. (2018). Reliability estimation of industrial building structures. *Magazine of Civil Engineering, 83*(7), 24-37. https://doi.org/10.18720/MCE.83.3

[20] Chybowski, L. & Zolkiewski, S. (2015). Basic Reliability Structures of Complex Technical Systems. *Advances in Intelligent Systems and Computing, 354*, 333-342. https://doi.org/10.1007/978-3-319-16528-8_31

[21] Pownuk, A. (2001). Calculation of reliability of structures using random sets. *AIMECH 2001, Symposium on Methods of Artificial Intelligence in Mechanics and Mechanical Engineering, Gliwice, 14-16 Nov.,* 217-220.

[22] Semko, V., Pichugin, S. F. & Leshchenko, M. (2017). Probabilistic analysis of thermal performance of the wall from light-gauge thin-walled steel structures. *Academic Journal. Series: Industrial Machine Building, Civil Engineering, 48*. 144-155. https://doi.org/10.26906/znp.2017.48.788

[23] Pashynskyi, V. A. & Plotnikov, O. A. (2017). Probabilistic method for analyzing thermal reliability of units of enclosing structures. *Bulletin of the Belarusian-Russian University, 3*(56), 129-135.

[24] DSTU B V.2.6-79: 2009. Constructions of houses and buildings. Seams connecting adjacencies of windows to the walls. General technical conditions. - [Effective from 01.04.2010]. K*.: Ministry of Regional Development of Ukraine, 2009*, - 26 p. (in Ukrainian)

[25] DSTU B V.2.6-189: 2013. Methods of heat-insulating material selection for warming of buildings. *K.: Ministry of Regional Development of Ukraine, 2014,* 51 p. (in Ukrainian)

[26] Kariuk, A. (2018). The Statistical Characteristics and Calculated Values for Air Temperature in Building's Cladding Design. *International Journal of Engineering & Technology, 7*(3), 608-613. https://doi.org/10.14419/ijet.v7i3.2.14600

[27] Gatti, P. L. (2005). Probability Theory and Mathematical Statistics for Engineers. *1st Edition, London, UK: CRC Press*. https://doi.org/10.1201/9781482267761

Authors' contacts:

Zeljko Kos, PhD, Associate Professor
(Corresponding author)
University North,
42 000 Varaždin, 104. brigade 3, Croatia
+38598757989 / zeljko.kos@unin.hr

Victor Pashynskyi, Science Doctor, Professor
Central Ukrainian National Technical University,
25006, Kropyvnytskyi, 8, Prospekt Universytetskyi, Ukraine
+380963652812 / pva.kntu@gmail.com

Mykola Pashynskyi, PhD, Associate Professor
Central Ukrainian National Technical University,
25006, Kropyvnytskyi, 8, Prospekt Universytetskyi, Ukraine
+380501648778 / mykola.pashynskyi@gmail.com

Stanislav Dzhyrma, PhD, Associate Professor
Central Ukrainian National Technical University,
25006, Kropyvnytskyi, 8, Prospekt Universytetskyi, Ukraine
+380953995883 / stas55871@ukr.net

Iryna Grynyova, PhD, Associate Professor
Odesa State Academy of Civil Engineering and Architecture,
65029, Didrikhsona St, 4, Odesa, Odessa Oblast, Ukraine
+380939799301 / irene.grinyova@gmail.com

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20231212053424

Original scientific paper
Received: 2023-12-12, Accepted: 2024-06-07

A Study of Anticipatory Failure Determination (AFD) based on Scenario Analysis Methods

Seock-Bum Yun, Yong-Won Song*, Young-Pyo Sue, Sang-Jun Lee

Abstract: Theory of Inventive Problem Solving (TRIZ) is a methodological algorithm that helps solve complex problems by generating original ideas. The Anticipatory Failure Determination (AFD) method is one of the problem-solving methodologies within TRIZ. It is a creative way to solve problems when you cannot find the cause of a problem or observe the situation in which it occurs. This study aimed to investigate the effectiveness of the Anticipatory Failure Determination (AFD) method, which is one of the problem-solving methodologies within the Theory of Inventive Problem Solving (TRIZ). TRIZ is a methodological algorithm designed to address complex problems by stimulating the generation of original ideas. It involves a shift in thinking: instead of looking for the cause of a problem (i.e., "why does it happen?"), it looks for ideas to solve the problem (i.e., "how can I make it happen?") by formulating a hypothesis and testing the hypothesis. In doing so, they explore and leverage resources that exist within or adjacent to the system to support the failure hypothesis. A resource is anything that is used to fulfill the ideal final result. In this study, scenario analysis, a methodology synonymous with foresight, is integrated into the AFD framework to enhance problem-solving capabilities. By incorporating scenario analysis, the formulation of failure hypotheses crucial to the AFD process is facilitated. Additionally, a case study demonstrating the application of AFD augmented by scenario analysis is presented, showcasing its efficacy in resolving complex problems. This integration aims to proactively mitigate future risks and prevent potential damages. Through this research, a comprehensive understanding of how scenario analysis enriches the AFD methodology is provided, empowering practitioners to pre-emptively address emerging challenges. Ultimately, the goal is to fortify problem-solving capacities and foster resilience in the face of evolving complexities.

Keywords: Anticipatory Failure Determination (AFD); Inversion Analysis; Scenario Method; TRIZ

1 INTRODUCTION

The recent experience of the COVID-19 pandemic has made the world acutely aware of the need to prepare for changes resulting from uncertainty. A number of unexpected events, such as the territorial war between Russia and Ukraine and the bankruptcy of global financial firms, have led to a growing interest in predicting the future in various fields. It is impossible to predict the future with any accuracy. However, it can be said that forecasting is not a prophetic statement, but a process of clearly identifying the current situation and presenting a highly probable realisation. This simply means that future research is about predicting the future and preparing to respond flexibly, quickly and wisely to situations that are likely to arise from the current situation.

The future is full of possibilities and opportunities, but it is also full of uncertainties and vagaries, and it would be foolish to face it without countermeasures and preparations, because it always contains risks and failures. To manage future failures and risks, it is important to understand the current situation and examine possible scenarios, taking into account trends in change. Seeing what could happen in the future and preparing for failure is a great way to reduce risk. As a way of managing failure and risk, this study investigated a method of predictive failure assessment based on scenario analysis. Scenario analysis, a concept first used during the Second World War, is a methodology for predicting the future by forming hypotheses to estimate the development of different possible alternatives that may emerge in the future. It is characterised by the fact that it does not simply predict intuitively, but derives different situations by reflecting current situations and trends, rather than making a definitive prediction. In essence, anticipating failure is not just about avoiding negative outcomes; it's about fostering resilience, innovation, and continuous improvement, ultimately enabling organizations to thrive in dynamic and uncertain environments.

Anticipatory Failure Determination (AFD) is a TRIZ (Theory of Inventive Problem Solving) inversion analysis method. Inversion analysis is the redefinition of a problem into a creative problem. It generates new ideas or solutions by reversing the function of a system or component. Inversion analysis can be traced back to the psychological phenomenon of denial, where people refuse to think about unpleasant and negative things, thus rejecting the causes of problems and failures. However, approaching a problem through the lens of change analysis makes it easier to uncover possible causes of failure and unpleasantness that are not found by more conventional methods. Failure Analysis and Predictive Failure Detection share the ability to generate innovative solutions and improve design. Combining the concepts of both methods can help develop creative solutions, mitigate future risks and provide a broader approach to problem solving.

Inversion analysis in TRIZ is used as a methodology for product development or technical problem solving, but failure prediction based on the scenario analysis method can reduce potential risks and prepare countermeasures. Inversion analysis in TRIZ is a problem-solving methodology that aims to prevent problems from occurring by identifying the cause of failure in advance, or when it is difficult to identify the root cause of a particular problem. AFD aims to prevent or solve problems by predicting the situation in which failure will occur in advance, by establishing and verifying hypotheses about "how the normal process of the system will fail" through a change in thinking. In the process of generating hypotheses about possible failures, scenario analysis, one of the foresight methods, is used to facilitate the discovery of more accurate, rational and consistent failure hypotheses. In addition, to validate the hypothesis, the scenario analysis process should consider whether there are resources within the system that can cause the failure to occur, to ensure reasonableness and plausibility when using the scenario analysis method. AFD is not limited to technical issues but can be applied in many areas such as corporate business and strategic planning, politics, economics, international energy issues, international political forecasting and human resource management issues. In particular, it is suggested that the worst-case scenario setting method implemented in the scenario analysis method can be applied to the predictive failure decision method to prepare optimal responses to situations that would cause a lot of risk if they occurred in reality.

This study aimed to investigate the theoretical system of the predictive failure decision method based on the scenario analysis method for achieving ideality through resource utilisation as a problem-solving method that can be used in various fields, not only in product development or technical problem solving. In addition, the case of the anticipatory failure determination method based on scenario analysis was examined together with the theoretical previous studies.

2 LITERATURE REVIEW

2.1 Anticipatory Failure Determination of Theory of Inventive Problem Solving (TRIZ)

TRIZ is a problem-solving methodology based on logic and data. It is a powerful tool that provides repeatability, predictability and reliability in problem solving through a structured, algorithmic approach to problem solving [1, 2]. The Russian acronym for "theory of inventive problem solving", developed by the Soviet scientist G. S. Altshuller argues that anyone can develop creativity by thinking in terms of objective laws. TRIZ is a methodology based on the study of hundreds of thousands of patents to provide an objective method for problem-solving and has empirical objectivity as a tool for using knowledge.

Classical TRIZ was first introduced in the former Soviet Union, where an early system of problem solving methods was built and developed. Its tools include the ARIZ algorithm, the concept of resources based on matter field analysis, 40 principles of invention, contradiction, separation and eight evolutionary patterns, and many other methods. He also introduced the concept of ideality as a way of thinking that clarifies the purpose of problem solving. Classical TRIZ was further developed by interested scholars after the establishment of the Technical School in Kishinev. The more systematic development of TRIZ methodology during this period included the expansion and systematisation of the "resource" approach and the introduction of the concept of transition analysis in Anticipatory Failure Determination (AFD).

AFD is based on TRIZ, the Theory of Inventive Problem Solving developed by the Russian Altschuler [3], and on the concepts of Subversion Analysis or Sabotage Analysis in R&D problem solving methods developed by Voloslav Mitrofanov [4]. AFD is a problem-solving methodology proposed by Boris Zlotin, an inventor and scientist who was a student of Voloslav Mitrofanov. Boris Zlotin developed the methodology of TRIZ in the 1970s. In 1992, together with patent attorney Alla Zusman and experts from the United States, he founded Ideation International Inc. and ushered in the era of I-TRIZ. During this time, he developed Predictive Failure Determination (AFD), which is a reversal of the traditional TRIZ problem solving algorithm and is useful for failure analysis and prediction. Along with the development of the methodology, he also developed software as an analysis tool.

Predictive Failure Analysis can be used for problem solving when the root cause of a problem cannot be determined. Some problems are too difficult to analyse for their causes to use TRIZ or ARIZ troubleshooting algorithms. In these cases, inversion analysis can be used. In particular, it can be used when the nature of the problem makes it impossible to observe how the problem occurs. The current method of AFD is classified by I-TRIZ into the concepts of Failure Analysis and Failure Prediction and applied to problem solving.

Failure analysis is applied to identify causes and find solutions to problems that have already occurred. It is a systematic procedure for determining the root cause of a failure or undesired phenomenon in a system and correcting it. Failure Prediction is used to identify and prevent possible failures that have not yet occurred and to eliminate their causes. In particular, the idea process for Failure Prediction is a systematic procedure for proactively identifying and preventing all risks or harmful situations that may be associated with a system [5].

The process for Failure Analysis proposed by I- TRIZ consists of seven steps.

1) Formulate the problem, which includes naming the problem encountered and describing the failure.
2) Identify the success scenario of the original system by outlining the step-by-step outcome of the normal process.
3) Identify where the problem occurred in order to reduce the scope of the analysis by identifying where in the step-by-step process of the system the problem occurs.
4) Formulate the transformed problem and turn the problem into a creative problem by asking "How can we make it happen?" and formulate it by maximising or totalising the problem that has occurred.
5) Find solutions by using existing knowledge or using technology and knowledge through exploration. Also, identify the resources that exist in the system or in the environment. This is because there is a specific resource that causes the problem.
6) Formulate a hypothesis about the failure and design a test to validate the hypothesis.
7) Fix the failure. This step ensures that the failure will not occur again.

Next, the failure prediction process involves eight steps.

1) Formulate Problem by identifying and formulating all the important problems that may occur in the current system.
2) Identify the success scenarios of the original system, the same as in the Failure Analysis step, outlining the outcome of each step in the normal course.

3) Formulate an inverted problem using the scenario method to describe all possible failures of the problem.
4) Record the possible failures within the system and construct a scenario tree by assuming when the problem occurs, the harmful end state and the worst case scenario.
5) Investigate the resources in or around the system that enable the scenario to be realised. This may lead to new failure scenarios.
6) Use a checklist to generate more ideas for failure scenarios.
7) Identify the final harmful state, identify the failure scenarios for it, and suggest ways to maximise it.
8) Derive a solution to solve the problem.

The process of driving failure analysis and failure prediction is shown in Fig. 1 [5].

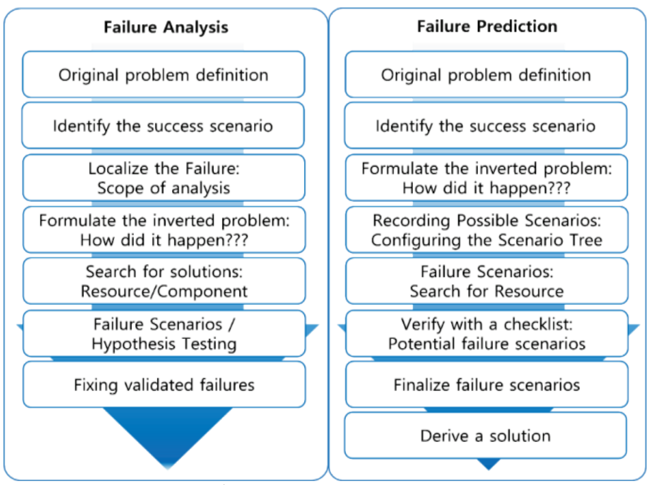


Figure 1 Anticipatory failure determination process [5], rewrite by quoting

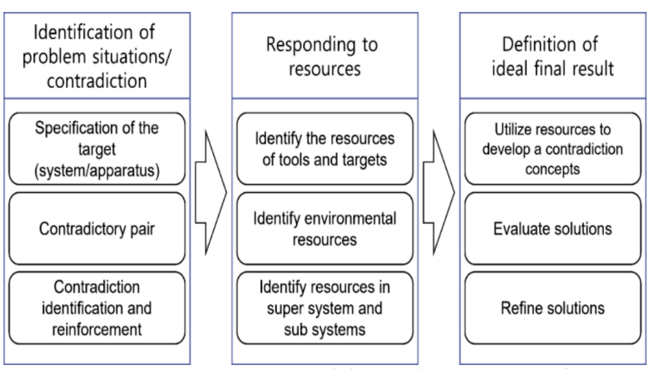


Figure 2 Invention problem solving process [6] (Theory of Inventive Problem Solving and Case Study – TRIZ Level 1 pp 52)

In this study, exploration of resources was essential in the process of problem solving through AFD. Furthermore, in order to realise the ideal in problem solving, it was essential to use the resources that exist in the system, and it can be said that problem solving ability is directly related to the ability to explore and use resources. The use of resources to realise the ideal of problem solving is, first, the use of resources that are generally available without cost within the system. In particular, it is most desirable to use casual or discarded resources. Second, to use resources available in the external environment. When an infinite amount of resources from the external environment is used to solve the problem, it is not limited by the difficulty of obtaining or the cost due to the scarcity of the resources. Common examples of resources from the external environment are air, water, temperature, sun and wind. Third, use tools to minimise costs. In this way, the use of resources takes into account the ideality of the problem to be solved, and the resources with the lowest cost and continuous procurement are selected and used to solve the problem. In addition, since there are always resources that cause failure when troubleshooting through AFD, eliminating resources that cause problems in the system can also be a way to realise the ideal. Fig. 2 shows the process of creative problem solving by using resources [6].

3 SCENARIO ANALYSIS METHOD

Scenario analysis is one of the methods for forecasting the future. It is a skilled and professional methodology that allows us to imagining of possible futures [7] and to presenting different situations that may occur in the future, based on consistency and logic [8]. The scenario analysis technique is a method of reasoning that starts from the present and logically derives different situations in order to know "what will happen" and develop a response to the facts that have occurred.

Scenario analysis was originally developed in the military field to analyse the relationship between military developments and strategy, and is said to have originated at the RAND Institute under the leadership of Herman Kahn, who served in the United States Air Force. It has since been popularised in the private sector and used as a management strategy by companies. Scenario analysis is a strategic tool that helps decision-makers in organisations to be prepared by examining possible outcomes and influencing factors and considering multiple alternatives [7, 9-14].

A classic example of the use of scenarios in business management is the case of Royal Dutch Shell. Pierre Wack, who worked as a scenario planner in the London office, used scenario analysis to predict the rise in oil prices caused by the fourth Middle East war, and responded quickly to the oil shock by reducing capital investment and improving the quality of refined oil products, greatly improving profit margins and growing the company to become the second largest oil company in the world. The company also used scenario planning to anticipate the collapse of the Soviet Union, securing the right to develop resources in Russia and growing into the global company it is today. There are countless other examples of company's successfully managing environmental change through scenario analysis.

In his study, Schoemaker [7] identified two common errors in scenario planning. The first is underestimation of environmental change and the second is overestimation. He argued that most people make the error of underestimation by imagining the future without considering the rate of change. He presented the steps to develop a scenario analysis as a 10-step process.

In the scenario construction process, step 1 is setting the time horizon of the issue to be analysed and the scope of the scenario analysis. Step 2 is identifying the key stakeholders in the issue to be analysed, and step three is identifying the underlying trends in the issue to be analysed, which he said is helpful to list in a chart or diagram and categorise the impact on current strategy as positive, negative or uncertain. Step 4 is recognising the key uncertainties that could have a significant impact on the issue, taking into account a range of factors including political, economic, social, technological and legal, and it is important to understand the relationships between the uncertainties.

Step 5 synthesises the information from steps one to four to form an initial scenario. A simple approach is to place all positive factors in one scenario and all negative factors in the other, creating an extreme initial scenario.

Step 6 is analysing of the scenarios created in step five to see if they are internally consistent or plausible. The internal consistency test should examine whether the scenarios are valid for the time period analysed. It should examine whether the outcomes caused by the uncertain factors can realistically occur in the real world. It is also important to ensure that the outcome of the uncertainty is one that key stakeholders are likely to choose. Step 7 is identifying some common themes, refined by the consistency and persuasiveness check in Step 6. This process identifies strategically relevant themes and organises the possible outcomes and situations.

Step 8 is reviewing the revised scenario for any overlooked or missing elements. This may require further research. Step 9 is developing a quantitative model, i.e., after further research, internal consistency is checked and a quantified scenario model is produced. Step 10 is finally developing the scenarios into decision scenarios. Through an iterative review of the process, the scenarios are finalised to confirm the strategy and generate new ideas.

The scenario building process is shown in Fig. 3.

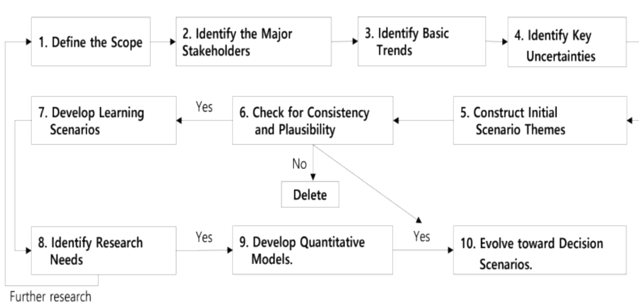


Figure 3 Constructing Scenarios Process [7]

In recent years, this process has been streamlined and often involves six steps. Identifying key issues are done in step one and key trends in step two. Selecting the driving uncertainties are done in step three and creating initial scenarios in step four. In step five, each scenario is given a title, and in step six, a response strategy for each scenario is developed.

When presenting scenarios using the scenario analysis method, it is common to present three to four scenarios. In practice, you can create many scenarios, but each scenario requires implications and explanations that can hinder decision-making. For this reason, scenarios should be limited and alternative scenarios should be developed [15].

3.1 Research Resources

A resource is anything that can be used to solve the problem that has occurred and create the desired situation [16]. In AFD analysis, when a scenario analysis method is used to formulate a failure hypothesis, there must be a resource within or adjacent to the system to support the hypothesis. In other words, a problem situation must have a cause (resource).

The concept of resources in TRIZ is very broad. From a detailed point of view, resources can be divided into basic resources, which can be classified as time, space, matter, field, energy and information, and induced resources, which are generated by the combination of basic resources. A function can be seen as an induced resource that combines matter, energy and information, and a person can be seen as an induced resource that combines matter, energy and information [16]. In addition, super-effects can also be seen as exploitable resources. Through systems thinking, available resources are not limited to the working system, but are explored through the nine windows such as the super system and sub-system that make up the system [2]. In addition, super effect is a new changed resource that occurs after problem solving, which suggests the best use of resources to solve the problem [17].

AFD aims to identify existing resources to reveal the occurrence of harmful or undesirable situations that threaten to disrupt the normal functioning of the system.

Fig. 4 shows the range of resources in terms of types and locations where they can be used.

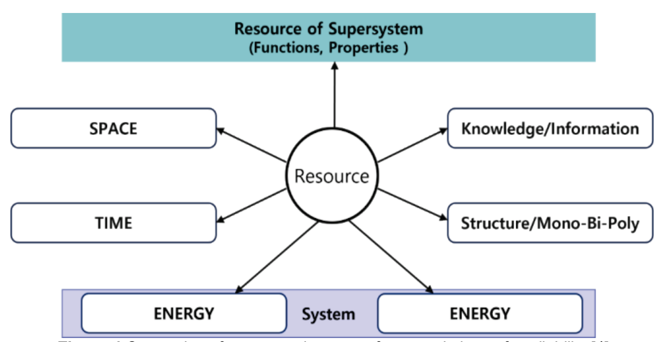


Figure 4 Categories of resources in terms of type and place of availability [4]

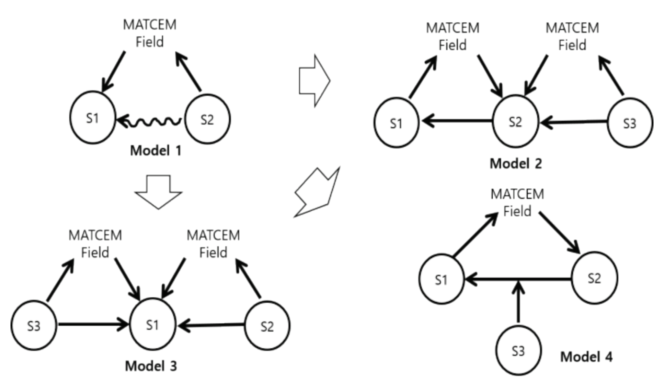


Figure 5 Standard Solution of Ideal Final Result

During the AFD process, it is necessary to identify the existence of resources that support the hypotheses generated by the scenario analysis, and subsequently to identify scenarios in which the resources are the cause of failure during the hypothesis validation process. It is important to note that resources may change in form or nature over time. They can also be combined with other resources to create new types of resources [16]. In TRIZ, the ideal solution is derived by using the standard solution to solve the problem by using other resources to solve the problem. Fig. 5 shows an example of how a standard solution becomes an ideal solution.

Model 1 represents a problem with a substance-field acting on the relationship between substance S1 and substance S2 [18]. As a standard solution to model 1, model 2 solves the problem by adding a new substance S3 and substance-field S2. Model 3 solves the problem by adding a new substance S3 and a field to substance S1. Model 4 solves the problem by adding a new substance S3 to the process of interaction between substance S1 and substance S2. The solution ideas of all models are sequentially considered and applied to implement the appropriate problem-solving model solution by applying the eight fields of MATCEMIB (Mechanical, Acoustic, Thermal, Chemical, Electric, Magnetic, Intermolecular, Biological) [19]. Thus, it can be said that in the AFD process, resources are an important factor affecting the cause and solution of the problem.

4 CASE STUDY FOR AFD

4.1 Case Study for Failure Analysis

4.1.1 Problem Cases of Mortar Shell Explosion Anomalies

The usefulness of the Inversion Analysis method is illustrated by solving the problem of anomalous explosions of mortar shells during the Second World War. The case is taken from the book "Theory of Inventive Problem Solving and Case Study - TRIZ Level 3".

Problem situation: During the Second World War, mortar shells were designed to fly unexploded after being fired, hitting enemy trenches where the impact would cause the projectile to hit the fuse and explode. As designed, the bombs fell into enemy territory, detonated and were used effectively, but towards the end of the war a problem arose where the bombs would explode in mid-air seconds after being fired. This was not a one-off mistake, but an ongoing problem, and the urgency of the fighting meant that the cause had to be identified and dealt with quickly. However, it was not possible to conduct experiments in the field and it was difficult to identify the cause. Inversion analysis can be used for problems like this where root cause analysis is difficult. Inversion Analysis can be done as follows:

Step 1, formulate the problem, name the problem, and describe the failure.

The problem is that mortar rounds are exploding in the air instead of reaching enemy territory and detonating, and this problem is called "Anomalous Mortar Shell Explosion". Because of this problem, friendly forces are damaged by the explosion and are unable to hit the enemy.

Step 2, create your first success scenario.

Decide to fire → Prepare mortar and shells → Fire mortar → Flight of shells → Reach target → Firing pin hits fuse → Explosion.

Step 3, identify the location of the problem. During the step-by-step process, identify the problem area.

The area where the problem occurs after the bombardment decision is made is identified as the area where the problem occurs during the process from mortar launch to shell flight, since the mortar launch proceeds normally without any problem. Set the analysis domain of the problem to Mortar Launch to Target.

Step 4, set the problem to Transitioned. The cause of the problem is that it is difficult to analyze the cause or observe the shell after it is fired, so it is difficult to analyze the cause with RCA (Root Cause Analysis) or CECA (Cause Effect Chain Analysis). Therefore, Inversion Analysis is used to turn the creative problem into "How can we make the shell explode in flight after being fired? The goal is to maximize the problem so that the mortar shells explode continuously and constantly during flight, rather than intermittently during flight.

Step 5, explore resources and use existing knowledge and skills to derive a solution.

From a TRIZ problem solving perspective, the resources that can be used for the explosion inside the system or in the adjacent environment during the flight after the mortar shell is fired are time, space, substance, field, energy, and information. Since the scope of the problem is the physical phenomenon of a mortar shell exploding, the substance-field should be given focus. The available resources include the outer casing of the mortar shell, the inner lining, the firing pin, the spring, and the metal plate that prevents the firing pin from escaping. According to the MATCEM analysis, there are various environmental resources such as air resistance, inertia, heat of friction with air, chemical reactions of the charge, sparks, and magnetic generation.

Step 6, develop a hypothesis for the failure (the shell always explodes in flight) and conduct a test to verify the hypothesis. In this case, the process of hypothesizing a situation where mortar shells can always explode in flight is determined by the scenario analysis method.

The first step of the scenario analysis is to determine the scope of the analysis, and the analysis is focused on shells in flight after firing. The fourth step of the analysis is to use resources to apply factors that may affect the object of analysis. In particular, focus on factors that change from before launch. The step five of the scenario analysis is where an initial hypothesized scenario is created. The chosen initial hypotheses were:

Hypothesis 1. An enemy spy was present and planted the timer.

Hypothesis 2. A chemical reaction of the internal charge occurred during the flight of the mortar shell.

Hypothesis 3. The mortar shell's firing pin ejector metal plate lost or malfunctioned during flight, causing the firing pin to strike the detonator.

Hypothesis 4. Mortar shells detonate during flight due to the heat of friction with air.

In Scenario Analysis step six, hypotheses that are plausible and consistent are selected. In step seven, the selected hypothesis is refined by testing its feasibility. In step eight, overlooked aspects are revisited, and in steps nine and 10, the final scenario is determined. The final hypothesis is also summarized after this process.

In this case, Hypothesis 3 was selected which states that "The metal plate that prevents the firing pin from dislodging during the flight of the mortar shell lost its function or failed and the firing pin struck the detonator." Hypothesis 3 was chosen the most likely because it is the most likely change that could have occurred during flight compared to before the mortar was fired, and also because it is the most likely change compared to before the initial explosion.

Based on the selected hypothesis, "How can a mortar shell be made to explode in flight?" was analyzed. The analysis showed that the metal plate to prevent firing pin disengagement was designed to perform two functions: to keep the firing pin from being disengaged by the acceleration generated when the mortar shell is fired, and to absorb energy to prevent the accelerated firing pin from hitting the detonator. And when the mortar shell flies and hits the target, the firing pin hits the detonator by inertia, causing it to explode normally. Therefore, to perform this function, in the early days of mortar shell production, mortar shell designers made a copper metal plate to prevent the firing pin from disengaging.

However, as the war progressed and copper plates became scarce, factory engineers substituted readily available steel plates for copper. Unlike copper, steel plates are elastic, and while inertia does not affect the firing pin in the early stages of a mortar shell's firing, elasticity occurs when inertia is lost. By changing the material of the metal plate to prevent the firing pin from disengaging compared to before the first explosion, it was found that the firing pin performed the function of preventing the firing pin from disengaging, but an unnecessary function, elasticity, occurred, causing the firing pin to hit the detonator a few seconds after firing. Thus, the hypothesis was verified.

In step 7, the cause of the abnormal explosion of the mortar shell was confirmed to be a malfunction (elasticity) of the "metal plate to prevent firing pin disengagement". To improve this, the problem can be solved creatively by changing the "firing pin disengagement prevention metal plate" to a ductile copper material that does not produce elasticity.

4.2 Case Study for Failure Prediction

4.2.1 Problem Case that Causes "Company S" to Fail

We will present the case of Failure Prediction using the example of Company S, which has grown into a global company. The case is based on Daeje, Jin "Manage Your Passion" [20], as well as interviews and press releases from related parties.

In November 1994, a strategy meeting was organized by the head of the memory division of S Company. Executives from the strategic planning department were invited to the strategy meeting, and the topic was "Company S (semiconductor business) is in danger of going out of business.

The problematic situation was to select a possible scenario that could lead to the failure of Company S (semiconductor business) in 1997, three years later. In addition, countermeasures for the situation in which the continued business promotion fails need to be prepared. In other words, the anticipatory failure determination method and scenario analysis method to the situation that predicts the failure of the business should be applied. The analysis steps were as follows:

Step 1, formulate the problem, name the problem that occurred, and explain the failure.

Three years later, Company S (a semiconductor company) is facing a difficult business situation that may lead to its failure. This situation can be attributed to a number of environmental changes, but there are a number of uncertain situations that can be considered. Define the problem by considering the most likely trends within the industry. Industry trends include integration between companies, technology development, and exploration of new business areas.

Step 2, Company S wants to increase its market share by developing new technologies while expanding its reach to similar businesses. Create a scenario of successful business development that Company S wants.

Maintain market share → Develop new high-performance semiconductors → Create new customers and increase existing supply → Increase market share → Expand non-memory business.

Step 3, formalize the problem using a scenario analysis method by listing all the possible failures that could lead to the problem through inversion analysis. Formulate the problem by listing the ways in which the problem could occur.

In other words, hypothetically create a situation that causes a problem in driving the normal success scenario. For example, a strong new competitor enters the industry that is not an existing competitor. Or a new technology is developed that produces a product that is incomparable to what is currently being produced. Depending on the geographical environment, there are various geographical environmental changes, such as earthquakes, floods, and droughts. Assume that changes occur, such as a change in the political situation in a country that leads to increased regulation, or a change in consumer needs that leads to a decrease in demand for existing products.

Step 4, record possible failure scenarios within the system, and construct an initial scenario by synthesizing the time of problem occurrence, harmful end state, and extreme situation.

Hypothesis 1. The emergence of a strong competitor in the industry or the expansion of the business leads to a decline in market share, which leads to a decline in the productivity of the company and the failure of Company S.

Hypothesis 2. The development of new technologies leads to the mass production of substitutes for existing semiconductors and the semiconductor industry enters a period of decline, causing Company S to fail.

Hypothesis 3. Due to changes in the geographical environment (earthquake, water damage, lack of industrial water, etc.), the semiconductor infrastructure collapses, making it difficult for Company S to produce products.

Hypothesis 4. Due to the deterioration of the international situation, import restrictions on production materials, materials, parts, and production equipment occur, making it difficult to produce products and causing Company S to fail due to a decrease in production rate and sales.

Step 5, explore the resources that exist in or around the system in which the initial scenario could be realized.

Looking at the various resources that could cause problems, there are established companies that could be strong competitors. These include Intel, Toshiba, Hitachi, and Texas Instruments. There are not many resources that can cause geographic change, but there are many new competitors, and there is a high probability of mergers among semiconductor companies, so there are many resources for developing new technologies. In addition, South Korea, where S Company is located, has a complex and delicate international relationship with the United States, China, Japan, and North Korea, so there is a lot of volatility in international affairs.

Step 6, use the checklist to get more ideas for failure scenarios.

In a strategy meeting, the memory division of Company S reviewed the various possible scenarios and resources that could cause problems, and finally settles on two failure scenarios. The finalized scenarios were,

Hypothetical Scenario 1.

The current chairman of Company I, Mr. Grove, retires and a new chairman is appointed. After taking office, the new chairman announces a business policy to focus on the memory semiconductor business as well as the current CPU production. Company I, a powerhouse in the memory semiconductor industry that integrates semiconductor devices that control memory, declares that it will determine the specifications of memory semiconductors based on its patents and technologies and will not allow other companies to use its technology. Company S, whose main business is memory semiconductors, is facing a serious crisis in its business due to the emergence of strong competitors and technological threats.

Hypothetical Scenario 2.

Due to the delicate international situation in Northeast Asia, Japan enacts a law to prevent the export of semiconductor manufacturing equipment to South Korea, such as exposure equipment needed to produce D-RAM, in order to punish South Korea economically.

This means that Company S will not be able to procure the materials, components and equipment needed to produce semiconductors, causing a significant disruption in the production of its main product, D-RAM.

Step 7, check the final harmful state, assuming the worst-case scenario.

If the scenario is realized, in the worst case, Company S may go bankrupt and the country's core business, the semiconductor industry, may collapse. In particular, if the two hypothetical scenarios occur simultaneously, the situation will be so serious that both Company S and the country will have no countermeasures.

Step 8, derive a solution to resolve the problem.

Company S did not disclose a clear plan for countermeasures based on scenario management. However, through scenario management, Company S anticipated a scenario in which the company might fail, so it prepared a medium- to long-term preparedness plan internally.

Subsequently, the failure scenario was realized. However, it can be judged that it was able to adequately prepare for the failure scenario. Company S has now become a world-class company in the semiconductor industry. Here is an example of a real failure scenario,

Fact 1. In 1997, Mr. Grove, Chairman of Company I, retired and Mr. Barrett, Senior Vice Chairman, took over. Company I then entered the memory semiconductor industry and introduced a memory product called "iRAM" to the market. In fact, the failure scenarios predicted by scenarios and predictive failure decision methods were realized. However, unlike the failure scenario that expected the worst, Company S cleverly solved the problem through thorough preparation. First, it was able to maintain price competitiveness due to its production cost advantage through technology development. In addition, it maintained product quality through the superiority of D-RAM's heat control technology. By reducing the defect rate through product quality control, the company was able to maintain the trust of its customers. In this way, the company was able to overcome the crisis wisely.

Fact 2. On July 4, 2019, Japan's Ministry of Economy, Trade, and Industry imposed export restrictions on fluorine polyimide, photoresist, and hydrogen fluoride, which are key materials for semiconductors and displays, to South Korea due to deteriorating diplomatic relations. This is the situation predicted in Failure Scenario 2, which means that it is difficult to substitute by regulating items corresponding to materials and components. Company S is in an urgent situation to produce products on time to deliver the contracted products to its customers. In addition, the company does not have much inventory and urgently needs to procure parts.

However, Company S has been trying to locate photoresist, which is the most difficult to replace, for more than 20 years. Hydrogen fluoride has also been localized, and the company has supported technology development by domestic companies. On the one hand, the company put pressure on the Japanese government, pointing out that the Japanese company could go bankrupt if it failed to deliver its products. On the other hand, Company S worked with other domestic S-corporations to localize their supply chains. By preparing for the worst-case failure scenario through scenario management and contingency planning, they were able to overcome the risk. In the end, a major Japanese company invested more than 100 billion yen to build a photoresist plant in Korea. This outcome is the result of the failure scenario that highlighted the superiority of Company S.

5 DISCUSSIONS

The integration of scenario analysis into the Anticipatory Failure Determination (AFD) methodology has yielded promising results in enhancing problem-solving efficacy. The study demonstrates that by employing scenario analysis, organizations can systematically formulate and test failure hypotheses, thereby preemptively addressing potential risks and vulnerabilities. This aligns with previous research by Chybowski, Leszek & Gawdzińska, K. & Souchkov, Valeri

[4], which emphasizes the importance of early-stage application of AFD in system development for risk mitigation and performance optimization.

Furthermore, the findings underscore the broader implications of leveraging anticipatory approaches in problem-solving. By anticipating and proactively addressing failures, organizations can not only minimize costs associated with downtime and repairs but also safeguard their reputation and maintain customer trust. Moreover, the integration of scenario analysis into the AFD framework fosters a culture of continuous improvement and innovation. By systematically exploring potential failure scenarios, organizations can uncover new insights, identify areas for enhancement, and drive iterative innovation. Several studies emphasize the role of anticipatory approaches in stimulating creativity and innovation within organizations. Additionally, the study underscores the importance of interdisciplinary collaboration in addressing complex challenges. By bridging the domains of future prediction and problem-solving methodologies, researchers and practitioners can leverage diverse perspectives and expertise to develop holistic and robust solutions.

Furthermore, the research highlights the need for organizations to embrace a proactive mindset in navigating uncertainty and change. By embracing anticipatory approaches like AFD, organizations can position themselves to effectively adapt to evolving circumstances and seize emerging opportunities. The study underscores the transformative potential of integrating scenario analysis into the AFD methodology for enhancing problem-solving efficacy. By anticipating and proactively addressing failures, organizations can mitigate risks, foster innovation, and drive continuous improvement. This research contributes to the growing body of literature on anticipatory problem-solving approaches and underscores the importance of proactive strategies in navigating complex and uncertain environments. The study has achieved its objectives by effectively integrating scenario analysis into the Anticipatory Failure Determination (AFD) method within the Theory of Inventive Problem Solving (TRIZ), thereby enhancing the understanding of predictive failure analysis and problem-solving methodologies. This integration has provided insights into proactive approaches for identifying and addressing potential failures, ultimately contributing to more robust problem-solving strategies.

6 CONCLUSIONS

Foresight is the process of trying to predict what will happen in the future by taking the current situation and adding the changes that are expected to occur in the future. There are many ways to predict the future, and there is no right answer. However, the reason for this difficulty is that there are many risks and damages that can be caused by an unprepared future. In other words, the purpose of forecasting is to prepare for the risks that will come in the future. One of the different methods of foresight is scenario analysis, which is a method of preparing different possibilities of what might happen in the future. Scenario analysis does not determine the future. It assumes all possible scenarios and prepares for the worst, based on the most likely to happen scenario.

The TRIZ perspective on foresight defines a situation where risk and harm occur as a problem situation and uses creative problem reversal to solve problems, where the cause of the problem is difficult to identify or the process of the problem is difficult to observe. By switching the problem, it solves the problem by switching to "How can I cause failure?" instead of worrying about "Why does failure happen? This is how predictive failure analysis works. It is a method of creating a hypothesis that causes a problem and validating it. In this process, there are resources that cause failure.

This study applied a method based on the scenario analysis method in the process of using the anticipatory failure determination method and examined cases. In the process of using the anticipatory failure determination method to turn the problem into a creative method and solve it, the scenario analysis method is applied at the stage of setting the failure hypothesis. In other words, the failure hypothesis is analyzed using the scenario analysis method to determine the situation that will cause the failure by examining the hypothesis that is likely to be realized and is consistent.

In addition, in the process of establishing failure hypotheses through the scenario analysis method, it is easy to derive failure hypotheses for problem solving by exploring and applying resources that cause problems that exist in the system or resources that allow ideal problem solving among the troubleshooting methods. In other words, as a way to solve problems that are difficult to identify and observe, it is easy to solve problems by setting hypotheses for the anticipatory failure determination method based on the scenario analysis method using the exploration of system resources. TRIZ's anticipatory failure determination method, which solves problems by shifting thinking, is a creative problem-solving method, and it is necessary to actively conduct research in this area in the future. Creativity is learned, not born. The convergence of research on future prediction and problem-solving methods underscores the growing recognition of the interconnectedness between anticipation and effective resolution of challenges. As organizations navigate increasingly complex and uncertain environments, the need to anticipate future scenarios becomes imperative for informed decision-making and risk mitigation. By integrating predictive techniques with robust problem-solving methodologies like AFD, researchers and practitioners can not only identify potential failures in advance but also develop proactive strategies to address them, ultimately enhancing organizational resilience and adaptability in the face of dynamic change. This synthesis of future prediction and problem-solving methodologies represents a pivotal advancement in addressing the intricacies of modern-day challenges and fostering sustainable success.

7 REFERENCES

[1] Barry, K., Domb, E. & Slocum, M. S. (2010). TRIZ-what is TRIZ. *The TRIZ journal,* 603-632.
[2] Yun, S. B., Song, Y. W. & Sue, Y. P. (2023). A study on dew condensation problem at curling stadium using ARIZ-85C. *Asia-pacific Journal of Convergent Research Interchange, 9*(2), 13-28. https://doi.org/10.47116/apjcri.2023.02.02

[3] Sarno, E., Kumar, V. & Li, W. (2005). A hybrid methodology for enhancing reliability of large systems in conceptual design and its application to the design of a multiphase flow station. *Research in Engineering Design, 16*, 27-41. https://doi.org/10.1007/s00163-005-0003-1
[4] Chybowski, L., Gawdzińska, K. & Souchkov, V. (2018). Applying the anticipatory failure determination at a very early stage of a system's development: Overview and case study. *Multidisciplinary Aspects of Production Engineering, 1*(1), 205-215. https://doi.org/10.2478/mape-2018-0027
[5] Kaplan, S., Visnepolschi, S., Zlotin, B. & Zusman, A. (2005). New tools for failure and risk analysis, anticipatory failure determination (AFD) and the theory of scenario structuring (ed.) *Ideation International Inc.* Farmington Hills.
[6] Kim, I. C., Park, Y. S., Song, Y. W., Joo, J. M. & Choi, S.W. (2022). *Theory of Inventive Problem Solving and Case Study – TRIZ Level 1*, GS Inter Vision, 52.
[7] Schoemaker, P. J. (1995). Scenario planning: A tool for strategic thinking. *MIT Sloan Management Review.*
[8] Porter, M. E. (1985). Technology and competitive advantage. *Journal of Business Strategy, 5*(3), 60-78. https://doi.org/10.1108/eb039075
[9] Wack, P. (1985). *Scenarios: Shooting the rapids. How medium-term analysis illuminated the power of scenarios for shell management* (No. REP-9208. CIMMYT.).
[10] Wack, P. (1985). Scenarios: Uncharted waters ahead. *Harvard business review, 63*(5), 72-89.
[11] Heijden, K. A. (1996). *Scenarios: The art of strategic conversation*. John Wiley & Sons.
[12] Schwartz, P. (1996). *The Art of the Long View: Planning for the Future in an Uncertain World*, Currency and Doubleday.
[13] Gausemeier, J., Fink, A. & Schlake, O. (1998). Scenario management: An approach to develop future potentials. *Technological Forecasting and Social Change, 59*(2), 111-130. https://doi.org/10.1016/S0040-1625(97)00166-2
[14] Godet, M. (2000). The art of scenarios and strategic planning: Tools and pitfalls. *Technological forecasting and social change, 65*(1), 3-22. https://doi.org/10.1016/S0040-1625(99)00120-1
[15] Courtney, H., Kirkland, J. & Viguerie, P. (1997). Strategy under uncertainty. *Harvard business review, 75*(6), 67-79.
[16] Kim, I. C., Park, Y. S., Song, Y. W., Joo, J. M. & Choi, S. W. (2022). *Theory of Inventive Problem Solving and Case Study – TRIZ Level 1*, GS Inter Vision, 51.
[17] Kim, I. C., Park, Y. S., Song, Y. W., Joo, J. M. & Choi, S. W. (2022). *Theory of Inventive Problem Solving and Case Study – TRIZ Level 3*, GS Inter Vision, 71.
[18] Dobrusskin, C., Belski, A. & Belski, I. (2014). On the effectiveness of systematized substance-field analysis for idea generation. *Innovator*, 1, 123-27.
[19] Belski, I. (2007). *Improve your thinking: Substance-field analysis*. Melbourne: TRIZ4U.
[20] Jin, D. J. (2022). *Manage passion*. Gimmyoung, 91-94. (in Korean)

Authors' contacts:

Seock-Bum Yun, Adjunct Professor
Graduate School of Entrepreneurial Management Chung-Ang University,
84 Heukseok-ro, Dongjak-gu, Seoul, 06974, Republic of Korea
acesys@cau.ac.kr

Yong-Won Song, Professor
(Corresponding author)
Nano & Semiconductor Engineering, Technical University of Korea,
237 Sangidaehak-ro, Siheung-si, Gyeonggi-do, 15073, Republic of Korea
ywsong@tukorea.ac.kr

Young-Pyo Sue, Team Leader
Korea Institute of Industrial Technology,
89, Yangdaegiro-gil, Ipjang-myeon, Seobuk-gu, Cheonan-si, Chungcheongnam-do, Republic of Korea
sueyp@kitech.re.kr

Sang-Jun Lee, Ph.D. Candidate
Graduate School of Entrepreneurial Management Chung-Ang University,
84 Heukseok-ro, Dongjak-gu, Seoul, 06974, Republic of Korea
gokbal@cau.ac.kr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20231227064230

Original scientific paper
Received: 2023-12-27, Accepted: 2024-04-18

Contextual Ambiguity Framework for Enhanced Sentiment Analysis

Pinky Yadav*, Indu Kashyap, Bhoopesh Singh Bhati

Abstract: Negation is a universal linguistic phenomenon that affects the performance of Natural Language Processing (NLP) applications, especially opinion mining data. Many words exists in sentences that have multiple interpretations or sentiments depending on how they are placed with respect to the negation word in the sentence. A cutting-edge framework is designed to tackle the nuanced challenge of detecting contextual ambiguity through negation in sentiment analysis. The approach uniquely combines advanced natural language processing techniques with deep linguistic insights, enabling the accurate interpretation of sentiment in complex sentences where negation plays a key role. The framework identifies negation cues and their scope, then assesses their impact on sentiment, considering contextual dependencies and word semantics. The model's innovation lies in context-sensitive algorithms that adeptly handle different sentence structures and idiomatic expressions, a notable advancement over traditional sentiment analysis tools. Particularly effective in interpreting sarcastic or ironic statements, the framework significantly outperforms existing models in accuracy, especially in negation-heavy contexts. This advancement enhances sentiment analysis applications like social media monitoring and customer feedback analysis, offering a more nuanced understanding of public opinion.

Keywords: contextual ambiguity; linguistic nuances; machine learning; natural language processing; sentiment interpretation; text analytics

1 INTRODUCTION

Contextual ambiguity can be defined as the presence of such words in sentences that have multiple interpretations or sentiments depending on how they are placed and what the grammar usage has been. It is a very challenging task in sentiment analysis as the models need to consider the surrounding words, phrases, or sentences to accurately determine the sentiment. Consider the word "sick" in the following sentences:

a) "The new album by the band is sick! I love it."
b) "I feel sick today; I can't go to work."

In sentence (a) "sick" is used in a positive context to mean something is excellent. In sentence (b) "sick" is used in a negative context to indicate feeling unwell. The sentiment of the word "sick" varies based on the surrounding context. Contextual ambiguity poses challenges for sentiment analysis models because they must consider the broader context to accurately determine the sentiment of a text. Simple keyword-based approaches may not work well in these cases. To resolve contextual ambiguity, sentiment analysis models need to consider the surrounding words, phrases, and even the tone of the text to make an accurate sentiment prediction [15]. They should analyze the entire sentence or paragraph, not just isolated words. Some sentiment analysis models use sentiment lexicons and predefined rules to help disambiguate sentiment in context. Lexicons assign sentiment scores to words based on their typical usage, and rules define how to handle certain context-specific cases. In some cases, domain-specific or user-defined rules can be added to sentiment analysis models to handle contextual ambiguity unique to a particular application or industry [16]. Therefore, this research work delves into the world of contextual ambiguity in sentiment analysis, examining the challenges it poses and its significance. It also explores the impact of negation words on sentiment analysis, illustrating how they can reverse or modify sentiment and complicate the task. The objective is to present the complexities of this problem and provide a new approach that addresses these challenges effectively. Addressing Contextual Ambiguity offers various compelling motivations such as Enhanced Communication, Improved User Experience and Effective Informational Retrieval. It also leads to precision in decision making, Advancements in Artificial Intelligence and Ethical considerations.

1.1 Contextual Factors

In addition to providing an objective account of events, texts frequently communicate the sentiments of authors or people involved in the recounted event. The emotional disposition is conveyed through the selection and organization of language. Sentiment is supported by context, which gives statements complex interpretations that are frequently dependent on language, situational, or cultural nuances. The same phrase might express completely different feelings in different contexts, so ignoring this context can result in a serious misreading of sentiments. While certain words consistently exhibit either positive or negative connotations, others are prone to undergo contextual shifts in valence as a result of neighboring words and the overall structure of the text. Instances of valence-shifting encompass the subsequent examples:

- The concept of negation refers to the logical operation of denying or contradicting a statement or proposition. The predominant form of valence-shifting, seen by the phrase "She doesn't like this laptop," involves the utilization of the negation "doesn't" to reverse the affective connotation of the verb "like."
- Defective verbs are lexical items employed to convey various notions such as capacity, certainty, permission, request, competence, suggestion, order, obligation, or advice. These verbs are inherently incomplete and require the presence of another verb that possesses semantic content and modifiers to convey a complete meaning. When a statement contains a word that is deficient, the emotional significance of the sentence is sometimes diminished due to the

inadequacy of the feeling word or phrase used in conjunction with the verb. The statement "The school should reduce the tuition fee" expresses a negative evaluation, whereas the statement "The school reduces the tuition fee" presents a factual observation.

- Words with heightened or lessened intensity: These lexical choices amplify or reduce the emotional significance of the associated term. In the given instance, the phrase "She studies hard" exhibits a heightened emotional connotation in comparison to "She studies quite hard," while possessing a diminished emotional connotation when contrasted with the statement "She studies very hard." The observed correlation between emotions (e.g. intensity) and the presence of reinforcing or mitigating words (e.g. "very" or "quite") suggests a causal relationship.
- The phenomena of contrast is observed when the text from the provided link displays a shift in emotional valence, indicated by the use of phrases. Typically, the emotional inclination of the entire sentence is only evident on one side. An instance of an insulting remark can be observed in the given statement, "This laptop is beautiful, but the price is too expensive." The term "too expensive" contributes to the pejorative nature of the remark, mostly owing to the presence of the linking word "but" earlier in the sentence.
- Incompatibility or inconsistency is evident in numerous instances where the emotional inclination of a text is largely reliant on the surrounding context.
- Valence-shifting is a phenomenon that occurs when the semantic interpretation of a word is subject to alteration within a particular context [11]. Machine learning techniques like bag-of-words and n-grams fail to account for the impact of negation structures and other sentiment valence shifter structures. This limitation is evident in statements like "I like this hotel but the price is quite high" and "I like this hotel but the price is too high." Based on a bag-of-words model, it is probable that the sentiment value of the two lines is equivalent, as they both contain the emotive words "like" and "high." Nevertheless, the bag-of-words approach fails to account for the presence of words such as "quite" and "too," which significantly influence the sentiment conveyed by the two statements. Sequence mining techniques are employed to uncover patterns related to valence-shifting. These patterns encompass several linguistic phenomena, including negation, contrast, intensification, and attenuation of polarization. Semantic Oriented CALculator (SO-CAL) is an innovative system that effectively addresses valence-shifting through the utilization of rule-based procedures and a mix of methodologies. This system incorporates rules and collections of words that have been annotated with emotional attributes. The researchers [14] and [15] utilize a dependency tree that incorporates comprehensive syntactic structure information in order to establish syntactic rules for assessing the influence of negation and other valence-shifting structures on the emotional polarity of sentences or the entirety of a document. In this analysis, straightforward yet efficient strategies are examined for employing rules to identify instances of contextual valence changing. The utilization of neural network methods, particularly the Long short-term memory (LSTM) network [16], has facilitated the advancement of deep learning models. In this regard, the attention mechanism has emerged as a potent approach for effectively representing context.

Although previous research work have investigated the role of gesture in conveying denial, there is a need for more investigation in the context of the interplay between scope-bearing elements, particularly the relationship between different types of quantification and negation. Interactions of this nature present a communication issue due to the production of semantically ambiguous phrases that can be interpreted in numerous ways, with the scope of negation varying depending on the intended interpretation. An example is provided in which the negator and quantifier are underlined. All the magnolias won't bloom [3]. The relevant interpretations are captured by the logical representations in the following: $\forall x$ [magnolia $(x) \rightarrow \neg$bloom (x)], $\neg\forall x$ [magnolia $(x) \rightarrow$ bloom (x)]

To demonstrate the lack of consistency in empirical findings regarding the English language [10] conducted perception experiments. These experiments focused on the prosodic cues used to disambiguate socially ambiguous sentences. Specifically, sentence-final falling intonation was used for sentences such as "all>not," while sentence-final rising intonation was used for sentences like "not>all." The results showed that participants were able to interpret the intended meaning with a relatively high degree of success, ranging above 63% depending on the sentence type. This success rate was significantly higher than chance. Again, these rates were significantly higher than chance. In contrast, the initial production experiment [2] demonstrated that speakers tend to provide prosodic cues with limited reliability in a semi-naturalistic environment. This context involved the speakers reading scripted ambiguous lines that were embedded within disambiguating paragraph-length texts. In the context of the production investigation, it was observed that speakers commonly employed descending intonational contours when articulating negation combined with quantification utterances. According to the findings of Syrett et al. [17], it was determined that there is no direct correspondence between intonation contour and scopal interpretation in the context of language production. Therefore, it was concluded that the patterns of intonation cannot be considered as either a necessary or sufficient condition for resolving the ambiguity in sentences with multiple scopal interpretations [12].

The remaining sections of the paper are organized as follows: The next section discusses prior work on sentiment analysis for contextual ambiguity. Section 3 elaborates on the proposed methodology and model building process. Section 4 offers the results. Finally, Section 5 discusses the conclusions of the research, and outlines future directions.

2 LITERATURE REVIEW

The field of sentiment analysis has evolved significantly since its inception. Early foundational work by Pang and Lee [6] outlined basic approaches for classifying sentiments in texts, primarily using bag-of-words models and basic machine learning techniques. These techniques laid the groundwork for more advanced sentiment analysis. Negation handling is a well-acknowledged challenge in sentiment analysis. Wiegand et al. [2] highlighted the complexities that negation introduces in determining sentiment orientation, emphasizing the need for sophisticated linguistic analysis. Similarly, Councill et al. [3] demonstrated the impact of negation on sentiment analysis accuracy. Contextual ambiguity in sentiment analysis has been explored extensively. Studies by Cambria et al. [4] emphasized the importance of context in interpreting sentiments, especially in idiomatic and sarcastic expressions. This is further supported by research of Reyes and Rosso [5], who focused on the role of irony and sarcasm in sentiment analysis.

The integration of advanced NLP techniques has significantly improved sentiment analysis. Socher et al. [6] introduced deep learning approaches, which leverage neural networks for better understanding of contextual nuances. This is further elaborated by Zhang et al. [7], who explored the use of deep learning for negation and speculation identification.

Arabic language poses challenges for automatic processing due to its multiple dialects, ambiguous syntax, and limited high-quality datasets. Duwairi et al. [21] introduced a novel framework for augmenting Arabic sentences using the language's rich morphology, synonymy lists, and grammatical rules. The approach, focused on sentiment analysis, significantly increased the size of initial datasets and improved accuracy by 42 % through reliable rule-based augmentation. Aoumeur et al. [22] introduced a new dataset, CASAD, derived from art books and labeled by human experts. Unlike previous methods relying on word frequency, this method employs word embedding techniques Word2Vec to extract deep relations in formal Arabic language features. The dataset is evaluated using machine learning algorithms such as SVM, LR, NB, KNN, LDA, and CART, with statistical methods for validation and reliability. Results show that the Logistic Regression with Word2Vec approach achieves the highest accuracy in predicting topic-polarity occurrences in classical Arabic texts.

Alshutayri et al. [23] explores sentiment analysis on social media, focusing on Arabic Twitter data. It employs machine learning algorithms like Naïve Bayes, Logistic Regression, and Support Vector Machines, with bigrams and unigrams as features. Logistic Regression achieved the highest accuracy at 63.40 %. Additionally, a deep learning approach using Long Short-Term Memory (LSTM) neural network achieved a higher accuracy of 70 %, surpassing related works in the field. The effectiveness of these methods in analyzing sentiments expressed in Arabic tweets is demonstrated. Cumaoğlu et al. [24] introduced a new dataset gathered from Twitter, focusing on Arab opinions about Turkey across various topics. The dataset is multi-dialectic Arabic and covers fields like the Turkish economy, tourism, food, and politics. A deep learning-based Arabic Sentiment Analysis (ASA) approach using Word2Vec and Bidirectional Encoder Representations for feature extraction is employed. Bidirectional long short-term memory, Convolutional neural networks, feedforward neural networks, and a transformer auto classifier based on AraBERT are applied for binary classification. The results show that AraBERT outperforms Word2Vec, and the transformer auto classifier achieves the highest accuracy in classifying positive and negative emotions in the dataset.

Refai et al. [25] addressed the challenge of dataset adequacy in Arabic language models by proposing a new Data Augmentation (DA) method using AraGPT-2. Existing approaches for Arabic DA are limited, relying on traditional methods like paraphrasing. The proposed method leverages AraGPT-2 for augmentation and evaluates generated sentences using various metrics. The augmented dataset is then tested on sentiment classification tasks using the AraBERT transformer. Results demonstrate improved performance across different sentiment Arabic datasets, with F1 scores increasing by 4 % in AraSarcasm, 6 % in ASTD, 9 % in ATT, and 13 % in MOV.

Díaz et al. [25] worked on negation and speculation detection in natural language processing. It defines these phenomena, discusses the need for their processing, reviews existing research, and provides a comprehensive list of resources and tools. This research introduced new datasets and scripts, offering an overview of the current state of the art in negation and speculation detection. Mahany et al. [1] introduced NSAR, the first Arabic corpus designed for the purpose of review analysis, which has been annotated with negation and speculation markers. The dataset consists of 3,000 words that span various categories, revealing that 29% of these sentences entail negation, while just 4% of them contain conjecture. The presence of a high level of inter-annotator agreement serves as evidence of the dependability of the data. This particular resource holds significant importance in the field of Arabic Natural Language Processing (NLP), since it provides valuable insights into the prevailing linguistic characteristics seen in reviews.

Significant progress has been achieved in the field of supervised relation extraction (RE) using neural networks (NNs) in recent years, particularly with the introduction of deep learning models such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) [18]. Within this particular paradigm, the task of relation extraction (RE) is approached as a classification problem. In order to accomplish this objective, informative features are carefully constructed using training data, and a variety of classification models are then trained. The taught classifiers are subsequently utilized to make predictions about relationships, either using the pipeline approach or the joint learning approach. The pipeline approach involves the sequential execution of named entity recognition (NER) and relation extraction (RE) operations [19].

The first step in the research experiment involves conducting Relation Extraction (RE) on sentences that contain identified entity pairs. Afterwards, entities that have relationships are merged into triples and presented as the predicted outcome. Nevertheless, this methodology involves the division of Named Entity Recognition (NER) and Relation Extraction (RE) activities, hence rendering the performance of RE vulnerable to the outcomes of NER. On the other hand, the collaborative learning approach entails the concurrent detection and extraction of many sorts of relationships between items [20]. Both methodologies commence by acquiring the fundamental vector representation of characters, subsequently extracting sentence attributes through the use of diverse neural network models. The procedure is finalized by employing nonlinear classifiers for the purpose of relation classification.

Zheng et al. focused specifically on the detection of irony and sarcasm in text, noting that these linguistic devices often invert the literal sentiment, posing a significant challenge for automated systems. Their research highlighted the need for models to recognize contextual cues that indicate sarcasm or irony.

In the work by Poria et al. [3], the authors demonstrated that the sentiment polarity of a word or phrase can change dramatically based on the context in which it appears. This research underlines the need for sentiment analysis models to incorporate broader contextual analysis rather than relying solely on individual word sentiments.

The advancement in deep learning techniques has provided new opportunities for addressing contextual ambiguity. Socher et al. [4] showed how recursive neural networks could be used to understand the compositional effects of sentiment in sentences, accounting for the influence of context.

The emergence of contextual word embeddings, like those introduced by Devlin et al. [5] in BERT (Bidirectional Encoder Representations from Transformers), represents a significant leap in capturing the contextual nuances of language for sentiment analysis. These models are designed to understand the meaning of a word in the context of the surrounding text, thereby providing a more accurate sentiment interpretation.

3 METHODOLOGY

The aim is to enhance the performance of collaborative recommender systems by addressing challenges related to word ambiguity, data sparsity, and mistake rate. Textual evaluations are to be incorporated into the matrix of ratings provided by users for items. Textual reviews have the potential to mitigate the issue of data sparsity and enhance the quality of suggestions. The aim is to address the issue of domain sensitivity in recommender systems by including contextual information in the sentiment analysis model. The majority of sentiment-based recommender systems commonly employ conventional techniques that lack semantic analysis. Consequently, these systems encounter challenges related to data sparsity and domain sensitivity. There have been limited recommendations on the inclusion of domain sensitivity information in order to enhance sentiment accuracy and address the issue of data sparsity [10]. The manner in which domain sensitivity information can enhance the quality of suggestions remains unclear.

3.1 Data Collection

The Amazon Movie Reviews dataset is a rich and extensive collection of user-generated reviews sourced from the Amazon website, one of the world's largest online retailers. This dataset forms a critical part of the web data used in sentiment analysis and natural language processing research. It encompasses 8 million reviews, offering insights into consumer opinions, preferences, and viewing experiences. Each review typically includes a textual comment along with a star rating, providing a dual perspective of qualitative and quantitative data. The sheer volume and variety of the reviews make this dataset an invaluable resource for training and testing machine learning models, especially those focused on sentiment analysis.

3.2 Data Preprocessing

Data preprocessing is an essential stage in any text analysis task. It involves transforming the raw data into a format that can be easily analyzed. Several preprocessing steps are used, including Data Cleaning, Tokenization, Removing stopwords, Removing punctuations, and Lowercasing.

3.3 Proposed Algorithm

There are two primary stages that are involved. The process initially starts with conversion of textual reviews into numerical ratings. The incorporation of sentiment grading into collaborative combination is a secondary consideration. Context-based sentiment analysis is a computational approach that involves the conversion of domain-specific textual reviews into numerical ratings. The CF method utilizes the extended user-item matrix obtained from the previous phase in order to calculate the recommendation value.

Inputs:

R: A set of textual reviews.
S: Traditional numerical ratings.
Algorithm Steps:
Textual Review Preprocessing and Numerical Conversion

Step 1.1 Text Preprocessing:
- Segment phrases and annotate with part-of-speech (POS) tags.
- Apply custom tagging for domain-specific features.

Step 1.2 Opinion Word Extraction:
- Extract opinion words using predefined linguistic rules.

Step 1.3 Sentiment Scoring:
- Apply sentiment analysis techniques to convert reviews into numerical sentiment scores.

Step 1.4 Merge Scores:
- Combine S with sentiment scores to create an augmented dataset.
- Extended User-Item Matrix Construction

Step 2.1 Matrix Formation:
- Construct an extended user-item matrix incorporating both traditional ratings and sentiment scores.
- Collaborative Filtering with Sentiment Analysis

Step 3.1 Score Generation using OWs:

Step 3.2 Score Generation using Negation

Step 3.3 CNN/K-Means based distance Clustering of the scores:
- Compute review-review similarity using cosine similarity on the extended matrix.
- A probabilistic model, denoted as p(rai, nri) is utilised for generating representation space.

Step 4.1 Contrastive Loss:
- Applying the NTXent Loss values to the clusters.

Step 4.2 Identifying Class for each cluster:
- Attract puts them in same class, repel puts them in unknown class.

Step 4.3 Domain-Specific Tailoring:
- Unsupervised method puts the reviews in same class.
- Unsupervised method requires user guidance

Step 5.1 Comparative Analysis:
- Compare the performance of the scores against general-learned scores.

Step 5.2 Accuracy Assessment:
- Evaluate sentiment analysis accuracy and recommendation relevance.

System Output:
- Final polarity of the statement based on the context, whilst understanding contextual ambiguity.

End

The Collaborative Combination (CC) algorithm incorporates both rating and sentiment rating. A sample of 1000 textual evaluations sourced from Amazon's domains pertaining to TV and Film series was utilized for the purpose of conducting a comparative analysis. The presentation of experimental data within the domain of television and film is focused. The commencement of dataset analysis is initiated. The input datasets consist of textual reviews and numerical ratings. The conversion process commences by conducting pre-processing on textual data to prepare it for textual review input. Opinion words are extracted by the use of specific rules, which are defined for the contextual ambiguous words. The pre-processing endeavor encompasses the tasks of phrase segmentation, part-of-speech (POS) annotation, and the incorporation of custom tags. Traditional sentiment analysis involves the transformation of reviews into numerical ratings. The sentiment analysis of user movie reviews is employed to ascertain a discernible inclination toward specific ratings. A training set is utilized in machine learning to estimate a probabilistic model, denoted as p(rai, nri). The emotion score of each review was computed through the utilization of unsupervised machine learning techniques. This methodology yields quantitative ratings. The creation of an extended user-item matrix is achieved through the merging of the input data.

Fig. 1 illustrates the augmented user-item rating matrix, wherein sentiment ratings are incorporated to reflect user preferences. However, it should be noted that the rating matrix exhibits sparsity. The task of accurately predicting and providing recommendations for dependable products necessitates considerations beyond just ratings. Sentiment analysis is utilized to examine the contextual information present in user ratings. The enlarged user-item matrix is augmented with a sentiment rating derived from sentiment analysis.

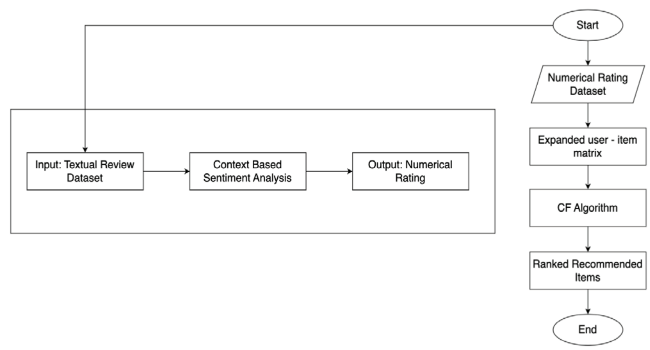


Figure 1 Flow chart for ranked recommendation

The feedback provided by all users is utilized in the computation of suggestions. Typically, a matrix is employed to represent the relationship between users and items.

The items in the user-item matrix are assigned ratings by users. In order to identify the closest neighbors, the neighborhood-based approach initially identifies a cohort of users who exhibit the highest degree of similarity to the active user. The predictive power of active user ratings can be determined by calculating the weighted aggregate of their ratings. The comparison of users is conducted through the utilization of cosine similarity.

The evaluations for second and third-process adjectives are consolidated. Subsequently, the scores of noun and verb phrases obtained from the second stage are utilized. The sentiment lexicon has been carefully calibrated and is responsive to the characteristics of the corpus. The efficacy of the model was enhanced through testing on Amazon TVs, films, and technological devices. The sentiment lexicon generator incorporates two sentiment ratings in addition to the lexicon. The scores of each word are computed and then averaged. The opinion words and negative words are identified and are kept for training the model. The model mentioned in Fig. 1 is used to depend on the opinion words and negative words one by one to create two different scores of the statement showing dependency on the respective kind of words being used in the review.

In order to identify which scoring provides better correlation with the actual score, K means based distance is calculated, via applying the NT-Xent loss, which is a loss function used in Contrastive Learning, which follows the concept of attracting similar scores in one class, and repelling the different scores in different classes, requiring supervision from the user, which thus combines both, supervised learning as well as unsupervised learning. From the attract domain, where the dimensionality is increased to 1000 due to the application of the NT Xent loss, the final polarity of the review can be found, which is now dependent on both, contextual words as well as negative words. The implementation of contextual rules enhances the level of accuracy. The inclusion of valence shifters such as negations, intensifiers, and diminishers yielded enhanced outcomes. In order to evaluate the accuracy of this approach, the lexicons generated using domain sensitivity were compared to the general lexicon. The primary investigation employed a dataset consisting of Movie Review information.

The language developed by Bing Liu served as the baseline for the experiment. The Bing Liu lexicon is a sentiment classification tool that is applicable across several domains. The newly generated vocabulary was evaluated against the existing lexicon in order to assess the performance of the model. The experimental results indicate that the lexicon generated in this work exhibits superior performance compared to general lexicons. This can be attributed to the fact that the lexicon was constructed using a corpus from the same domain. The lexicon was utilized to compute a numerical sentiment score for each review paper.

The electronic product review dataset underwent a similar testing procedure. The accuracy of both models is assessed. The subsequent empirical illustration highlights the significance of employing a sentiment lexicon that is attuned to the specific domain in order to enhance the performance of the model. The classification of the statement "I would suggest that you go read the book instead of watching the movie" as neutral within a general-purpose vocabulary can be attributed to its absence of sentiment-carrying terms. Nonetheless, a vocabulary that is sensitive to the domain of movies and has been specifically designed for this purpose would correctly classify the sentence as having a negative polarity. This is because the writer intends to express a negative sentiment towards the movie by suggesting that the audience should "go read the book" instead. Based on empirical evidence, the lexicon that is attuned to the specific subject in question deems the term "book" to possess negative connotations. A privacy-preserving mechanism used refers to any technique or approach designed to protect individuals' sensitive information while still allowing useful analysis or processing to take place. These mechanisms are crucial in situations where data needs to be shared or analyzed while ensuring that the privacy of individuals is respected and upheld. Privacy-preserving mechanism encompass a broad range of techniques and tools designed to enhance privacy in various contexts. This can include tools for anonymizing web browsing, secure messaging protocols, and privacy-focused operating systems or applications. These mechanisms can be applied in various contexts, including healthcare, finance, social media, and government, to ensure that sensitive data is handled responsibly and ethically while still allowing for meaningful analysis and processing.

The process of domain adaptation involves the refinement of the sentiment lexicon in order to assign domain-specific polarities to phrases, hence enhancing the accuracy of the sentiment analysis model. The numerical scores derived from dataset documents can be utilized as input for collaborative filtering recommender systems subsequent to converting textual data into sentiment ratings. The collaborative filtering system utilized user sentiment to make predictions, rankings, and recommendations for movies. The advantages of proposed model includes a) Contextual Comprehension These embeddings grasp word meanings based on their context within a sentence. b) Ambiguity Resolution: Contextual embeddings help clarify polysemous words by capturing their various senses in different contexts. c) Capturing Word Relationships: Contextual word embeddings capture intricate dependencies between words in a sentence, incorporating both preceding and following words. d) Transfer Learning: Pre-trained on large text datasets using unsupervised methods, contextual word embeddings capture general language patterns and semantics. e) Leading Performance: Contextual word embeddings have demonstrated superior performance on a wide array of NLP benchmarks and tasks, such as sentiment analysis, named entity recognition, question answering, and machine translation.

4 RESULTS

The research utilized datasets from Amazon that consisted of movie ratings and textual evaluations. Additionally, records pertaining to electrical devices sold on Amazon were also employed. Various hyperparameters used in the training Process are Learning rate, Training Epochs/Batch size, Number of layers, Negation Sensitivity, Majority Voting Threshold and Dropout rate. Three different models were employed: the basic ratings-based collaborative filtering (CF) model, an enhanced CF model incorporating sentiment ratings, and a model using contextual sentiment ratings. The experiment baselines include Rating CF and sentiment CF. The experiment utilized the Bing Liu lexicons as the foundational framework. The contextCF will augment the lexicons by including the newly developed domain lexicon. The present research work utilizes benchmark datasets sourced from Amazon. Polarity Identifier system leads to more nuanced interpretations of emotion, whether they have to do with figurative language like irony or with identifying which entities or attributes of entities are truly being discussed in an opinionated remark. This is evident in the development of resources as well. For example, we have observed that polarity has been added to corpora at the sentence level, as opposed to the previous practice of adding it at the entity or global document levels. As a result, the Identifier System must constantly improve in accuracy and sophistication. As a result, they are trying to undertake deeper semantic processing, which has some success with linguistic data that is cleaner.

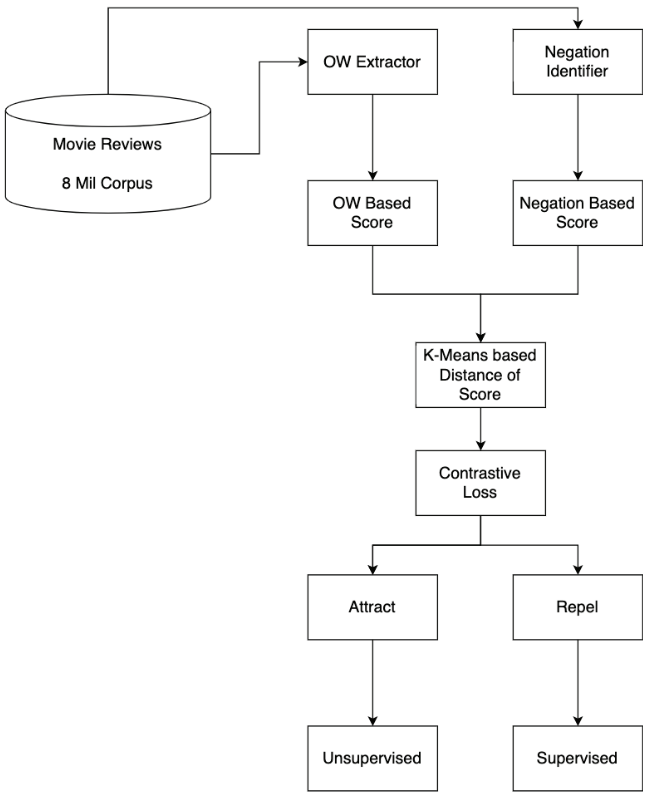


Figure 2 Polarity identifier systems

The experiments were conducted on two distinct domains in order to substantiate the efficacy of the proposed technique. The first dataset has a total of 50,000 ratings and 1000 written reviews, with a scaling range of 1 to 5. These ratings and reviews pertain to various categories such as televisions, movies, and electrical devices. In order to mitigate the risk of overfitting, the dataset is divided into three subsets, namely training, evaluation, and testing data [23].

Table 1 Comparison of the proposed algorithm with RCF & SCF

Model	Amazon TV and Movies			Amazon Electronic Products		
	Accuracy	RMSE	MAE	Accuracy	RMSE	MAE
Rating CF	80.80	3.14	2.47	89.8	2.49	3.34
Sentiment CF	84.35	2.57	3.43	92.7	3.04	3.80
Proposed	90.70	3.45	4.33	99.3	3.95	3.79

The dataset was partitioned into two sets: 80% for training purposes and 20% for testing purposes. The test set comprises items that were subjected to random testing for each user, whereas the training set is utilized to train the recommendation model. The evaluation of proposals generated by a recommender system should be based on their anticipated accuracy and the value they provide to users. The performance of recommender systems is significantly influenced by the level of accuracy they exhibit. The evaluation of recommender system performance is typically conducted using metrics such as Mean Absolute Error (*MAE*) and Root Mean Square Error (*RMSE*). The evaluation of recommender systems commonly employs standard measurements such as *RMSE* and *MAE* to gauge the accuracy of predictions [19].

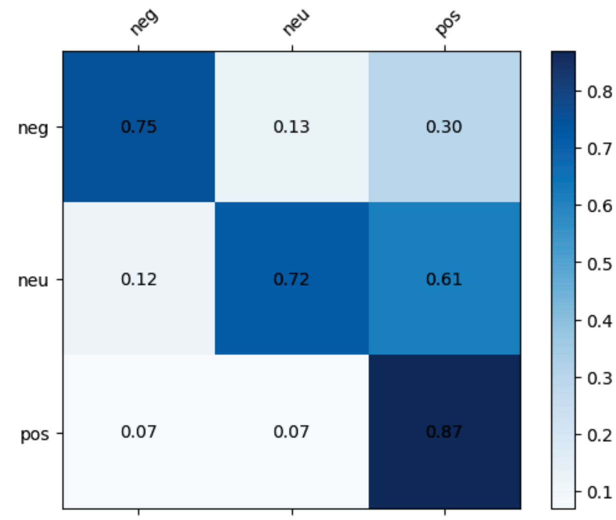


Figure 3 Confusion matrix

5 CHALLENGES

Despite the progress made in addressing contextual ambiguity and negation in sentiment analysis, several challenges remain. These include:

- Contextual Sensitivity: Capturing precise contextual information and understanding nuanced expressions of sentiment remains a challenge. Models that over-rely on context may struggle when the surrounding text is insufficient to determine sentiment.
- Ambiguity Resolution: Resolving ambiguity in sentiment analysis is an ongoing challenge, particularly in creative and informal texts. Handling phrases that can be both positive and negative is a complex task.
- Cross-Linguistic Variability: Different languages and cultures may have unique linguistic nuances, making sentiment analysis in multilingual settings challenging.

Real-Time Sentiment Analysis: In applications like social media monitoring, real-time sentiment analysis is crucial. Balancing the need for rapid analysis with accurate results poses a constant challenge.

Upon examination, it may be concluded that a variety of unique obstacles significantly affect how accurate predictive models are. Sentiment analysis models must be updated to reflect new language trends and issues. Since political discourse is constantly changing. If outdated characteristics are relied upon, sentiment analysis models may become significantly inaccurate. Separating fleeting from persistent sentiment characteristics presents additional challenges that frequently lead to overfitting or the omission of emergent sentiment patterns, which eventually deteriorates model accuracy, particularly when training and test data distributions are different. Due to subjective interpretations and the ambiguity of political attitudes, the addition of human annotation to the active learning process brings an additional layer of variability and potential bias, introducing noise.

6 CONCLUSION AND FUTURE WORK

Contextual ambiguity and negation are critical challenges in the field of sentiment analysis. The complexities are explored that arise when dealing with the interplay of these two factors and examined existing approaches to mitigate them. From machine learning models to sentiment lexicons and rule-based systems, there are diverse strategies that can be employed to handle these challenges. It is clear that context-aware sentiment analysis is essential for accurate sentiment interpretation. As sentiment analysis continues to find applications in various domains, from marketing and customer service to public opinion analysis, addressing contextual ambiguity and negation becomes increasingly vital. The research underscores the need for further research and development in this field.

Given the advancements and challenges identified in the "Negation-Aware Contextual Ambiguity Framework for Enhanced Sentiment Analysis," future research can be directed towards several promising avenues:

1) Integration of Cross-Lingual Capabilities: Expanding the framework to support multiple languages, especially those with complex linguistic structures and varying syntax. This would involve training the model on diverse datasets across different languages, addressing the challenges of negation and contextual ambiguity in a multilingual context.
2) Incorporating Context-Aware Word Embeddings: Further exploration of advanced word embedding techniques, such as contextual embeddings (e.g. BERT, GPT-3), to enhance the understanding of context in sentiment analysis. Future work could involve fine-tuning these models specifically for detecting nuanced sentiment expressions influenced by negation.
3) Dealing with Evolving Slang and Internet Jargon: The dynamic nature of language, especially on social media platforms, poses a continuous challenge. Future research should focus on continuously updating the model to understand evolving slang, abbreviations, and internet jargon which could carry sentiments influenced by negation.
4) Exploring the Role of Emojis and Multimedia Content: As sentiment analysis extends beyond text to include multimedia content, future work could explore integrating analysis of emojis, images, and videos, which often accompany text in digital communication and can influence the interpretation of sentiment.
5) Real-Time Sentiment Analysis Applications: Developing and testing real-time sentiment analysis systems, especially in dynamic environments like social media and customer service platforms, where immediate sentiment interpretation can provide valuable insights for businesses and organizations.
6) Ethical Considerations and Bias Mitigation: Future research must also consider the ethical implications of sentiment analysis, particularly in ensuring that the models do not perpetuate biases present in training data. This involves developing methods to detect and mitigate biases related to gender, race, and cultural background in sentiment analysis models.
7) Explainable AI in Sentiment Analysis: Enhancing the transparency and explainability of AI models used in sentiment analysis. This is crucial for users to trust and understand the basis on which the model makes its sentiment predictions, especially in sensitive applications like mental health analysis and political sentiment tracking.
8) Hybrid Models Combining Rule-Based and ML Approaches: Investigating the synergy between rule-based and machine learning approaches in sentiment analysis to leverage the strengths of both methodologies, especially in handling negation and contextual ambiguities.

7 REFERENCES

[1] Mahany, A., Khaled, H., Nouh, E., Aljohani, N. & Ghoniemy, S. (2022). Annotated Corpus with Negation and Speculation in Arabic Review Domain: NSAR. *International Journal of Advanced Computer Science and Applications, 13*(7). https://doi.org/10.14569/IJACSA.2022.0130706

[2] Reyes, A. & Rosso, P. (2014). Making objective decisions from subjective data: Detecting irony in customer reviews. *Decision Support Systems, 53*(4), 754-760. https://doi.org/10.1016/j.dss.2012.05.027

[3] Poria, S., Cambria, E. & Gelbukh, A. (2016). Aspect extraction for opinion mining with a deep convolutional neural network. *Knowledge-Based Systems*, 108, 42-49. https://doi.org/10.1016/j.knosys.2016.06.009

[4] Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A. & Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1631-1642.

[5] Devlin, J., Chang, M. W., Lee, K. & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv* preprint arXiv:1810.04805.

[6] Pang, B. & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval, 2*(1-2), 1-135. https://doi.org/10.1561/9781601981516

[7] Wiegand, M., Balahur, A., Roth, B., Klakow, D. & Montoyo, A. (2010). A survey on the role of negation in sentiment analysis. *Proceedings of the Workshop on Negation and Speculation in Natural Language Processing*, 60-68.

[8] Councill, I. G., McDonald, R. & Velikovich, L. (2010). What's great and what's not: learning to classify the scope of negation for improved sentiment analysis. *Proceedings of the Workshop on Negation and Speculation in Natural Language Processing*, 51-59.

[9] Cambria, E., Poria, S., Bajpai, R. & Schuller, B. (2013). SenticNet 3: A common and common-sense knowledge base for cognition-driven sentiment analysis. *Proceedings of the AAAI Conference on Artificial Intelligence, 27*(1). https://doi.org/10.1609/aaai.v28i1.8928

[10] Zhang, Y., Jin, R. & Zhou, Z. H. (2018). Understanding AI-Enabled Sentiment Analysis: Techniques, Applications, and Challenges. *Annals of Data Science, 5*(2), 155-177.

[11] Maas, A. L., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y. & Potts, C. (2011). Learning Word Vectors for Sentiment Analysis. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, 142-150.

[12] Liu, B. (2012). Sentiment Analysis and Opinion Mining. *Synthesis Lectures on Human Language Technologies, 5*(1), 1-167. https://doi.org/10.1007/978-3-031-02145-9

[13] Duwairi, R. & Abushaqra, F. (2021). Syntactic-and morphology-based text augmentation framework for Arabic sentiment analysis. *PeerJ Computer Science*, 7, e469. https://doi.org/10.7717/peerj-cs.469

[14] Mahany, A., Khaled, H., Elmitwally, N. S., Aljohani, N. & Ghoniemy, S. (2022). Negation and Speculation in NLP: A Survey, Corpora, Methods, and Applications. *Applied Sciences, 12*(10), 5209. https://doi.org/10.3390/app12105209

[15] Wang, J., Yue, K. & Duan, L. (2023). Models and Techniques for Domain Relation Extraction: A Survey. *Journal of Data Science and Intelligent Systems*. https://doi.org/10.47852/bonviewJDSIS3202973

[16] Maña López, M. J. & Cruz Díaz, N. P. (2019). Negation and Speculation Detection. *Negation and Speculation Detection*, 1-105. https://doi.org/10.1075/nlp.13

[17] Syrett, C., Ebbels, S., Nicoll, H., Davies, K., White, S. & Zuniga-Montanez, C. (2023). Supporting adjective learning by children with Developmental Language Disorder: Enhancing metalinguistic approaches. *International Journal of Language & Communication Disorders, 58*(2), 629-650. https://doi.org/10.1111/1460-6984.12792

[18] Cai, R., Zhang, X. D. & Wang, H. F. (2016). Bidirectional recurrent convolutional neural network for relation classification. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, 1, 756-765. https://doi.org/10.18653/v1/P16-1072

[19] Li, F., Zhang, M. S., Fu, G. H. & Ji, D. H. (2017). A neural joint model for extracting bacteria and their locations. *Proceedings of the 21st Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 15-26. https://doi.org/10.1007/978-3-319-57529-2_2

[20] Zheng, S. C., Wang, F., Bao, H. Y., Hao, Y. X., Zhou, P. & Xu, B. (2017). Joint extraction of entities and relations based on a novel tagging scheme. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, 1, 1227-1236. https://doi.org/10.18653/v1/P17-1113

[21] Duwairi, R. & Abushaqra, F. (2021). Syntactic-and morphology-based text augmentation framework for Arabic sentiment analysis. *PeerJ Computer Science*, 7, e469. https://doi.org/10.7717/peerj-cs.469

[22] Aoumeur, N. E., Li, Z. & Alshari, E. M. (2023). Improving the polarity of text through word2vec embedding for primary classical Arabic sentiment analysis. *Neural processing letters*, 1-16. https://doi.org/10.1007/s11063-022-11111-1

[23] Cumaoğlu, İ. (2022). Sentiment analysis of Arabs in Turkey using deep learning on social media data. *Doctoral dissertation*.

[24] Refai, D. S. (2022). A New Data Augmentation Approach for Boosting Arabic Text Classification Performance Using Transformers and Similarity Measures. *Doctoral dissertation*, Princess Sumaya University for Technology (Jordan).

[25] Díaz, N. P. C. & López, M. J. M. (2019). *Negation and speculation detection (Vol. 13)*. John Benjamins Publishing Company.

Authors' contacts:

Pinky Yadav
(Corresponding author)
FET, MRIIRS (Manav Rachna International Institute of Research and Studies),
Sector – 43, Aravalli Hills, Delhi – Surajkund Road, Faridabad – 121004,
(Haryana), India
pinky.yadavphd@gmail.com

Indu Kashyap, Dr.
FET, MRIIRS (Manav Rachna International Institute of Research and Studies),
Sector – 43, Aravalli Hills, Delhi – Surajkund Road, Faridabad – 121004,
(Haryana), India

Bhoopesh Bhati, Dr.
Indian Institute of Information Technology, Sonepat (IIIT, Sonepat),
IIT Delhi Techno Park, Rajiv Gandhi Education City, Rai, Sonipat, Haryana – 131001,
India

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240425095548

Original scientific paper
Received: 2024-04-25, Accepted: 2024-06-06

Does Ageing Have an Impact on Color Preferences: A Study of Color Preference Trends in Croatia?

Željko Bosančić*, Teo Žeželj, Mile Matijević

Abstract: Attracting and retaining the attention of users, or motivating them to use the particular product being the main task of designers in graphic design for all materials, from printed to digital media. It is crucial to select colors that are more appealing to the target audience. The aim of this paper is to explore the differences in color preferences depending on age groups in a nationally representative sample of the Republic of Croatia consisting of 1,000 research participants. For this purpose, a questionnaire was created to collect information on the color preferences of participants, and as a result of this research, it was determined that there are statistically significant differences in the preferences for certain colors depending on the age groups of the participants. It has also been shown that there are both increasing and stable or decreasing trends in color preferences depending on age groups.

Keywords: color; color preference; communication materials; graphic design; visual communication

1 INTRODUCTION

Graphic communication strives to convey specific messages through carefully selected and harmonized visual elements. To successfully communicate ideas and intended messages, graphic designers use carefully chosen visual elements such as line, shape, color, texture, and typography. Among all the visual elements available to graphic designers, color stands out as one of the fundamental communication channels through which humans interact with their environment. Despite this deep-rooted significance of color in human nature, the semantic interpretation of colors is not universal but can vary significantly depending on numerous cultural, social, and individual factors [1, 2]. The experience of colors is one of the key elements that require special attention when creating visual communications that are intuitively understandable, aesthetically pleasing, and functionally aligned with design objectives. Understanding color preferences is extremely important in graphic design, both for conventional graphic and contemporary digital media, as it plays a crucial role in how viewers interpret visual messages. Through this understanding, designers can better communicate, connect with the target audience, and create visually attractive graphic products. On the other hand, the use of certain carefully selected colors is essential for the visual design. In the case of advertising, the goal is to create visually interesting products where the correct use of colors can significantly reduce cognitive load when using the user interface. This approach also facilitates easier learning for new users, often based on the intuitive use of graphical elements, positively affecting the learning path. Given that colors are one of the fundamental communication channels between humans and their natural environment, it is clear that a deeper understanding of colors can contribute to building systems of human communication and interaction with the environment. Therefore, it is evident that thoughtful and conscious use of colors in graphic design is not just advisable but necessary [3-4]. Given this undeniable importance, it is crucial for graphic designers to understand how people perceive colors and to continuously research market preferences and refine their expertise in the color domain to achieve the desired effect in the designs they create [2].

Color preference, as a concept, refers to individual and collective inclinations towards certain colors or color palettes, and understanding these inclinations can lead to the difference between successfully and unsuccessfully designed communication materials [5-6]. By selecting colors according to user preferences, it is possible to increase brand recognition, improve the perception and understanding of information, and stimulate desired reactions from observers, one of the most important being the retention of attention [6-13]. When creating digital content, designers must consider color semantics to meet user needs and improve the user experience [14].

An individual's age has a significant impact on the preference for certain colors, and color preference changes from early childhood to mature age. Numerous studies have shown that as people age, there are changes in color perception, experience, and preferences [15-20]. Some research has concluded that children often show a marked preference for bright and vivid colors, while older individuals tend to prefer subtler, more neutral shades, and darker tones [21-23]. There are general tendencies for people, regardless of age, to prefer cool colors like blue over warm colors like orange and yellow [17, 24-26].

Research on color preferences has typically been conducted in two ways: through color stimulation, where participants are visually shown a color, and through naming colors, where only the name of the color is used without visual stimulation [17]. In this paper, research will be conducted using color stimulation.

2 RESEARCH DESCRIPTION

For the research, a questionnaire was created consisting of basic demographic questions to monitor the completion rate of the research participant sample, as well as questions about color preferences. Eighteen colors were selected for the examination, which are part of the X-Rite ColorChecker (Fig. 1). These colors were chosen because, in addition to the colors of additive and subtractive synthesis (Red, Green, Blue, Cyan, Magenta, and Yellow), it includes 12 colors that represent the coloring of objects and phenomena in nature, as defined by the description of the tool itself, colors of the most common photographic motifs [27]. Due to the time constraints of the research, it was necessary to rationalize the number of colors on which the research would be conducted, and by the estimation of researchers and authors, this set of colors is useful for understanding color preferences. Also, a peculiarity of this system is that it is used in monitoring color reproduction through digital systems and is adapted for reproduction in the RGB color space.

Figure 1 Display of tested colors with hexadecimal color code

Research participants received the questionnaire via email and completed it on a computer; it was not possible to complete it with phones or tablets. Studies have shown that the results of testing color preferences online, where conditions are not fully controlled (such as screen calibration, lighting, etc.), do not significantly affect the outcomes [24]. Also, the very nature of human-computer interaction, especially in terms of graphical user interfaces, is based on their interaction with personal computers, which significantly contributes to the relevance of the collected data.

The sample of respondents was created in cooperation with Ipsos Croatia, representing a nationally representative sample in the Republic of Croatia, with 1,000 participants carefully distributed by gender, age, and place of residence, to ultimately represent the views on color preferences of the population of the Republic of Croatia. For the purposes of the research, participants aged between 18 and 65 years were selected and divided into three age groups: "younger segment of the population," covering participants aged 18 to 29, "middle segment of the population," covering participants aged 30 to 45, and "older segment of the population," covering participants aged 46 to 65.

After answering the demographic questions, research participants were shown one color on the screen, for which they had to respond how much they liked it among the offered answers "I really like it," "I kind of like it," "I neither like nor dislike it," "I kind of dislike it," and "I really dislike it," as shown in figure 2. After they answered the question, the next color was displayed, and this continued until they had given their response for all 18 tested colors. The colors were presented in a random order, meaning that the sequence of colors shown was different for each research participant, so as not to influence their responses.

Figure 2 Display of color preference question

3 RESULTS ANALYSIS

The study involved 1,000 participants, of which 198 were in the younger segment of the population, 356 in the middle segment, and 446 participants in the older segment, in order to answer the question of whether a person's age affects color preference.

3.1 Statistical Analysis

The data were processed in the JASP 0.18.1.0 statistical analysis software. To check the normality of the data, i.e., whether the data fit or deviate from a normal distribution, it was necessary to conduct the Shapiro-Wilk test, which showed p-values of $< .001$ for all tested data. This value indicates that non-parametric statistical methods are needed for further data analysis. In this case, when comparing more than two groups of independent variables, in this instance, three age groups, the Kruskall Wallis non-parametric test was used. Tab. 1 shows that the Kruskall Wallis test revealed a statistically significant difference between the observed color and age groups for the colors Dark skin ($p < .001$), Light skin ($p < .001$), Foliage ($p = 0.023$), Purplish blue ($p = 0.023$), Moderate red ($p = 0.032$), and Yellow green ($p = 0.022$). For all other observed colors, the test showed that there is no statistically significant difference between the observed colors as the dependent variable and age groups as the independent variable, leading to the conclusion that color acceptance in the Republic of Croatia is consistent across all age groups.

Table 1 Significance of differences in color preferences among age groups where 1 represents the younger segment of the population, 2 - the middle segment of the population, and 3 - the older segment of the population.

Color	Age Group	Mean Rank	Kruskal Wallis	p
Dark Skin	1	2.578	15.497	**<.001****
	2	2.525		
	3	2.827		
Light Skin	1	3.434	18.430	**<.001****
	2	3.042		
	3	3.031		
Blue Sky	1	3.495	0.913	0.634
	2	3.556		
	3	3.594		
Foliage	1	3.091	7.545	**0.023***
	2	2.924		
	3	3.17		
Blue Flower	1	3.505	1.018	0.601
	2	3.638		
	3	3.585		
Bluish Green	1	3.712	2.759	0.252
	2	3.801		
	3	3.684		
Orange	1	3.005	4.696	0.096
	2	2.916		
	3	3.078		
Purplish Blue	1	3.662	7.571	**0.023***
	2	3.89		
	3	3.886		
Moderate Red	1	3.702	6.870	**0.032***
	2	3.539		
	3	3.455		
Purple	1	3.419	1.088	0.581
	2	3.565		
	3	3.509		
Yellow Green	1	3.152	7.615	**0.022***
	2	3.413		
	3	3.444		
Orange Yellow	1	3.071	2.373	0.305
	2	2.997		
	3	3.103		
Blue	1	3.783	5.876	0.053
	2	3.994		
	3	3.973		
Green	1	3.596	1.319	0.517
	2	3.671		
	3	3.706		
Red	1	3.793	0.842	0.656
	2	3.775		
	3	3.688		
Yellow	1	3.359	1.607	0.448
	2	3.435		
	3	3.484		
Magenta	1	3.348	4.803	0.091
	2	3.601		
	3	3.512		
Cyan	1	3.934	4.501	0.105
	2	3.994		
	3	3.843		

*$p < 0.05$, **$p < 0.001$

To determine the nature of the differences between specific age groups, for each color where the Kruskall Wallis test showed a statistically significant difference, a Post-hoc Dunn test with Bonferroni correction was conducted, given that it controls the probability level of Type I error in multiple comparisons. The results of these tests are presented in Tab. 2, and a Box plot graph visualizing the statistically significant differences (graph 1) was created.

Table 2 Dunn post-hoc analysis of the significance of differences in color preferences between age groups (where 1 represents the younger segment of the population, 2 the middle segment of the population, and 3 the older segment of the population)

Color	Age Group	z	p	p_{bonf}
Dark Skin	1-2	0.363	0.717	1.000
	1-3	-2.683	0.007	**0.022***
	2-3	-3.669	<.001	**<.001****
Light Skin	1-2	3.819	<.001	**<.001****
	1-3	4.009	<.001	**<.001****
	2-3	0.052	0.959	1.000
Foliage	1-2	1.500	0.134	0.267
	1-3	-0.713	0.476	0.476
	2-3	-2.728	0.006	**0.019***
Purplish Blue	1-2	-2.484	0.013	**0.039***
	1-3	-2.539	0.011	**0.033***
	2-3	0.047	0.963	1.000
Moderate Red	1-2	1.570	0.116	0.349
	1-3	2.611	0.009	**0.027***
	2-3	1.179	0.238	0.715
Yellow Green	1-2	-2.293	0.022	0.066
	1-3	-2.674	0.008	**0.023***
	2-3	-0.352	0.725	1.000

*$p < 0.05$, **$p < 0.001$

Based on the results of the Post-hoc analysis shown in Tab. 2 and the visualization of statistically significant differences displayed in Fig. 3, it can be observed that the color Dark Skin is more favored by the older segment of the population compared to the younger and middle segments. For the color Light Skin, it is evident that it is more liked by the younger population compared to the middle and older segments.

For the color Foliage, it can be seen that the older segment of the population prefers this color more compared to the middle segment. The results also suggest that the color Purplish Blue is more favored by the middle-aged and older population compared to the younger segment, while the color Moderate Red is more liked by the younger part of the population in comparison to the older segment. Conversely, the color Yellow Green is more favored by the older segment of the population compared to the younger age group.

A closer examination of the data reveals that three of the previously mentioned colors are preferred more or less by certain segments of the population compared to the other two age groups, specifically Dark Skin, Light Skin, and Purplish Blue. The color Dark Skin is more liked by the older segment of the population compared to the younger and middle segments. The color Light Skin is more favored by the younger segment of the population compared to the middle and older segments, while, conversely, the color Purplish Blue is less liked by the younger segment of the population compared to the rest of the population.

3.2 Trend Analysis

For the purpose of analyzing trends across all 18 tested colors, a graph was created that displays the trends of average ratings for each color across the three age groups (Fig. 4).

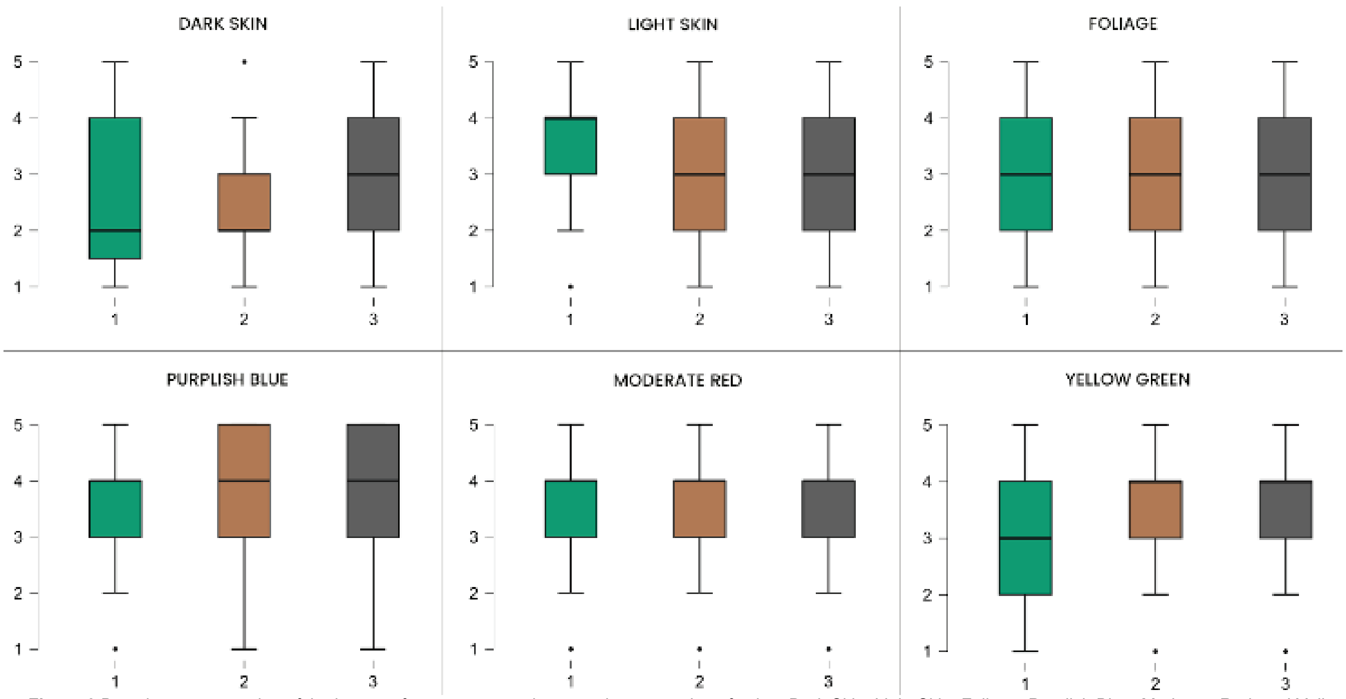


Figure 3 Box plot representation of the impact of age groups on the attractiveness rating of colors Dark Skin, Light Skin, Foliage, Purplish Blue, Moderate Red, and Yellow Green - Visualization of statistically significant differences where the x-axis represents the age groups (1 - younger segment of the population, 2 - middle segment of the population, and 3 - older segment of the population) and the y-axis represents the color preference ratings. **Box plot explanation: Box** represents the interquartile range (IQR) from 25% to 75% of values, **Thicker line** represents median, **Whisker** represents minimum and maksimum value while **Outliers** represents extreme values

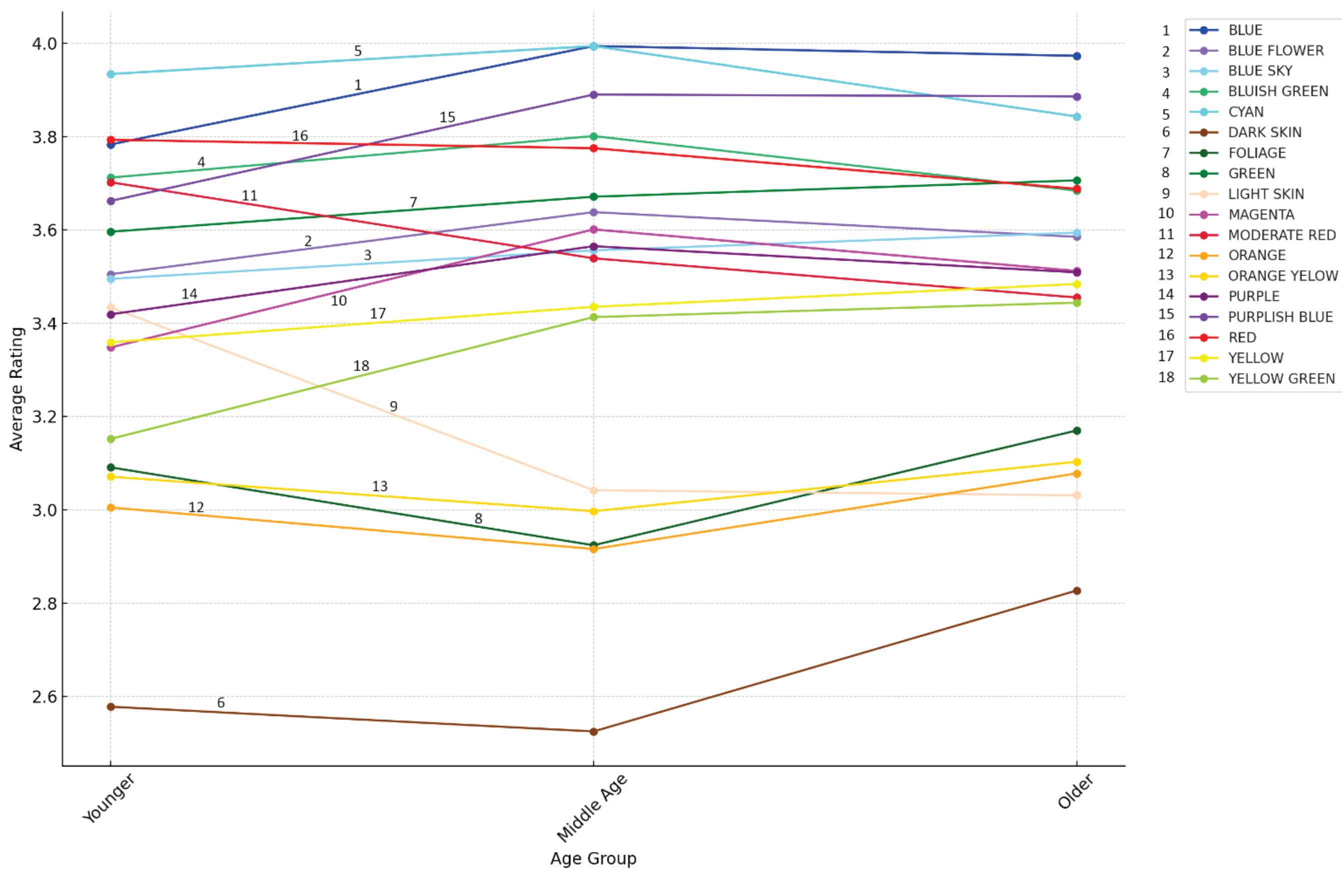


Figure 4 Analysis of color preference trends among age groups

Through the analysis of average color ratings across three age groups (Fig. 4), it can be concluded that general trends in the preference for shades of blue, such as Blue, Blue Sky, Bluish Green, Cyan, and Purplish Blue, indicate they are generally highly rated among age groups with slight variations, suggesting a wide acceptance of blue shades. The colors Blue and Cyan show a specific trend of acceptance and have high average ratings across all age groups, while Blue Flower and Blue Sky show less variation in average ratings among age groups, which may suggest that specific color shades have a more consistent perception across different age groups. Colors containing green shades, such as Green, Yellow Green, and Foliage, have medium to high ratings, reflecting a generally positive perception of the color green. The color Light Skin shows the greatest difference in preference among age groups, with a preference trend significantly declining from the younger to the older part of the population. The results also show that the color Dark Skin is the least accepted of all the colors among all three age groups of research participants.

From the results, it can be inferred that regularities in the data may indicate that there are slight preferential differences among age groups of respondents, but there are also similarities in color perception and acceptance. It can be concluded that preferences for some colors, such as blue and its variants, which typically symbolize peace, trust, stability, and purity [28, 29], show consistency and universality in preferences among all age groups. In contrast, colors containing green shades have medium to high ratings, reflecting a generally positive perception of green color associated with nature, growth, and renewal, while more neutral colors such as Light Skin, Orange, and Orange Yellow show weak preferences within the population. The color Dark Skin, which closely resembles the color Brown, shows general unacceptability across all three age groups of research participants, which is not surprising as studies have shown that people associate it with dirt, mud, decay, etc. [29, 30].

The trend analysis provided several key insights into color preference results relative to the age groups of research participants. Some colors show an increase in popularity with age, while others have more stable ratings or even a decline, indicating that color perception and preference vary with the age group of respondents. Most colors show relatively stable average ratings across age groups, with certain fluctuations suggesting that, although there are personal preferences, there is also a certain consistency in color perception among different age groups in the territory of the Republic of Croatia. Some colors have more pronounced trends, either increasing or decreasing, which may reflect specific cultural, aesthetic, or psychological preferences that change with age.

4 CONCLUSIONS

The aim of the paper was to discover whether differences exist and to analyze trends in color preferences depending on the age of research participants in the Republic of Croatia, and the results showed that there are certain differences in preference influenced by the age of the participants. It was also shown that there are certain increasing and decreasing, as well as stable trends in color preferences depending on the age groups of research participants.

From all the above, certain guidelines can be provided in terms of using specific colors depending on the age groups for whom the designed graphic product is intended. It can be suggested to use shades of blue color, and to a lesser extent shades of green color regardless of age groups, while it is recommended to avoid using colors Light Skin, Orange, and Orange Yellow and similar ones. Colors that show a pronounced declining trend in preference among age groups, such as Moderate Red and Light Skin, could be used when creating content for the younger segment of the population but should be avoided in other age categories. Of course, colors that are not recommended should be used if a graphic solution is being designed for, for example, a product campaign that is precisely such a color, but in that case, it is necessary to choose a combination of colors (e.g., in the background) that will neutralize the negative impact and present the product in a way that it becomes attractive to targeted users.

This research has opened up possibilities for conducting further studies that should include additional social and cultural criteria such as gender, place of residence, place of upbringing, etc., to gain a more detailed and precise insight into color preferences. Additionally, research with color combinations should be conducted to get an insight into to possibly make "unattractive" colors more attractive to targeted users, all in order for designers to gain a better understanding of all aspects of color preferences and their combinations when designing graphics.

5 REFERENCES

[1] Gunther, K. & van Leeuwen, T. (2021). *The Grammar of Visual Design*, 3rd ed. Routlege, London. https://doi.org/10.4324/9781003099857

[2] Landa, R. (2005). *Designing Brand Experience: Creating Powerful Integrated Brand Solutions*. Cengage Learning.

[3] Mohr, C., Jonauskaite, D., Dan-Glauser, E. S., Uusküla, M. & Dael, N. (2018). Unifying research on colour and emotion: Time for a cross-cultural survey on emotion associations with colour terms. *Progress in Colour Studies*, MacDonald, L. W., Biggam, C. P. & Paramei, G. V. Eds., Amsterdam: John Benjamins Publishing Company, 209-222. https://doi.org/10.1075/z.217.11moh

[4] O'Connor, Z. (2015). Colour, contrast and gestalt theories of perception: The impact in contemporary visual communications design. *Color Research & Application, 40*(1), 85-92. https://doi.org/10.1002/col.21858

[5] O'Connor, Z. (2011). Colour psychology and colour therapy: Caveat emptor. *Color Research & Application, 36*(3), 229-234. https://doi.org/10.1002/col.20597

[6] Lewandowska, A. & Olejnik-Krugly, A. (2021). Do Background Colors Have an Impact on Preferences and Catch the Attention of Users? *Applied Sciences, 12*(1), p. 225. https://doi.org/10.3390/app12010225

[7] Gjoni, A. (2022). Color Coordination as a Powerful Design Tool. *ADR, 10*(02), 221-230. https://doi.org/10.4236/adr.2022.102016

[8] Aslam, M. M. (2006). Are You Selling the Right Colour? A Cross-cultural Review of Colour as a Marketing Cue. *Journal of Marketing Communications, 12*(1), 15-30. https://doi.org/10.1080/13527260500247827

[9] Jost, T., Ouerhani, N., Wartburg, R. V., Müri, R. & Hügli, H. (2005). Assessing the contribution of color in visual attention. *Computer Vision and Image Understanding, 100*(1-2), 107-123. https://doi.org/10.1016/j.cviu.2004.10.009

[10] Kawasaki, M. & Yamaguchi, Y. (2012). Effects of subjective preference of colors on attention-related occipital theta oscillations. *NeuroImage, 59*(1), 808-814. https://doi.org/10.1016/j.neuroimage.2011.07.042

[11] Swasty, W. & Adriyanto, A. R. (2017). Does Color Matter on Web User Interface Design. *CommIT (Communication and Information Technology) Journal, 11*(1), p. 17. https://doi.org/10.21512/commit.v11i1.2088

[12] Singh, S. (2006). Impact of color on marketing. *Management Decision, 44*(6), 783-789. https://doi.org/10.1108/00251740610673332

[13] Labrecque, L. I. & Milne, G. R. (2012). Exciting red and competent blue: the importance of color in marketing. *J. of the Acad. Mark. Sci., 40*(5), 711-727. https://doi.org/10.1007/s11747-010-0245-y

[14] Xiaoxiao, B. & Wenming, L. (2018). Research on the application of color semantics in the human-computer interaction design of smartphones. *MATEC Web Conf., 176*, p. 04001. https://doi.org/10.1051/matecconf/201817604001

[15] Ou, L.-C., Luo, M. R., Woodcock, A. & Wright, A. (2004). A study of colour emotion and colour preference. Part I: Colour emotions for single colours. *Color Res. Appl., 29*(3), 232-240. https://doi.org/10.1002/col.20010

[16] Gong, S.-M. & Lee, W.-Y. (2017). Colour preference model for elder and younger groups. *Journal of the International Colour Association, 18*, 33-42.

[17] Dittmar, M. (2001). Changing Colour Preferences with Ageing: A Comparative Study on Younger and Older Native Germans Aged 19–90 Years. *Gerontology, 47*(4), 219-226. https://doi.org/10.1159/000052802

[18] Ou, L.-C., Luo, M. R., Sun, P.-L., Hu, N.-C. & Chen, H.-S. (2012). Age effects on colour emotion, preference, and harmony. *Color Res. Appl., 37*(2), 92-105. https://doi.org/10.1002/col.20672

[19] Owsley, C. (2011). Aging and vision. *Vision Research, 51*(13), 1610-1622. https://doi.org/10.1016/j.visres.2010.10.020

[20] Xi, Y., He, R., Guo, C. & Huang, M. (2019). Colour Difference Discriminations of Young and Old Observers Based on Different Displays. *Proceedings of the 29th Quadrennial Session of the CIE*, Washington DC, USA: International Commission on Illumination, CIE, 917-922. https://doi.org/10.25039/x46.2019.PO014

[21] Frank, M. G. & Gilovich, T. (1988). The dark side of self- and social perception: Black uniforms and aggression in professional sports. *Journal of Personality and Social Psychology, 54*(1), 74-85. https://doi.org/10.1037/0022-3514.54.1.74

[22] Franklin, A., Bevis, L., Ling, Y. & Hurlbert, A. (2010). Biological components of colour preference in infancy. *Developmental Science, 13*(2), 346-354. https://doi.org/10.1111/j.1467-7687.2009.00884.x

[23] Biggam, C. P. (2012). *The Semantics of Colour: A Historical Approach*. Cambridge University Press.

[24] Yu, L., Westland, S. & Li, Z. (2021). Analysis of experiments to determine individual colour preference. *Color Research & Application, 46*(1), 155-167. https://doi.org/10.1002/col.22589

[25] Sengupta, A., Halder, A., Biswas, S., Saha, S. & Dutta, T. (2020). Impact of Age on Color Choice/Preference. *International Journal of Scientific Advances, 1*(2), 92-95. https://doi.org/10.51542/ijscia.v1i2.3

[26] Taylor, C., Schloss, K., Palmer, S. E. & Franklin, A. (2013). Color preferences in infants and adults are different. *Psychon Bull Rev, 20*(5), 916-922. https://doi.org/10.3758/s13423-013-0411-6

[27] McCamy, C. S., Marcus, H. & Davidson, J. G. (1976). A color-rendition chart. *J. Appl. Photogr. Eng., 2*(3), 95-99.

[28] Hanada, M. (2018). Correspondence analysis of color–emotion associations. *Color Research & Application, 43*(2), 224-237. https://doi.org/10.1002/col.22171

[29] Fugate, J. M. B. & Franco, C. L. (2019). What Color is Your Anger? Assessing Color-Emotion Pairings in English Speakers. *Front. Psychol., 10*, p. 206. https://doi.org/10.3389/fpsyg.2019.00206

[30] Palmer, S. E. & Schloss, K. B. (2010). An ecological valence theory of human color preference. *Proc. Natl. Acad. Sci. U.S.A., 107*(19), 8877-8882. https://doi.org/10.1073/pnas.0906172107

Authors' contacts:

Željko Bosančić, MSc
Faculty of Graphic Arts, University of Zagreb,
Getaldićeva ulica 2, 10000 Zagreb, Croatia
zeljko.bosancic@grf.unizg.hr

Teo Žeželj, PhD
Faculty of Graphic Arts, University of Zagreb,
Getaldićeva ulica 2, 10000 Zagreb, Croatia
teo.zezelj@grf.unizg.hr

Mile Matijević, Associate Professor Dr.
Faculty of Graphic Arts, University of Zagreb,
Getaldićeva ulica 2, 10000 Zagreb, Croatia
mile.matijevic@grf.unizg.hr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240502085916

Original scientific paper
Received: 2024-05-02, Accepted: 2024-05-08

Methodology for the Development of Production Systems in the Automotive Industry

Filip Knapp*, Michal Šimon

Abstract: The goal of this research was creation and validation of a methodology for the development of production systems of suppliers of the automotive industry. The theoretical starting points analyze the available literary sources and point to the insufficient treatment of the issue by a comprehensive industry standard representing a gap in the current state of knowledge. The practical part of the research is divided into two stages. The subject of the first stage is the standardization of the methodology through a structured scientific procedure in intensive cooperation with a panel of experts. The subject of the second stage is testing the methodology on authentic projects of the application sector using case studies. The results of the testing showed that in all projects managed using the proposed methodology there was an increase in performance compared to projects managed in a traditional way. The successful confrontation of the proposed solution with reality (practice) is underlined by the confrontation with the current state of knowledge (theory). The practical benefits of the research are the possibility of implementing the methodology into the pedagogical process at training workplaces and, above all, its application into engineering practice to develop new production systems of suppliers of the automotive industry.

Keywords: automotive industry; development; methodology; production system; project management; quality assurance; standardization

1 INTRODUCTION

Over the past few decades, the automotive sector has experienced significant development. There's been a rise in pressure to lower car prices and speed up development time, while at the same time, products are becoming more complex and demanding due to increased competition and higher expectations from consumers. As a result, car makers (OEMs) are increasingly moving ever larger portion of the development and production towards suppliers. Suppliers took over much of the research and development as well as production, leading to an overall increase of up to 80%. This shift has created intricate supplier networks and car production heavily reliant on suppliers. This change has also defined a new set of qualifications for suppliers, who now must take full responsibility for product development, production systems, launch, and meeting the specified requirements and costs set by car manufacturers [15, 16, 22, 35]. This presents new challenges for suppliers.

The key to success is mastering the development stage, including activities from product quality design, through design, implementation and verification of the capability of production systems, to the course of verification production and the release of the process and product into serial production. The reason is the fact that in the development phase there are many more non-conformities (up to 75%) than in the implementation phase and serial production itself.

Moreover, practical experience indicates that addressing non-conformities during the development phase requires significantly lower costs than addressing them during product realization and usage (referred to as The Rule of Ten). By identifying and eliminating non-conformances early, suppliers can minimize cost and resource consumption, leading to improved project outcomes. In this context, modern project management (PM) methods and procedures tailored to efficiently handle intricate, time- and resource-constrained tasks with long durations and high levels of uncertainty, such as development activities, become crucially significant. However, for project management methods and procedures to be truly effective and capable and represent the best current solution, they need to be standardized. Despite today's existence of thousands of suppliers for one car manufacturer (for example, the supply network of the Volkswagen concern includes more than forty thousand suppliers), today there is a lack of a comprehensive industry standard that would serve suppliers as a solid basis for the standardization of development activities. Such a standard would provide automotive industry suppliers with a robust framework for standardizing development activities [35]. The results of the research on the current state of knowledge show that both academic and industrial social groups are quite extensively engaged in the planning and development of new products. However, the same does not apply to studies related to the issue of planning and development of new production systems, which also require extensive care and the mastery of which is considered a strategic competence of enterprises [2, 3]. The primary objective of the research was to address this deficiency in the current state of knowledge and create a comprehensive industry standard that will help automotive industry suppliers effectively manage production system development activities, prevent non-conformities in time, reduce the costs of their elimination and increase the probability of success of new projects.

2 THEORETICAL FRAMEWORK

The initial step of the theoretical part of the research was a traditional examination of the current state of knowledge, the purpose of which was to verify, whether some resource had already dealt with the issue, to what extent and in what place it is possible to follow up on the research. A systematic literature review (SLR) served this purpose. Due to the very limited formal scope of this article, only publications dedicated solely to planning and managing the quality of production systems within the automotive sector are listed. Other relevant sources are mentioned in Chapter 4.1.

2.1 Web of Science

The systematic literature review was launched in the Web of Science (WoS) global scientific database. The search string employed variations of the terms: production / manufacturing, process / system, creation / development, methodology and automotive. The search terms were organized by the use of search (Boolean) operators, parentheses and truncation symbols to maximize the relevance of the outputs. In addition, the research was limited to contributions published since 2010 in order to eliminate out-of-date contributions. Based on the entered search terms and restrictions, the search engine returned 377 articles (gross hits) subsequently sorted according to relevance. After excluding duplications and non-relevant papers, a total of 12 studies remained for a closer examination. This showed that only article entitled "Reference model for the implementation of new assembly processes in the automotive sector" by the authors Emilio C. Baraldi & Paulo C. Kaminski (2018) [4] corresponded to the research objectives in its focus.

2.1.1 Reference Model for the Implementation of New Assembly Process in the Automotive Sector

The analysis of the article showed that the authors were led to the research by the absence of literature and studies dealing with the development of production systems in the automotive sector. The research was thus initiated by the same motive as the research presented in this article. As Baraldi and Kaminski state: "The purpose of the research was to present a consistent and effective reference model for the implementation of new assembly processes in the automotive sector". The study of the article showed that the proposed reference model is the result of the analysis and synthesis of five literary sources that describe with different approaches and objectives the various activities that must occur during the production process development. The five sources are: 1. APQP by AIAG (1995), 2. ABNT ISO/TS 16949 (2010), 3. The product development process (PDP) model by Rozenfeld et al. (2006), 4. VDI 5200 Factory planning (2011) and Proposal of framework to manage the automotive product development process (PDP) by Silva and Kaminsky (2017).

The resulting reference model takes the form of a graphical model embedded in the timeline. The model is vertically divided into two levels. The higher of the two levels (level N) forms a basic axis divided into five phases of different lengths depending on their duration. The individual phases are bounded by quality gates. The lower of the two levels (level N-1) contains the activities that need to be completed within each stage to move from one quality gate to another. The authors list a total of 52 activities. The activities are processed in the form of a simplified Gantt chart and each activity is provided with an acronym explained in a legend placed below the model. The legend lists the full names of phases, quality gates and individual activities, for which it also lists the responsibilities for execution. Individual phases and their content are color coded. The model mentions two auxiliary techniques supporting development activities, which are PFMEA and Control plan. The graphic model is accompanied by a simple text of 1.5 pages briefly describing the content of the individual phases. The paper can be briefly characterized as a basic guide rather than as a full-fledged methodology applicable to the effective management of complex projects.

In an effort to find more relevant articles, the authors decided to change the search string and perform a new search round with variations of the terms: production / manufacturing, process / system, quality, planning and automotive. This time the search engine returned 195 articles (gross hits). The analysis continued in the same way as in the previous round. No further relevant article matching the research needs was however found.

Since the papers of the WoS database fundamentally did not meet the needs of the research, the authors decided to focus on literature dealing with the issue of project management with the stress on automotive industry standards.

2.2 Industry Standards of Project Management

The research of industry standards of PM in automotive industry revealed that they are created by automotive industry trade groups. These groups unite car makers and key suppliers with the aim of exploring shared interests, exchanging practical knowledge, and establishing guidelines to streamline collaboration across all tiers of the industry's supply chains. American Automotive Industry Action Group (AIAG) and German Verband der Automobilindustire (VDA) proved to be the most popular and active worldwide in terms of standardization of PM procedures.

There were 80+ titles discovered in the databases of the AIAG group. The following titles conformed to the research limitations (see chapter 7.2 - categories "Application Sector" and "Scope") and were subjected to closer examination [35]:

- Advanced Product Quality Planning [5;6],
- AIAG & VDA FMEA [7],
- Costs of Poor Quality Guide [8],
- Effective Problem Solving Guide [9],
- Layered Process Audit Guideline [10],
- Measurement System Analysis [11],
- Production Part Approval Process [12], and
- Statistical Process Control [13].

There were 40+ titles discovered in the databases of the VDA group. The following titles conformed to the research limitations and were subjected to closer examination:

- VDA Minimizing risks in the supply chain [14],
- VDA Maturity level assurance for new parts [15],
- VDA2 Quality assurance prior to serial production [16],
- VDA4 Quality assurance in the process landscape [17],
- VDA Customer specific requirements [18],
- VDA 6.3 Process audit [19], and
- VDA 14 Preventive quality management methods in the process landscape [20].

A detailed analysis showed that the only publication with the character of a complex methodology dealing with the quality assurance of manufacturing systems is APQP.

2.2.1 APQP

According to the publication [5], APQP is a standard designed to product and production system quality planning in the early stages of car development, focusing on planning and preparation before serial production. The standard is made in the form of a graphic model, expanding plain text in the scope of 45 pages and appendices. The graphical model consists of phases, quality checks and activities. Horizontally, the model is divided into two levels. The higher of both levels (level N) consists of five overlapping phases. The mentioned stages align with the conventional steps of the Product Creation Process. Throughout these stages, the methodology suggests employing 11 auxiliary tools and techniques (listed in one of the appendices), such as CP, MSA, PFMEA, or PPAP that enhance the likelihood of success in development endeavors. For practicality, the five phases are condensed into five primary activities comprising the N-1 level. The content of each phase is outlined in the methodology through suggested crucial inputs and outputs (49x), where the output of one phase seamlessly feeds into the input of the next phase. Each input and output is subsequently described briefly in plain text. Validation of the procedure's accuracy is facilitated by predefined questionnaires provided in the methodology's appendix. There are 15 checklists in total and they are tied to various key activities such as Control Plan, D-FMEA, Floor Plan, New Equipment, Process Flowchart, P-FMEA, and Product/Process Quality [35].

2.2.2 Remaining AIAG and VDA Publications

The remaining AIAG and VDA publications are recognized as established tools and techniques for ensuring product and production system quality, which can be effectively implemented within the automotive sector [35]. No other truly relevant sources were discovered.

2.3 Research Design and Hypotheses

The results of the SLR showed that the current state of knowledge encompasses various information sources that are highly relevant to crafting a methodology aimed at guaranteeing the quality of production systems developed by suppliers in the automotive industry. These sources differ not only in the time and place of creation and the purpose for which they were created, but above all in their scope and method of processing. The research however confirmed that the current state of knowledge does not offer a comprehensive industry methodology for the development of production systems, which would guide the user through all the important steps of the development stage (= Production System Creation Process = PSCP) and minimize the related project risks. Based on the results of the theoretical part of the research and confirmation of the gap in world knowledge, the objectives of the practical part of the research, the research question and hypotheses were determined.

The main research objectives were defined as:

1) Creation of a methodology for the development of production systems of automotive industry suppliers, and
2) Validation of the methodology in practice on authentic projects of the application sector.

The research question based on the main objectives is: "Will automotive industry suppliers that deploy the proposed methodology for the development of production systems, achieve higher project performance than suppliers that do not deploy the methodology?"

2.3.1 Research Hypotheses

In order to evaluate the impact of the applied methodology on project performance and to answer the research question, the research hypotheses were deliberately formulated to verify or disprove the impact of the methodology on the three main objectives of the project, which are:

1) <u>Time</u> required to achieve the project results,
2) <u>Quality</u> of the project product (main output), and
3) <u>Costs</u> required to achieve the project results.

The research hypotheses are:

H1: Automotive industry suppliers that deploy the proposed methodology for the development of production systems, <u>will achieve a lower number of delayed activities</u> than suppliers that do not deploy the methodology.

H2: Automotive industry suppliers that deploy the proposed methodology for the development of production systems, <u>will achieve a lower number of non-conformities</u> than suppliers that do not deploy the methodology.

H3: Automotive industry suppliers that deploy the proposed methodology for the development of production systems, <u>will achieve lower costs for resolving non-conformities</u> than suppliers that do not deploy the methodology.

The defined hypotheses were supportively related to the proposed methodology and its verification in practice. The following chapter describes the steps that led to the fulfillment of the main research objective #1.

3 PROPOSAL OF CONSTRUCTION OF THE METHODOLOGY

Based on the knowledge gained in the theoretical part of the research, it was decided that the methodology will be processed in the form of a complex model of the Production System Creation Process consisting of two basic forms of notation – PROCESS MAP and PROCESS CARDS. The model will be divided into 4 levels (N - N-3) depending on the purpose and detail they are supposed to treat [35].

The PROCESS MAP was determined as the basis of the model. The process map was inspired by the graphic model of Product Creation Process of the VW concern (including the brands such as Audi, Porsche, Seat, Škoda, or VW), which the authors analyzed in detail within the framework of the SLR [2, 5]. This model visualizes on a single page in a natural way the most important process steps of the car development stage and their interrelationships. Stages, quality gates and milestones play a key role in this model. By way of visualization and the symbols used, this form corresponds to general models of the project life cycle (PJLC) - see, for example, Jan Doležal - Project management [23], or PMI - PMBOK [3]. Such a form of notation is compatible with all key world standards and concepts of project management and at the same time corresponds to the habits of the application sector.

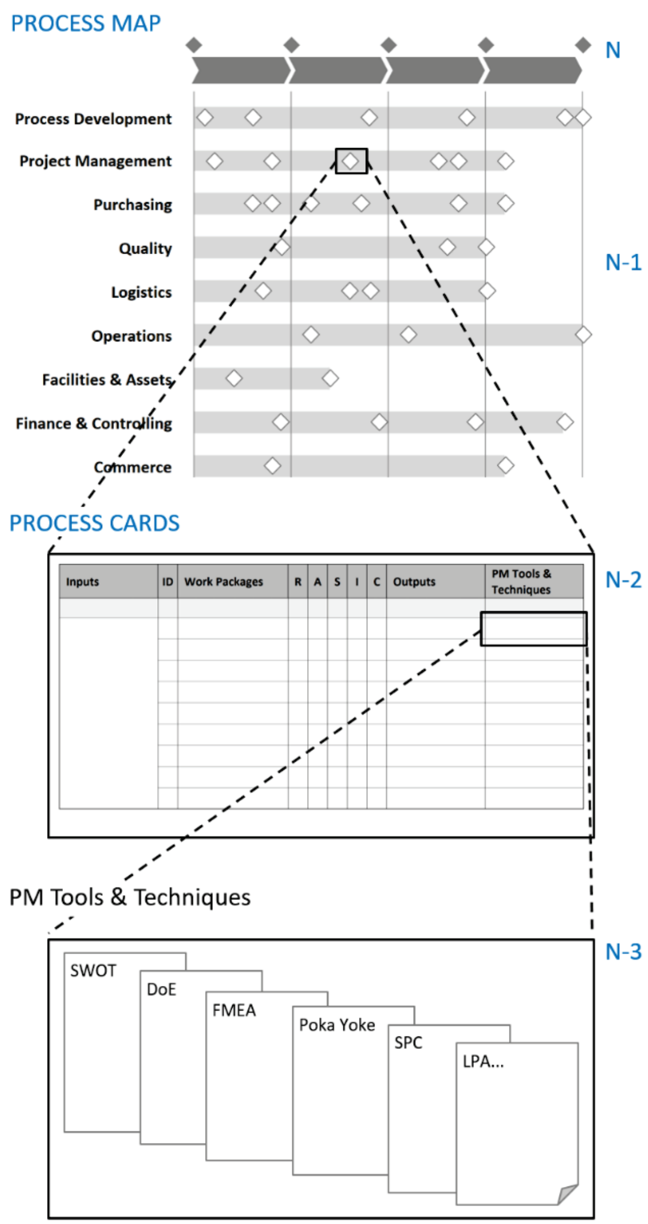


Figure 1 Proposed Model of Production System Creation Process [35]

The proposed process map consists of two levels:

- **N** level forms the main axis of the model divided into the basic stages. The boundaries between the individual stages form quality gates, following the model of the Maturity Level Assurance publication by VDA. The main supplier milestones will come to the axis, which follow on the main milestones of the OEM. These form the basic interface of mutual cooperation.
- **N-1** level makes up the PM success factors of the Production System Creation Process. PM success factors are defined by professional literature as circumstances and influences that lead directly or indirectly to ensuring the quality of the project's product (here production systems), while the correct identification and treatment of these factors will help the success of the project and the achievement of project goals (Doležal, 2016) [23]. Success factors take the form of milestones placed in a matrix according to responsibilities and the required date of fulfillment in the so-called "swimlanes".

The task of the process map is to clearly visualize WHAT is to be done by WHO and WHEN.

In order for the methodology to serve its purpose well, the process map is further complemented by PROCESS CARDS clearly characterizing the individual process steps.

The process cards also consist of two levels:

- **N-2** level consists of the A4 landscape format process cards themselves, which characterize the individual process steps in detail by describing their:
  - Inputs,
  - Process Steps,
  - Responsibilities,
  - Outputs, and
  - PM Tools & Techniques forming N-3 level.
- **N-3** level provides a reference to general PM tools and techniques well applicable to the quality management of automotive supplier production systems. However, the methodology will only refer to these techniques as proven solutions described in detail in separate books, which do not need to be invented or repeated again.

Process Cards are intended to correct deficiencies in the current state of knowledge and show users HOW. See Chapter 6 for a better understanding of the differences in a traditional approach and the proposed methodology.

4 CREATION OF THE METHODOLOGY

The standardization of the methodology consisted of two steps in terms of time and according to the nature of the activities performed. These were the mapping of the business process designated as the PSCP and its subsequent modeling.

4.1 Mapping

Mapping began in the initial phase of the research with the analysis and synthesis of the current state of knowledge.

Scientific method: Analysis and synthesis of the current state of knowledge through systematic literature review

Systematic literature review, also referred to as content analysis of documents, is characterized by experts as a systematic, explicit, and repeatable procedure designed for the identification, evaluation, and synthesis of results created by researchers, academics, and practitioners (Fink, 2014) [24]. Unlike observation or questioning, it does not collect data in the field (field research), but directly evaluates data that is already collected and available (Jesson, 2011) [25].

Selection of resources:

- Selection criteria: content conforms to the research limitations - categories "Application Sector" and "Scope"
- Area: WoS scientific database, literature, database of OEMs organized in VDA (BMW, Daimler, Porsche, VW), database of AIAG and VDA trade groups.

Procedure: The purpose of the review is already described in the introduction of Chapter 2, it was planned for 1 year and was conducted through ICT. Among others, the following sources were analyzed: the three most popular general PM standards - PMBOK Guide [3], IPMA ICB4 [26] and PRINCE2 [27], the Product Creation Process (PCP) of the VW Group [2, 5], or specific requirements of the VW group for its suppliers [18, 28]. Due to a very limited formal scope of this article, the sources are not described in more detail.

Benefit to the research:

- Inspiration for the form and content of the Process map
- Source of 9 key N level milestones
- Source of 30 N-1 level success factors
- Inspiration for the form and content of Process cards
- Source of 200+ level N-3 tools and techniques of PM.

Since the available literary sources did not offer the comprehensive information needed to develop the methodology, especially the success factors of the N-1 level, the authors decided on additional mapping in the field based on cooperation with experts from the application sector who best understand the needs of the industry. Due to the nature and objectives of the research, a Delphi method was chosen as the appropriate method.

Scientific method: Delphi method with a panel of experts.

In a broader sense, the Delphi method can be defined as a process of structuring group communication, in a narrower sense then, as a method of collecting expert opinions through multi-round questioning with controlled feedback between individual rounds. Among the basic features of the method are the anonymity of experts, controlled feedback and statistical determination of the consensus of experts' opinions. The Delphi method can be advantageously applied in areas in which it is necessary to obtain expert opinions on issues that are more difficult to ascertain through statistical analysis or other standard methods. The method is suitable for data collection in the field (field research) [31, 29, 30].

Selection of experts:

- Selection criteria:
  - Has $\geq$ 8 years of experience in project management in the field of production system development in Tier1/2 supplier companies,
  - Holds a professional certificate in project management,
  - Works in middle or senior management.
- Final number: 25 (the number recommended by the methodology is 15-35).

Procedure: The research was aimed at identifying the success factors of the PSCP (see Chapter 3), it was attended by 25 experts selected in accordance with the research limitations and the author of the research in the role of moderator, the research was determined to be three-round, planned for 5 months and conducted via e-mail.

Benefit to the research: Source of 48 N-1 level success factors.

4.2 Modeling

4.2.1 Process Map Modeling

The modeling started with the modeling of the Process Map, which forms the basis of the PSCP model. Due to the nature and objectives of the research, it was determined that the modeling will take place through a group interview with a panel of experts.

Scientific method: Guided semi-structured group interview with a panel of experts.

Group interviews belong to questioning. As the name implies, it is an activity connected with interpersonal contact of the researcher with a selected panel of respondents. On the basis of the discussion, under the guidance of an expert moderator, information is obtained about the opinions of the panelists on the chosen topic - issue. The advantage of the method is not only the ability to capture data, but also to penetrate deeper into the motives and attitudes of the respondents and obtain deeper information about the topic [31]. The method also allows participants to search for majority agreement (consensus) on the issue being addressed, which increases the validity of the outputs [32].

Selection of experts:

- Selection criteria:
  - Comes from a panel of 25 experts who participated in the Delphi survey (conforms to the same criteria),
  - Comes from a different enterprise than the other candidates (to ensure diversity).
- Final number: 7 (the number recommended by the methodology is 6-9).

Procedure:

The research investigation was focused on the modeling of the process map according to the construction design presented in Chapter 3, the interview was attended by 7 experts selected in accordance with the research limitations and the author of the research in the role of moderator, the investigation was determined to be two-round, it was planned for one day and it was placed in the production plant of the author of the research. The participation of 7 selected experts enabled majority agreement of respondents on a solution (= consensus).

Benefit to the research: Creation of a process map.

- **N** level is made up of:
  - the basic axis of the model divided into 4 stages,
  - 5 quality gates RG0-RG4,
  - 9 strategic milestones of the company, largely following on from the main milestones of the OEM.

- **N-1** level is made up of 78 milestones (key success factors) located in swimlanes. Each process step is provided with a unique identifier paired with a corresponding descriptive process card.

4.2.2 Process Cards Modeling

The second modeling step was the modeling of process cards. The processing of process cards was left to the individual enterprises participating in testing the methodology in practice. The reason is simple. As stated by IPMA [26], each organization is original and what works in one may not suit another. However, the authors standardized the process card template and provided specific examples for the initial (#0 - "PJ PRESETS") and final process step (#77 - "HOP") of the PSCP model. In addition, they provided the enterprises with more than 200 project management tools and methods discovered during SLR, which were sufficient to cover the needs of the enterprises

Selection of industrial enterprises: see Chapter 5

Benefit to the research: Creation of 78 process cards.

All the planned steps of the practical part of the research designed to fulfill the main research objective #1 were completed by modeling the PSCP model. The following chapter describes the steps that led to the verification of the validity of the proposed methodology in practice and to the fulfillment of the main research objective #2.

5 VERIFICATION OF THE METHODOLOGY IN PRACTICE

The chapter describes the confrontation of the proposed solution with reality (practice). This took place by testing the methodology on authentic projects of the application sector. Due to the nature and objectives of the research, it was determined that the testing of the methodology will take place through Case Studies.

Scientific method: Case Studies on authentic projects of application sector.

A case study is one of the qualitative research approaches. Unlike a statistical survey, in which a limited amount of data is collected from many individuals (cases), a case study collects a large amount of data from one or a few cases. The method thus enables capturing the complexity, details, relationships and processes taking place in a given microenvironment [33]. It is essential to select the case that will be examined in such a way that it represents a certain type or group of similar cases. Research using case studies takes place in the field (field research) [34].

Selection of industrial enterprises:

- Selection criteria:
- Conforms to the research limitations - category "Application Sector",
- Represents a global enterprise,
- Operates in the Czech Republic,
- Reaches the required maturity of the project management system.

Note: Maturity was tested according to the methodology "Analysis and assessment of project management in organizations" by IPMA. The criteria were set at ≥ 80%.

- Final number: 5

Selection of KPIs:

For monitoring and evaluating the performance of the projects, traditional project indexes evaluating the fulfillment of the scope over time were chosen. The indexes compare the scope that should have been achieved by the evaluation date and the scope that was actually achieved. Performance Index = 1 (equivalent to 100 percentage points) means that the project is running according to plan, $PI > 1$ signals the achievement of a higher scope than originally planned, and $PI < 1$ signals the achievement of a lower scope than originally planned. The nature of the indexes enables mutual comparison of performance of different projects regardless of their scope or character.

Re H1: As stated in Chapter 2.3, the purpose of H1 was to verify or disprove the positive influence of the proposed methodology on **Time**. The selected KPI suitable for verifying or disproving the H1 was Schedule Performance Index, refered to as **SPI**. The source of input data for evaluating project performance through SPI is the Project Master Plan. The project master plan contains all the significant activities of the cross-sectional project team that need to be carried out within the project from its initiation to completion in order to achieve the desired project results. The SPI was calculated according to the following formula:

SPI = actual number of completed activities / planned number of completed activities

Re H2: H2 was intended to verify or disprove the positive influence of the proposed methodology on the **Quality** of the project product (= production system). The selected indicator suitable for verifying or disproving the second hypothesis was the Quality Performance Index, referred to as **QPI**. This indicator expresses the degree of maturity of the production system reflected in the quality of its main output - the product. For this purpose, a key characteristic of the product is systematically monitored, which is directly influenced by the maturity (capability) of the production system from which the product originates. The QPI was calculated according to the following formula:

QPI = actual number of characteristics conforming to requirements / planned number of conforming characteristics

Re H3: H3 was intended to verify or disprove the positive influence of the proposed methodology on **Costs.** The appropriate indicator chosen to verify or disprove the third hypothesis was the Costs Performance Index, referred to as **CPI**. For the purposes of practical testing and verification of the validity of the proposed methodology, the consumption of External Failure Costs was specifically monitored and evaluated. External Failure Costs represent the costs incurred to resolve failures after delivery to the customer. In a narrower sense, the costs indicate, whether the production system has reached the required level of maturity in time and is able to produce the product in accordance with the customer's requirements. In a broader sense, they indicate the state of maturity of the entire project and how well the control mechanisms across the value stream are set to prevent the delivery of non-conforming products to the customer. The source of input data for evaluating project performance through the CPI is the Project Budget, or its sub-item called the External Failure Costs Plan. The CPI was calculated according to the following formula:

CPI = planned cost of external nonconformities / actual cost of external nonconformities

Procedure: The purpose of confronting the proposed solution with reality was to confirm or refute the research hypotheses and answer the research question. Testing took place simultaneously in all 5 companies selected according to the selection criteria. In each company, 2 projects were managed in a traditional way and 2 according to the proposed methodology (20 projects in total). For the sake of mutual comparability, projects of an appropriate scale intended for OEMs organized in the VDA business group with identical customer requirements were selected for testing. The testing was planned for a total of 3 years and was managed through regular meetings of steering committee. By applying the methodology in practice, real data were obtained and the outputs were compared with the assumptions that were considered during the design of the methodology. After the tests were completed, the results of both approaches were averaged and compared to each other.

5.1 Re H1: Verification of the Impact on Time

SPI - E #1: Ø Projects before (0.83) × after (0.93) = Δ +0.10
SPI - E #2: Ø Projects before (0.87) × after (0.93) = Δ +0.06
SPI - E #3: Ø Projects before (0.93) × after (0.94) = Δ +0.01
SPI - E #4: Ø Projects before (0.85) × after (0.91) = Δ +0.07
SPI - E #5: Ø Projects before (0.79) × after (0.90) = Δ +0.11

In terms of verifying or disproving the positive impact on **Time**, the results showed the following:
- For all projects managed with the help of the proposed methodology, there was an improvement in performance compared to projects managed in a traditional way,
- The smallest improvement occurred in Enterprise #3, by an average of 0.01 points (1%),
- The biggest improvement occurred in Enterprise #5, by an average of 0.11 points (11%),
- The overall average improvement was 0.07 points (7%).

Based on the practical testing conducted and based on a comparison of the performance of projects managed in the traditional way and projects managed with the help the proposed methodology, it is possible to state that **hypothesis H1 was CONFIRMED**.

5.2 Re H2: Verification of the Impact on Quality

QPI - E #1: Ø Projects before (0.80) × after (0.93) = Δ +0.13
QPI - E #2: Ø Projects before (0.77) × after (0.94) = Δ +0.17
QPI - E #3: Ø Projects before (0.92) × after (0.96) = Δ +0.04
QPI - E #4: Ø Projects before (0.77) × after (0.94) = Δ +0.17
QPI - E #5: Ø Projects before (0.82) × after (0.90) = Δ +0.08

In terms of verifying or disproving the positive impact on **Quality**, the results showed the following:
- For all projects managed with the help of the proposed methodology, there was an improvement in Performance,
- The smallest improvement occurred in Enterprise #3, by an average of 0.04 points (4%),
- The biggest improvement occurred in Enterprise #2 and Enterprise #4, by an average of 0.17 points (17%),
- The overall average improvement was 0.12 points (12%).

Based on the results of practical testing, it is possible to state the **hypothesis H2 was CONFIRMED**.

5.3 Re H3: Verification of the Impact on Costs

CPI - E #1: Ø Projects before (0.33) × after (0.76) = Δ +0.43
CPI - E #2: Ø Projects before (0.71) × after (0.95) = Δ +0.24
CPI - E #3: Ø Projects before (1.04) × after (1.12) = Δ +0.08
CPI - E #4: Ø Projects before (0.86) × after (1.38) = Δ +0.53
CPI - E #5: Ø Projects before (1.05) × after (3.07) = Δ +2.02

In terms of verifying or disproving the positive effect on **Costs**, the results showed the following:
- For all projects managed with the help of the proposed methodology, there was an improvement in performance,
- The smallest improvement occurred in Enterprise #3, by an average of 0.08 points (8%),
- The biggest improvement occurred in Enterprise #5, by an average of 2.02 points (202%),
- The overall average improvement was 0.66 points (66%),
- Cost experienced more dramatic fluctuations in performance over time than Quality and Time.

Based on the results of practical testing, it is possible to state that the **hypothesis H3 was CONFIRMED**.

The results of testing the methodology in practice showed that all research hypotheses were confirmed and the answer to the research question is YES. The proposed methodology can thus be **declared valid**, despite pseudo-verification on a relatively small sample of Case Studies.

By verifying the validity of the proposed methodology in practice, the main goal of research #2 was fulfilled. A successful confrontation with reality was followed by a confrontation with world knowledge (theory).

6 DISCUSSION

The chapter confronts the proposed solution with the current state of knowledge analyzed in the theoretical part of the research. In order to make the comparison objective, only resources intended primarily to ensure the quality of production systems in the automotive sector are compared.

6.1 Reference Model by Baraldi and Kaminski (2018)

Analysis of the paper showed that the methodology was designed for the purpose of implementation of new production processes in the automotive sector, it is a simple guide through the entire Production System Creation Process (PSCP), it defines responsibilities of participants and it can be tailored to a specific project. It is processed in the form of a simple graphic model providing a very rough instructions on WHAT to do, WHO should do it and WHEN to do it. What cannot be found in the model, is the answer to the question HOW to do it. The model generally exhibits an excesive generalization of content and a significantly narrower scope compared to the sources from which it originates. For comparison, while the APQP methodology describes the process of product and process quality planning on 108 pages, the model by Baraldi and Kaminski fits on 3 pages. Among the auxiliary tools and techniques used in the planning and development of production systems, it mentions two, namely the PFMEA and the Control Plan. These tools are however not elaborated or analyzed in any further way. For the reasons stated above, it can be stated that the paper represents more of a schematic model than a full-fledged methodology, with the help of which it would be possible to effectively manage complex development projects. The publication can from the point of view of fulfilment of the goals of this research be marked with the label "BASIC".

6.2 APQP by AIAG (2008)

A detailed analysis of APQP showed that the methodology was designed to product quality planning in the early stages of car development, it guides the user through the entire Product Creation Process (PCP) and it can be tailored to the needs of a specific project. It is processed in the form of a simple graphic model supplemented by a modest explanatory text, validation of the accuracy of the process is facilitated through predetermined sets of questions provided in the methodology's appendix. The methodology provides the user a basic instructions on WHAT to do, WHEN to do it and WHY. This publication however does not provide information on WHO should perform the task and HOW to do it. Product and process quality assurance is thus looked at more from the point of view of a customer interested in the result, than from the point of view of a supplier who must achieve the result. Additionally, as the name suggests, APQP is a methodology designed primarily for product quality planning. Process quality planning is therefore partially addressed. The methodology also advises utilization of 11 supplementary analytical methods, like CP, FMEA, MSA, PPAP, or SPC, which enhance the likelihood of success of the development activities, but only the Control Plan is further elaborated from these techniques [35]. Based on the above information, it can be concluded that the methodology is intended for more experienced users, who will serve as a basis for standardizing their own PM procedures. The publication can be marked with the label "ADVANCED" from the point of view of fulfilling the research objectives.

6.3 PSCP by Knapp and Šimon (2023)

As can be seen from the paper, PSCP is a full-fledged methodology designed to plan and manage activities associated with the development of production systems in the automotive sector, it is a complex guide in the entire Production System Creation Process (PSCP), it defines the responsibilities of its participants and it can be tailored to the needs of a specific project. It is processed in the form of a clear but detailed PROCESS MAP supported by explanatory PROCESS CARDS providing the users not only instructions on WHAT to do, WHO should do it and WHEN to do it, but also HOW to do it. Process quality assurance is thus not looked at from the point of view of customer interested "just" in the result, but from the point of view of a supplier, who must achieve the results. The methodology also advises utilization of 200+ auxiliary analytical methods, such as DoE, FMEA, LPA, MSA, SPC, or SWOT, which enhance the likelihood of success of the development activities. However, it only refers to these as established tools that are described in separate literature and need not be described again [35]. The methodology is developed in such a way that it can be used even without previous expert experience in the field of PM and is therefore intended not only for experienced users, but also for beginners. The publication can be marked with the label "PROFESSIONAL" from the point of view of fulfilling the research objectives.

A brief description of individual solutions is complemented by a comparison of their scope.

Table 1 Comparison of the scope of individual solutions

		Solution		
Level	Subject	RM	APQP	PSCP
N	Stages Milestones	5 6	5 5	4 9
N-1	Process steps	52	49	78
N-2	Process cards	0	0	78
N-3	Tools and techniques of PM	2	11	200+

A comparison of the proposed methodology with the current state of knowledge showed that all three standards are intended to treat the quality of production systems, are based on a similar philosophy and show common features. Nevertheless, they differ significantly from each other, not only in terms of the time and place of creation and the environment from which they originate, but above all in the way of processing, formal complexity, scope and demands on the professional competence of users.

7 CONCLUSION

The final chapter summarizes the research results, states how the results could be applied in theory and practice, what are its limits and where the research could go next.

7.1 Research Results and Contribution to World Knowledge

The main objectives of the research were: creation of a methodology for the development of production systems of automotive industry suppliers and its validation in practice. The methodology was created through the structured scientific procedure described in the previous text, and the obtained materials were carefully assembled so that the proposed solution corrected the shortcomings of world knowledge and well fulfilled the purpose for which it was created. The proposed solution was tested in practice through case studies. The results of the testing showed that in all projects managed using the proposed methodology there was an increase in performance compared to projects managed in a traditional way. The successful confrontation of the solution with reality (practice) was then underlined by a successful confrontation with world knowledge (theory). The mutual comparison showed that the proposed methodology can be described as evolutionary, not only due to a significantly greater scope and level of detail, but also because it does not demand professional competence of the users, and contains elements that none of the available world knowledge solutions offers. The way of using the created and tested standard depends on the motives and needs of individual users. The following can be cited as an example:

- Training material: "Theory of project management within the automotive industry" intended for training workplaces,
- Methodology applicable in engineering practice in application sector enterprises for managing new projects,
- PM guideline of automotive industry suppliers meeting the requirements of ISO 9001 and IATF 16949 standards.

7.2 Research Limitations

Application sector:

- Automotive industry
- Tier1/Tier2 suppliers
- Serial suppliers (build-to-print).

Scope of the methodology:

- Project management (not program/portfolio)
- Development stage of the car life cycle.

Validity of the methodology:

- Pseudo-verification of the validity of the methodology through Case Studies on a relatively small sample: 5 enterprises, 20 projects.

7.3 Direction of Further Research

The recommendation for the direction of further research smoothly follows on from the results of the current work, which was artificially limited, mainly due to the formal limitation of the scope of the work.

- **Increasing the validity of the methodology**. For the reasons stated in Chapter 5, the validity of the methodology was tested on a relatively small sample of Case Studies (5 enterprises, 20 projects). The authors thus recommend a gradual controlled iteration through the DBR (Design Based Research) methodology based on new ideas, lessons learned (LL) and best practices (BP) obtained during the controlled deployment of the methodology on other projects of the application sector.
- **Increasing the horizontal scope of the methodology**. In its current state, the methodology covers the period of Production System Development. In the case of additional treatment of the period of Product Development (+ 0.8yr), the applicability of the standard can be extended to the entire Development stage of the car's life cycle.
- **Extending the scope of the methodology to program/portfolio management**.

Acknowledgement

The research was funded by the Internal Grant Agency of the University of West Bohemia in Pilsen in the framework of the research project "Environmentally sustainable production", grant number SGS-2023-025.

8 REFERENCES

[1] Nenadál, J. (2018). *Quality management for the 21st century*. Prague, Czech Republic: Management Press.
[2] Wab, G. & Wagner, R. (2017). *Projektmanagement in der Automobilindustrie.* 5th Updated and Revised Edition. Wiesbaden, Germany: Springer Gabler. (in Grman)
[3] Project Management Institute (2017). *A guide to the project management body of knowledge*. Philadelphia, PA: Project Management Institute.
[4] Baraldi, E. C. & Kaminski, P. C. (2018). Reference model for the implementation of new assembly processes in the automotive sector. *Cogent Engineering, 5*(1), 1482984. https://doi.org/10.1080/23311916.2018.1482984
[5] AIAG (2008). *Advanced Product Quality Planning and Control Plan*. Detroit, USA: AIAG.
[6] Stamatis, D. H. (2018). *Advanced product quality planning: the road to success*. Boca Raton, FL: CRC Press/Tailor & Francis Group. https://doi.org/10.1201/9780429401077
[7] AIAG (2019). *AIAG & VDA FMEA Handbook*. Detroit, MI: AIAG.
[8] AIAG (2012). *The Cost of Poor Quality Guide*. Detroit, MI: AIAG.
[9] AIAG (2018). *Effective Problem Solving Guide*. Detroit, MI: AIAG.
[10] AIAG (2014). *Layered Process Audit Guideline*. Detroit, MI: AIAG.
[11] AIAG (2010). *Measurement Systems Analysis*. Detroit, MI: AIAG.

[12] AIAG (2006). *Production Part Approval Process*. Detroit, MI: AIAG.
[13] AIAG (2005). *Statistical Process Control*. Detroit, MI: AIAG.
[14] VDA (2011). *Minimizing risks in the supply chain*. Berlin, Germany: VDA QMC.
[15] VDA (2009). *Product creation - Maturity level assurance for new parts*. Berlin, Germany: VDA QMC.
[16] VDA (2012). *VDA 2 Quality assurance prior to serial production – Production process and product approval*. Berlin, Germany: VDA QMC.
[17] VDA (2021). *VDA 4 Quality assurance in the process landscape*. Berlin, Germany: VDA QMC.
[18] VDA (2018). *Customer specific requirements*. Berlin, Germany: VDA QMC.
[19] VDA (2016). *VDA 6.3 Process audit*. Berlin, Germany: VDA QMC.
[20] VDA (2008). *VDA 14 Preventive quality management methods in the process landscape*. Berlin, Germany: VDA QMC.
[21] Volkswagen AG. (2019). Lieferantenleitfaden für Produktentwicklung. Retrieved from: https://smct-management.de/wp-content/uploads/upload_files/20171222-LF-_V4%20Lieferantenleitfaden%20Produktentwicklung.pdf (in German)
[22] Tecklenburg, G. (2010). Design of automotive body assemblies with distributed tasks under support of parametric associative design (PAD). *Dissertation*, University of Hertfordshire). Retrieved from: https://www.academia.edu
[23] Doležal, J. (2016). *Project management: comprehensively, practically and according to world standards*. Prague, Czech Republic: Grada Publishing.
[24] Fink, A. (2019). *Conducting research literature reviews: from the internet to paper*. Los Angeles, CA: SAGE publications.
[25] Jesson, J., Matheson, L. & Lacey, F. (2011). *Doing Your Literature Review: Traditional and Systematic Techniques*. Los Angeles, CA: SAGE publications.
[26] IPMA. (2015). *Individual competence baseline for project, programme and portfolio management*. Nijkerk, Netherlands: IPMA.
[27] AXELOS. (2017). *Managing successful projects with PRINCE2*. Norwich, United Kingdom: TSO.
[28] Volkswagen AG (2015). *VW99000 - Overall requirements for the performance of component development contracts*. Wolfsburg, Germany: Volkswagen AG, Group Quality Assurance.
[29] Gordon, T. (2009). *Futures Research Methodology - The Delphi Method.* Retrieved from: https://www.millennium-project.org/ publications-2/.
[30] Egerová, D. & Mužík J. (2010). Application of the Delphi method to identify factors influencing effectiveness of e-learning implementation for employee training in small and medium-sized enterprises. *E & M Economics and Managemen*t, 2, 137-152.
[31] Eger, L. & Egerová, D. (2017). *Základy metodologie výzkumu.* Pilsen, Czech Republic: University of West Bohemia. (in Czech)
[32] Chráska, M. (2016). *Metody pedagogického výzkumu. Základy kvantitativního výzkumu.* Prague, Czech Republic: Grada Publishing. (in Czech)
[33] Coombs, H. (2022). Case study research defined - single or multiple? https://doi.org/10.5281/zenodo.7604301
[34] Chrastina, J. (2019). *Case study – a method of qualitative research strategy and research design*. Olomouc, Czech Republic: Palacký University.
[35] Knapp, F. & Šimon. M. (2023). Standardization of Project Management Practices of Automotive Industry Suppliers - Systematic Literature Review. *Technical Journal, 17*(3), 432-439. https://doi.org/10.31803/tg-20230504094426

Authors' contacts:

Filip Knapp, Mgr. Ing. PhD
(Corresponding author)
University of West Bohemia in Pilsen, Faculty of Mechanical Engineering,
Univerzitní 22, 306 14 Pilsen, Czech Republic
Tel.: +420 377638401, E-mail: filip.knapp@seznam.cz

Michal Šimon, doc. Ing. PhD
University of West Bohemia in Pilsen, Faculty of Mechanical Engineering,
Univerzitní 22, 306 14 Pilsen, Czech Republic
Tel.: +420 377638400, E-mail: simon@fst.zcu.cz

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240502091110

Original scientific paper
Received: 2024-05-02, Accepted: 2024-05-08

Research Model of the Degree of Technological Humanism in Manufacturing Companies in the Transformation towards Industry 5.0

Željko Vukelić, Hrvoje Cajner*, Gordana Barić

Abstract: Adopting the concept of Industry 4.0 enables manufacturing companies to gain advantages from digital production management, though issues have been identified related to human resources, management and organisation. Problem areas concern workplace ergonomics to the general wellbeing and satisfaction of workers, and social responsibility. Industry 5.0 promotes progress based on human-machine collaboration and sustainable, human-centred value creation. The research model of the degree of technological humanism, which aims to ensure that technological progress improves human wellbeing in manufacturing companies in the transformation towards Industry 5.0, is a demanding area that has not been sufficiently researched or unambiguously defined so far. The presented model can facilitate the digital transition of manufacturing companies towards Industry 5.0, as it contains the components necessary for the transformation of organisational processes, thereby enhancing and accelerating the implementation of smart manufacturing.

Keywords: ergonomics; social responsibility; technological humanism; transformation towards Industry 5.0

1 INTRODUCTION

Companies today are facing conditions of high volatility and business uncertainty, which can hinder economic growth, and so organisations need to address a number of interconnected organisational changes. This is particularly true for manufacturing companies, and their possible courses of action can be challenging, resulting in problems or opportunities, depending on the approach taken. The current age of industry demands a deep integration of mature technologies and up-and-coming technologies, and the accompanying demands reflect not only on the fusion of information and communication technologies, but also on the convergence of information and operational technologies [1]. The industry sector is also facing increasing pressures to improve its agility and flexibility to better adapt to the highly dynamic production demands [2].

Digital production management has become the imperative for manufacturing companies, and digital transformation affects all areas of value creation [3]. The implementation of smart factory concepts and Industry 4.0 concepts increases company performance [4], including business agility [5], organisational efficiency, efficacy and profitability [6]. Adopting the Industry 4.0 concepts enables companies to gain an advantage from digital production management, though over time issues have been identified relating to human resources, management and organisation. Newer research has shown that hurdles within the company are more important and relevant for the adoption of digital production management than external and operational ones [7]. It has been identified that a hurdle such as a lack of technical skills and expertise is directly related to organisational culture and organisational factors, such as employee leadership, skills and expertise, and not purely technological factors. Accordingly, the greatest challenge for Industry 4.0 is not technology but people, since this concept has a strong impact on employees [8]. For that reason, focus should be placed on the entire organisation as a social-technical system, with consideration of the many interactions between people, technology, business and organisation, since digitalisation and people are still less than ideally synchronised [4], and people are a part of the entire production system [9]. This indicates the need to include a human element in the fundamental concept, which is the integration of human creativity throughout the industrial process [10] and caring for the emotional intelligence of employees as the organisation implementation of automation processes [11]. This is engrained within the new concept Industry 5.0, as an update to Industry 4.0.

2 INDUSTRY 5.0 AND TECHNOLOGICAL HUMANISM

The fundamental area for implementation of the concept Industry 4.0 includes production, logistics, operational management, and business organisation [12-16]. The new concept Industry 5.0 was designed to complement the digital development of industrial systems with purposeful and efficient cooperation between humans, machines and systems in a digital environment [17]. Industry 5.0 thus changes the paradigm and brings solutions that reduce the emphasis on technology, instead placing the focus on the holistic, sustainable creation of human-centric values [18, 19]. Industry 5.0 helps people and machines work together to improve the efficacy of industrial production [18, 17], and represents a tool for enabling sustainable and resilient industry [20], with synergy between humans and technology. With ongoing improvements to technology, humans have had to adjust in order to build a better society, while adopting industrial innovations [10], and Industry 5.0 brings benefits for both workers and society [18], bringing significant progress in quality in relation to the last industrial revolution. This can be considered a response to the demands of the new industrial paradigm that is human-centric, and includes structural, organisational, management, philosophical and cultural aspects, as well as those based on knowledge [21].

Technological humanism is a new concept that arose following the exceptionally rapid penetration of new technologies into all segments of human life and work. Industry 5.0 implements this concept into business and production processes and relationships. It recognises that technology has not supported transformation towards a globally sustainable society by providing tools that are aligned with the logic of pan-humanism, anthropo-relational humanism, and digital humanism in all social and business relationships [22]. Digital humanism describes, analyses and influences the complex interrelations between technology and humans, with the aim of creating a better society, with the complete abidance of universal human rights [23] and it should be interpreted as an interest for understanding the development of digitalisation, with the aim of influencing it technically, socially and politically [24]. In 2016, the Japanese government recognised the technical, social and political dimensions of technological humanism, when it presented the new concept Society 5.0 as an ideal Japan would strive to achieve, in which there is a high degree of convergence between the cyberspace (virtual space) and physical space (real space) [10]. Society 5.0 stands for a new type of society in which innovations in science and technology take a prominent place, in balance with social phenomena and problems that need to be solved, while ensuring economic development [25]. Industry 5.0, as defined by the European Commission, shares many properties with the Society 5.0 concept [26]. Technological humanism therefore is one of the most important new trends in the labour world, in which technology is in service of humans and their development, wellbeing, and fulfilment [27]. In the industrial sphere, technology in service of people has the purpose and goal of promoting social integration and human development, while the purpose of the company is to promote the alliance of all social agents in creating shared values [28]. The benefits of adopting the concepts of Industry 5.0 therefore enable a transition towards a "smart social factory" [18].

3 RESEARCH MODEL FOR DETERMINING THE DEGREE OF TECHNOLOGICAL HUMANISM IN THE TRANSFORMATION TOWARDS INDUSTRY 5.0

Technological humanism is complex in nature and has multiple dimensions. The technical dimension includes the consideration of various disciplines such as philosophy, social sciences, law and economics, and as such as a transdisciplinary concept with social, political and cultural implications of creating and using advanced forms of technology, and testing both the creation of benefits and potential threats [24]. To build a research model to examine the degree of technological humanism within production companies in their transformation towards Industry 5.0, it was first necessary to identify those components of Industry 5.0 in relation with the concept of technological humanism that enable and stimulate the necessary transformational processes for the implementation of smart production.

3.1 Employee Understanding and Knowledge Management

Strategic management of intellectual capital, innovativeness and creativity, technology and information together form the fundamental dimension that has become a trend in organisations, and forms the space for resolving challenges and improving competitiveness in the increasingly complex global context [29]. Today's business conditions place new goals before organisations, impacting their systems, managing information and knowledge, and the human sector. Industry 5.0 is revolutionising production systems around the world by eliminating redundant tasks for workers, and instead enabling development of innovative human potential [18]. Within this concept, people are asked to develop skills and new competencies, both the necessary soft skills and acquiring advanced technical skills that can pose a problem for workers [30]. The development of the human potential of a company should include the process of systematic improvement of knowledge and expertise within the organisation to improve its overall impact, while supporting the degree of understanding and education on the factors that could have an impact on both company operations and on themselves. Company employees who are efficiently informed and appropriately trained in new technologies are able to use them efficiently, which benefits the quality and speed of functioning of the organisation. This enables workers to understand that Industry 5.0 was designed to create strong cooperation, and not competition, between humans and machines [17], which was the greatest problem with previous industrial revolutions. The harmony between the organisation's knowledge management and implementation of Industry 5.0 enables the company to become more agile, effective and sustainable [29], while life-long learning should be a strategic goal of the organisation [31]. In that sense, the management of knowledge and employee understanding are identified as the first component needed to enable the transformation of organisational processes towards smart production and Industry 5.0.

3.2 Organisational and Operational Management

Transformation processes that enable the successful implementation of Industry 5.0 strongly depend on the effectiveness of a company's organisational and operational management. In order to adopt Industry 5.0, it is necessary to create conditions of a high level of interaction between machines and their operators, and the role of the business organisation in adequate decision-making [18]. Workers need to understand new technology to property handle intelligent machines and robot systems, which requires a change in the structure of the labour force, and places increasing demands on employees [32]. Physical and redundant tasks are becoming automated, so the fundamental goal of smart production is for humans and machines to be properly paired so as to increased efficacy of the process of using human brain power and creativity by integrating work processes and intelligent systems [33]. Due to the revolution in industry and society, it is expected that many work posts will be terminated, leading to unemployment and possible severe social issues [34]. The World Economic Forum has assessed that by 2025, 50% of all workers will require requalification due to the adoption of new technologies [35]. There is a need for the employment of highly qualified employees who are open to change and teamwork and are able to transfer knowledge [32], calling for the education of new employees and requalification of existing employees to use new technologies. This presents an upward spiral in which new jobs require new or newly retrained workers, thus creating the new concept Operator 5.0 [36]. It is evident that under such conditions, the company needs to strategically define its organisational and operational management. Digital strategies need to be developed that are focused on the interrelations between digital technology and people at various organisational levels in the processes that shape, transfer, implement, host and support the strategy [37]. Industry 5.0 evidently has a strong impact on work positions, and penetrates deep into every pore of organisational and operational management, setting new rules and patterns for future success. For that reason, organisational and operational management with considerations of the significance of the effects of Industry 5.0 on work posts is the second component of the transformational process aimed at achieving effective and smart production.

3.3 Organisation of an Ergonomic Workplace

The next step is to consider the workplace itself in its ideal form as posited by Industry 5.0, and accordingly to consider the ergonomics of the workplace with the aim of meeting the required conditions. Unlike previous revolutions, the paradigm of Industry 5.0 includes human health and safety as key factors in production systems [38], and following digitalisation, the role of the human operator within it requires changes in ergonomic principles [39]. Ergonomics is a science developed in response to the demands of the industrial revolution [40] and production companies need to treat it appropriately as it has a strong effect on business and production processes and their quality, including its effects on the workers themselves. This is a multidisciplinary science of three fundamental branches that include physics, cognitive and organisational aspects to achieve its goals, by applying other sciences such as physiology, anatomy, engineering, psychology, management and business system design [40]. Industry 5.0 has a tendency towards maximally reducing the physical engagement of workers, in particular problems of a physiological and anatomical nature. The specific effect of digitalisation on people is in its psychological effect and possible disturbances, which is a subject of wide discussion not only in the business world, but also within life in society overall. These disturbances have been defined as Cyber-syndrome, and include physical, social and emotional disturbances that have an effect on humans due to their excessive interactions with cyberspace [40]. Within the concept of Industry 5.0, human interaction with cyberspace is highly intensive and will only increase with the development of new technologies. For this reason, a special subsystem of ergonomics, called Cybergonomics, has been created. Its aim is to optimise security, productivity and health of individuals, and when applied properly, can facilitate Industry 5.0 in achieving its goal of protecting humans from the detrimental effects of new technologies and ensuring the necessary rules and adaptations for better interaction of the labour force and high technology [40]. Adopting advanced technology that requires more time and effort from workers [18] creates a challenge for the activities of organisational ergonomics, which in Industry 5.0 should create a resilient and sustainable business and production system that is human-centric, including the necessary organisational structures, policies and processes. The focus is not only on the need to reexamine existing work posts, but for innovations in those areas and directing the development of human-centric business solutions with new innovations [41]. Organisation of workplace ergonomics in line with the postulates of Industry 5.0 is therefore the third identified component within the company's transformation processes in its transition towards smart production.

3.4 Human Resource Management and Labour Psychology

The processes arising from Industry 5.0 and its influences on the workplace including its organisation and ergonomics certainly affect workplace satisfaction. The company is both a business and a social system, and it is necessary to consider the importance of human health and wellbeing in both these aspects, which is the subject of consideration in human resource management and labour psychology. Since Industry 5.0 places worker wellbeing at the forefront in production processes, new technologies should ensure wellbeing even outside the domain of business and economic growth, so as to become a resilient provider of overall prosperity [42]. Organisational management processes have the demanding task of fulfilling this aim. Companies need to recognise what technology can do for people, and focus on how to adapt technology to the requirements of workers, and not vice versa [18], and to strategically define an appropriate organisational system that enables it. Its efficacy is under the strong influence of workplace factors that affect worker satisfaction and motivation, and possible physical and mental stress at work [43]. An additional goal of digital transformation included in Industry 5.0 is to ensure that people lead purposeful and creative lives [21]. Reducing the need for physical work under conditions of digitalisation and automation enables people to improve their creativity, share ideas, and gain completely new knowledge and abilities for which there was no time or opportunity in the past. With this, technology is shifted from focusing only on improving productivity to creating happy and engaged employees [44]. Accordingly, human resource management must fulfil the conditions from the four main groups that are predictors of a happy and engaged employee: a well-defined role, development of skills and career management, social relations, and support of superiors that include the social factor [45]. Improving the capability of supervision and control in the company, based on the increasing accessibility of data, affects privacy and intimacy of workers and in that sense, it is important that technology deals with the issues of employee autonomy and privacy [18]. Achieving satisfaction at work within the systematic management of human resources and labour psychology within the organisation is therefore the fourth identified component of a successful transition to smart production.

3.5 Managing the Quality Profile of Human Resources

Once the requirements that Industry 5.0 sets for a company are known, the necessary processes and workplace organisation including optimal ergonomics and factors oriented towards achieving employee work satisfaction, the final component is management of the quality profile of employees by human resources, i.e., managing the creation and retention of a high-quality work force. Investing in people is an important determinant of Industry 5.0 as it enables employee retention. Within it, and with the further development of technology, the concept of quality in doing one's work changes in meaning over time. For that reason, companies are faced with the challenge of continuously having a high-quality work force under changing conditions, and these changes are now occurring faster than ever. On the one hand, workers require ongoing training and professional development so that they can develop their careers, making this a key phase in the future quality of the work force [31], while on the other hand, there is the inevitability of the need to replace workers that do not meet the required conditions. If certain workers are not open to change and do not possess the desire for new learning or requalification, the company is forced to hire a new work force. In this, it is important to note that new generations have different values and priorities, which means the company is also required to constantly adapt to the changing conditions on the labour market. Retaining a high-quality workforce requires financial instruments and may raise issues about employee care, visible through the optimal ergonomics of the workplace and a positive influence on their overall work satisfaction. Further, the feeling of belonging to a group and respecting the values of the work performed in the organisation are aspects that can influence organisational culture [46], and worker satisfaction. It can also achieve their explicit inclusion in the design and functioning of business processes [47], which in turn ensures their productivity and efficacy. Worker motivation is a key component of Industry 5.0 [48], and companies need to recognise employee interests in new knowledge and developing their job skills. Only once these and other requirements set by the current business environment are met, can companies expect to retain a high-quality work force and increase the company's attractiveness as a place to work, in line with the values and priorities of new generations.

3.6 Research Components and Model

Based on the above theoretical assumptions, the new research model has been designed to examine the degree of technological humanism with manufacturing companies in their transformation towards Industry 5.0 that contains the following components:

1) Information And Knowledge Management – accessibility and distribution of new information and knowledge
2) Workplace Organisation Management – how positions are organised and processes aligned with the identified elements of change
3) Workplace Ergonomics Management – ergonomics in work positions is aligned with the necessary changes in workplace organisation
4) Workforce Wellbeing Management – achieving worker satisfaction in changing working conditions
5) Long-Term Work Quality Management – ensuring lasting work quality by retaining existing and attracting new high-quality workers.

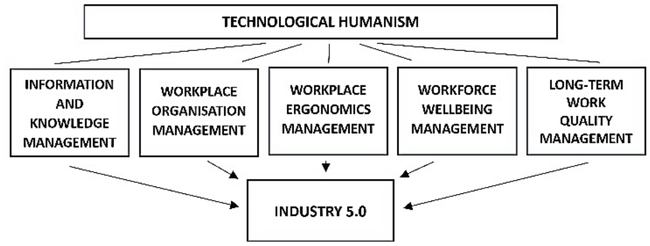


Figure 1 Graphic representation of the new model.

4 CASE STUDY OF RESEARCHING THE DEGREE OF TECHNOLOGICAL HUMANISM

In accordance with the newly created model, a case study was conducted on a sample of employees in two Croatian manufacturing companies to identify the degree of employee knowledge and awareness of Industry 5.0 and its influences on the company and on them personally. The research covered the five components of Industry 5.0 relating to the role of technological humanism, in line with the developed model. Both surveyed companies are small companies with annual revenues in the range of EUR 2–7 million. An additional goal of the case study was to determine whether there is a statistically significant difference in employee perceptions in relation to their work positions, and further to obtain their opinion on how the company could better organise the necessary educational and training processes.

The case study included all company employees, and their work positions were considered in six categories:

1) Auxiliary staff
2) Operational non-production worker
3) Operational production worker
4) Non-production department head
5) Production department head
6) Top management.

The survey questions 1-5 are presented below, and they followed the model components (1-5):

1) To what extent are you familiar with the potential that new technologies can offer the company you work in?
2) In your opinion, to what extent can the application of new technologies affect your work position?
3) In your opinion, to what extent can the application of new technologies affect or could affect the improvement of the ergonomics of your workplace?
4) In your opinion, to what extent can new technologies affect or could affect improvement of your satisfaction at work?
5) In your opinion, to what extent can new technologies affect or could affect your decision to be a long-term employee in your company?

For each question, nine areas addressing the technological aspects of Industry 5.0 were listed:

a) Digital collection of data during production and analysis
b) Automation of production and assembly
c) Use of autonomous robots

d) Automated warehousing systems
e) Real-time digital supervision of machine and device functioning
f) Automated needs planning for raw materials and other materials
g) Business decisions based on digitalised data
h) Use of artificial intelligence (AI) in production, organisation and management
i) Information support for measuring worker performance.

Employees were asked to give answers to the questions based on a five-point Likert scale. Statistical analysis of results was performed to identify possible correlations between factors that influence significant differences. The analysis of these associations was conducted by observing the mean values of responses to the areas a) to i). In total, 60 properly-filled out responses were obtained, with 77% of the total number of employees in the first company and 82% of the total number in the second. Responses from both companies were pooled into a single sample and examined on the basis of work position categories.

Tabs. 1 to 6 give an overview of the descriptive statistics results pertaining to the relationship of the classification determinants relating the type of work position (1-6) and the components of the research model (1-5). Statistically significant values are shown in bold.

Table 1 Total statistical data work position category

Variable	All Groups - Descriptive Statistics							
	Mean	Median	Mode	Freq.of Mode	Minimum	Maximu	Std.Dev.	Skewnes
Component 1	2,50	2,44	Multiple	7	1,00	4,67	0,84	0,51
Component 2	2,84	2,67	2,67	9	1,33	5,00	0,79	0,69
Component 3	2,89	2,89	3,00	8	1,11	4,78	0,82	0,48
Component 4	3,02	2,89	2,78	8	1,22	5,00	0,85	0,51
Component 5	2,95	2,78	Multiple	6	1,00	5,00	0,86	0,48

Table 2 Kruskal-Wallis test results for model component 1

Component 1	Multiple Comparisons p values (2-tailed); Component 1 Independent (grouping) variable: Work position classification Kruskal-Wallis test: H (5, N= 60) =8,453742 p =,1329					
	1	2	3	4	5	6
1		1,00	1,00	1,00	1,00	0,33
2	1,00		1,00	1,00	1,00	0,11
3	1,00	1,00		1,00	1,00	0,16
4	1,00	1,00	1,00		1,00	1,00
5	1,00	1,00	1,00	1,00		1,00
6	0,33	0,11	0,16	1,00	1,00	

Table 3 Kruskal-Wallis test results for model component 2

Component 2	Multiple Comparisons p values (2-tailed); Component 2 Independent (grouping) variable: Work position classification Kruskal-Wallis test: H (5, N= 60) =18,16744 p =,0027					
	1	2	3	4	5	6
1		0,28	0,69	1,00	1,00	1,00
2	0,28		1,00	0,27	1,00	**0,01**
3	0,69	1,00		0,64	1,00	**0,03**
4	1,00	0,27	0,64		1,00	1,00
5	1,00	1,00	1,00	1,00		0,54
6	1,00	**0,01**	**0,03**	1,00	0,54	

It is evident in model components 2, 3 and 4 there was a statistically significant difference in the responses with regard to the position (work category) of survey respondents. Specifically, work category 6, top management, showed significantly difference perceptions in relations to categories 2 and 3 (operational non-production and production workers).

Table 4 Kruskal-Wallis test results for model component 3

Component 3	Multiple Comparisons p values (2-tailed); Component 3 Independent (grouping) variable: Work position classification Kruskal-Wallis test: H (5, N= 60) =16,60810 p =,0053					
	1	2	3	4	5	6
1		0,33	1,00	1,00	1,00	1,00
2	0,33		1,00	1,00	0,12	**0,01**
3	1,00	1,00		1,00	1,00	0,08
4	1,00	1,00	1,00		1,00	1,00
5	1,00	0,12	1,00	1,00		1,00
6	1,00	**0,01**	0,08	1,00	1,00	

Table 5 Kruskal-Wallis test results for model component 4

Component 4	Multiple Comparisons p values (2-tailed); Component 4 Independent (grouping) variable: Work position classification Kruskal-Wallis test: H (5, N= 60) =13,58608 p =,0185					
	1	2	3	4	5	6
1		1,00	1,00	1,00	1,00	1,00
2	1,00		1,00	1,00	1,00	0,06
3	1,00	1,00		1,00	1,00	**0,04**
4	1,00	1,00	1,00		1,00	1,00
5	1,00	1,00	1,00	1,00		1,00
6	1,00	0,06	**0,04**	1,00	1,00	

Table 6 Kruskal-Wallis test results for model component 5

Component 5	Multiple Comparisons p values (2-tailed); Component 5 Independent (grouping) variable: Work position classification Kruskal-Wallis test: H (5, N= 60) =15,29599 p =,0092					
	1	2	3	4	5	6
1		0,57	1,00	1,00	1,00	1,00
2	0,57		1,00	0,86	0,09	0,15
3	1,00	1,00		1,00	0,25	0,43
4	1,00	0,86	1,00		1,00	1,00
5	1,00	0,09	0,25	1,00		1,00
6	1,00	0,15	0,43	1,00	1,00	

Figs. 2 to 4 illustrate the results of the case study for the identified significant differences between the model components 2, 3 and 4.

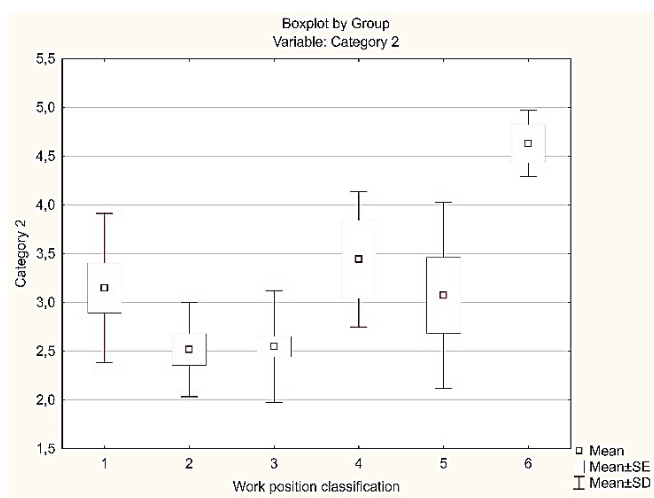


Figure 2 Graphic display of results for model component 2

The case study results indicate significant differences between work positions concerning knowledge and opinions concerning the processes surrounding the transformation of production companies towards Industry 5.0 within the aspect of technological humanism. Particularly in production companies, there is a wide spectrum of employees in different work positions and with varying qualifications and education levels. The results of the case study showed differences in the knowledge, opinion and position on transformation towards Industry 5.0, with statistical significance in one category, top management. This is important since top management is responsible for decision-making and defined policies and company transformation processes towards Industry 5.0. The results of the case study indicate the need to conduct further educational processes and training adapted to the various employee categories, and that more attention should be focused on the development and implementation of specific knowledge transfer programmes, instead of a universal approach to the entire organisation.

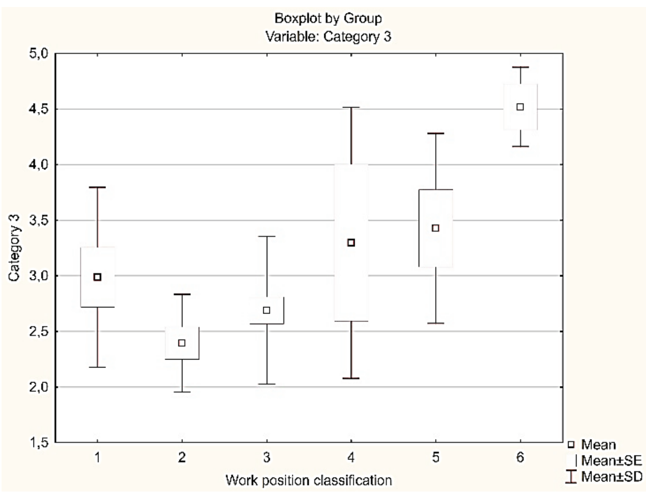


Figure 3 Graphic display of results for model component 3

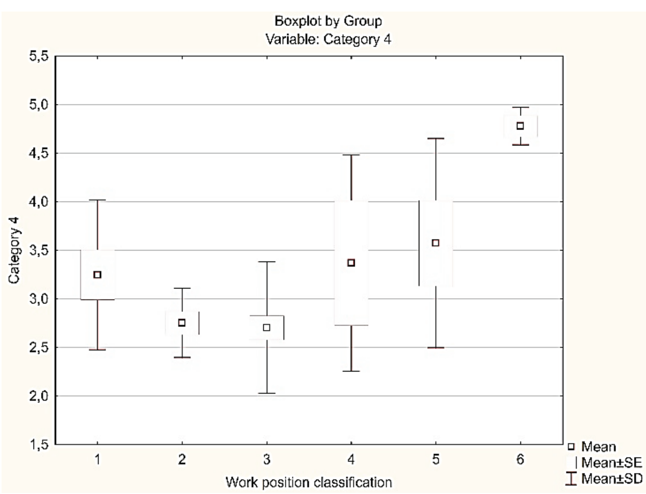


Figure 4 Graphic display of results for model component 4

5 CONCLUSION

Technological humanism is an important component in the transformation of traditional factories into smart ones, with the aim of securing both economic and social benefits. Previous concepts within Industry 4.0 have since been upgraded with the addition of a more human approach to the implementation of digital technologies that are developing at exceptional speeds and penetrating into all spheres of business and society. In that accelerated process, it is easy to neglect the importance of people who are the target group using these means, and instead concentrating on achieving only rapid implementation of new technologies and maximum economic effect. This can function in the short-term, though bringing negative consequences in the mid- and long-term. Technological humanism therefore is a concept that should serve as a guide for the successful transformation of manufacturing companies towards Industry 5.0, with long-term benefits.

This paper presents the new research model to examine the degree of technological humanism within a manufacturing company in transformation towards Industry 5.0, that through its components represents a template that such companies could follow to successfully organise their own business processes and strategic management at all levels. Its components include information and knowledge management, workplace organisation, workplace ergonomics, workforce wellbeing, and long-term work quality, which together form a basis for strategic definition and setting the appropriate organisational systems necessary for successful transformation of manufacturing companies towards Industry 5.0.

The conducted case study showed the differing perceptions of all structures of employees within the categories of technological humanism within Industry 5.0. Top management showed significantly difference perceptions in relations to operational non-production and production workers. The presented results are significant because they enable the creation of appropriate measures that are applicable to different categories of employees in the real sector of production. Primarily there is a need to clearly define the company policy and strategy of the new technologies implementation while informing all employees. In order for the process of transition towards Industry 5.0 to be successful, open communication and highly functional knowledge management that respects all categories of employees is essential. Here it is necessary to take into account that an approach according to employee age and education level is applied. At all organizational levels, it is necessary to promote openness to changes through effective change management. The corresponding activities should include a positive approach with focus on new opportunities, pointing out benefits for workers. There are new and specific tasks for HR management that should be considered to people with different skills and competencies, other than top management. HR departments should consider very seriously that the emergence of new technologies results in the emergence of new job profiles, and some of the existing jobs will disappear. One of the necessary measures that companies should take, especially for workers with a lower level of education and job profile, is the use of positive principles of work psychology. Employees often show big resistance to changes that can result on the lack of knowledge but also human fear of the new. Establishing an organizational culture that promotes constant learning and adaptation to changes is suggested. The task of top management is to foster employee motivation, and educate but also listen to employees.

In manufacturing companies, the aim therefore is to establish changes that are based on knowledge in relation to the work position classification, thereby enabling a successful process of appropriate education and learning that involves all company employees in line with their specific classification characteristics. This can enable the implementation of the smart factory concepts and transformation to Industry 5.0 to positively affect the entire organisation, and lead to business and social progress. These results therefore highlight the significance of technological humanism in today's world of rapid technological development.

6 REFERENCES

[1] Li, K., Zhang, Y., Huang, Y., Tian, Z. & Sang, Z. (2023). Framework and capability of industrial IoT infrastructure for smart manufacturing. *Standards, 3*(1), 1-18. https://doi.org/10.3390/standards3010001

[2] Ryalat, M., ElMoaqet, H. & AlFaouri, M. (2023). Design of a smart factory based on cyber-physical systems and Internet of Things towards Industry 4.0. *Applied Sciences, 13*(4), 2156. https://doi.org/10.3390/app13042156

[3] Klippert, M., Marthaler, F., Spadinger, M. & Albers, A. (2020). Industrie 4.0–An empirical and literature-based study how product development is influenced by the digital transformation. *Procedia CIRP, 91*, 80-86. https://doi.org/10.1016/j.procir.2020.02.152

[4] Koldewey, C., Hobscheidt, D., Pierenkemper, C., Kühn, A. & Dumitrescu, R. (2023). Increasing Firm Performance through Industry 4.0—A Method to Define and Reach Meaningful Goals. *Sci 2022, 4*, 39. https://doi.org/10.3390/sci4040039

[5] Mrugalska, B. & Ahmed, J. (2021). Organizational agility in industry 4.0: A systematic literature review. *Sustainability, 13*(15), 8272. https://doi.org/10.3390/su13158272

[6] Sony, M. (2020). Pros and cons of implementing Industry 4.0 for the organizations: a review and synthesis of evidence. *Production & Manufacturing Research, 8*(1), 244-272. https://doi.org/10.1080/21693277.2020.1781705

[7] Chauhan, C., Singh, A. & Luthra, S. (2021). Barriers to industry 4.0 adoption and its performance implications: An empirical investigation of emerging economy. *Journal of Cleaner Production, 285*, 124809. https://doi.org/10.1016/j.jclepro.2020.124809

[8] Ejsmont, K. (2021). The impact of industry 4.0 on employees—insights from Australia. *Sustainability, 13*(6), 3095. https://doi.org/10.3390/su13063095

[9] Sima, V., Gheorghe, I. G., Subić, J. & Nancu, D. (2020). Influences of the industry 4.0 revolution on the human capital development and consumer behavior: A systematic review. *Sustainability, 12*(10), 4035. https://doi.org/10.3390/su12104035

[10] Dautaj, M. & Rossi, M. (2021). Towards a new society: solving the dilemma between Society 5.0 and Industry 5.0. In *IFIP International Conference on Product Lifecycle Management* (pp. 523-536). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-94335-6_37

[11] Ganer, S. D., Kediya, S. O., Suchak, A. K., Dey, S. K. & Band, G. (2022). Analytical study of HRM practices in industry 5.0. *IOP Conference Series: Materials Science and Engineering, 1259*(1), p. 012041. https://doi.org/10.1088/1757-899X/1259/1/012041

[12] Agca, O., Gibson, J., Godsell, J., Ignatius, J., Davies, C. W. & Xu, O. (2018). *An Industry 4 readiness assessment tool*. Loughborough University.

[13] Çınar, Z. M., Zeeshan, Q. & Korhan, O. (2021). A framework for industry 4.0 readiness and maturity of smart manufacturing enterprises: a case study. *Sustainability, 13*(12), 6659. https://doi.org/10.3390/su13126659

[14] Schumacher, A., Erol, S. & Sihn, W. (2016). A maturity model for assessing Industry 4.0 readiness and maturity of manufacturing enterprises. *Procedia Cirp, 52*, 161-166. https://doi.org/10.1016/j.procir.2016.07.040

[15] Brozzi, R., D'amico, R. D., Pasetti Monizza, G., Marcher, C., Riedl, M. & Matt, D. (2018). Design of Self-assessment Tools to measure industry 4.0 readiness. A methodological approach for craftsmanship SMEs. In *Product Lifecycle Management to Support Industry 4.0: 15th IFIP WG 5.1 International Conference, PLM 2018*, Turin, Italy, July 2-4, 2018. *Proceedings, 15*, 566-578. Springer International Publishing. https://doi.org/10.1007/978-3-030-01614-2_52

[16] Ilanković, N., Živanić, D., Zelić, A., Guban, M. & Szabo, L. (2019). Basic Principles of Industry 4.0 as the Foundation for Smart Factories and Digital Supply Networks. *Proceedings of International Conference "Sustainable Logistics", 4*, 59-64.

[17] Babkin, A., Shkarupeta, E., Kabasheva, I., Rudaleva, I. & Vicentiy, A. (2022). A Framework for Digital Development of Industrial Systems in the Strategic Drift to Industry 5.0. *International Journal of Technology, 13*(7). https://doi.org/10.14716/ijtech.v13i7.6193

[18] Adel, A. (2022). Future of industry 5.0 in society: Human-centric solutions, challenges and prospective research areas. *Journal of Cloud Computing, 11*(1), 1-15. https://doi.org/10.1186/s13677-022-00314-5

[19] Hein-Pensel, F., Winkler, H., Brückner, A., Wölke, M., Jabs, I., Mayan, I. J., ... & Zinke-Wehlmann, C. (2023). Maturity assessment for Industry 5.0: A review of existing maturity models. *Journal of Manufacturing Systems, 66*, 200-210. https://doi.org/10.1016/j.jmsy.2022.12.009

[20] Breque, M., De Nul, L. & Petridis, A. (2021). *Industry 5.0: towards a sustainable, human-centric and resilient European industry*. European Commission, Directorate-General for Research and Innovation. Publications Office of the European Union.

[21] Carayannis, E. G. & Morawska-Jancelewicz, J. (2022). The futures of Europe: Society 5.0 and Industry 5.0 as driving forces of future universities. *Journal of the Knowledge Economy, 13*(4), 3445-3471. https://doi.org/10.1007/s13132-021-00854-2

[22] Hofkirchner, W. (2021). The Future of Anthroposociogenesis: Panhumanism, Anthroporelational Humanism and Digital Humanism. In *Proceedings, 81*(1), p. 114. https://doi.org/10.3390/proceedings2022081114

[23] Werthner, H., Ghezzi, C., Kramer, J., Nida-Rümelin, J., Nuseibeh, B., Prem, E. & Stanger, A. (2024). *Introduction to Digital Humanism: A Textbook*. https://doi.org/10.1007/978-3-031-45304-5

[24] Prem, E. (2024). Principles of digital humanism: A critical post-humanist view. *Journal of Responsible Technology, 17*, 100075. https://doi.org/10.1016/j.jrt.2024.100075

[25] Liliasari, S., Amsad, L. N. & Wahyudi, A. (2021). Innovative chemistry education: An alternative course models in the disruption era. *Journal of Physics: Conference Series, 1731*(1), p. 012023. https://doi.org/10.1088/1742-6596/1731/1/012023

[26] Carayannis, E. G. & Morawska-Jancelewicz, J. (2022). The futures of Europe: Society 5.0 and Industry 5.0 as driving forces of future universities. *Journal of the Knowledge Economy, 13*(4), 3445-3471. https://doi.org/10.1007/s13132-021-00854-2

[27] ESADE. (2020). Humanismo Tecnológico o cómo hacer al ser humano el centro de la automatización global. Conclusiones del Foro de Humanismo Tecnológico. (in Spanish)

[28] Hernando, A. (2022). Humanismo digital: la tecnología al servicio de las personas y del propósito de las empresas. Institute for Social Innovation. Barcelona. (in Spanish)

[29] Molina, R. I. R., Amaris, R. R. A., Raby, N. D. L. & Severino-González, P. (2024). Trends in the knowledge area of organizations in Industry 5.0: perspectives and theoretical references. *Procedia Computer Science, 231*, 571-576. https://doi.org/10.1016/j.procs.2023.12.252

[30] Liu, Y., Yuan, X., Xiong, Z., Kang, J., Wang, X. & Niyato, D. (2020). Federated learning for 6G communications: Challenges, methods, and future directions. *China Communications, 17*(9), 105-118. https://doi.org/10.23919/JCC.2020.09.009

[31] Li, L. (2022). Reskilling and upskilling the future-ready workforce for industry 4.0 and beyond. *Information Systems Frontiers*, 1-16. https://doi.org/10.1007/s10796-022-10308-y

[32] Saniuk, S. & Grabowska, S. (2022). Development of knowledge and skills of engineers and managers in the era of Industry 5.0 in the light of expert research. Zeszyty Naukowe. Organizacja i Zarządzanie/Politechnika Śląska. https://doi.org/10.29119/1641-3466.2022.158.35

[33] Nahavandi, S. (2019). Industry 5.0—A human-centric solution. *Sustainability, 11*(16), 4371. https://doi.org/10.3390/su11164371

[34] Huang, S., Wang, B., Li, X., Zheng, P., Mourtzis, D. & Wang, L. (2022). Industry 5.0 and Society 5.0—Comparison, complementation and co-evolution. *Journal of manufacturing systems, 64*, 424-428. https://doi.org/10.1016/j.jmsy.2022.07.010

[35] Schwab, K. & Zahidi, S. (2020, October). The future of jobs report 2020. In *World economic forum* (Vol. 20).

[36] Romero, D. & Stahre, J. (2021). Towards the resilient operator 5.0: The future of work in smart resilient manufacturing systems. *Procedia CIRP, 104*, 1089-1094. https://doi.org/10.1016/j.procir.2021.11.183

[37] Morton, J., Amrollahi, A. & Wilson, A. D. (2022). Digital strategizing: An assessing review, definition, and research agenda. *The Journal of Strategic Information Systems, 31*(2), 101720. https://doi.org/10.1016/j.jsis.2022.101720

[38] Leng, J., Sha, W., Wang, B., Zheng, P., Zhuang, C., Liu, Q., ... & Wang, L. (2022). Industry 5.0: Prospect and retrospect. *Journal of Manufacturing Systems, 65*, 279-295. https://doi.org/10.1016/j.jmsy.2022.09.017

[39] Trstenjak, M., Hegedić, M., Cajner, H., Opetuk, T. & Tošanović, N. (2023, June). Cognitive Ergonomics in Industry 5.0. In *International Conference on Flexible Automation and Intelligent Manufacturing* (pp. 763-770). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-38165-2_88

[40] Pouyakian, M. (2022). Cybergonomics: Proposing and justification of a new name for the ergonomics of Industry 4.0 technologies. *Frontiers in Public Health, 10*, 1012985. https://doi.org/10.3389/fpubh.2022.1012985

[41] Morawska-Jancelewicz, J. (2022). The role of universities in social innovation within quadruple/quintuple helix model: Practical implications from polish experience. *Journal of the Knowledge Economy, 13*(3), 2230-2271. https://doi.org/10.1007/s13132-021-00804-y

[42] Maddikunta, P. K. R., Pham, Q. V., Prabadevi, B., Deepa, N., Dev, K., Gadekallu, T. R., ... & Liyanage, M. (2022). Industry 5.0: A survey on enabling technologies and potential applications. *Journal of Industrial Information Integration, 26*, 100257. https://doi.org/10.1016/j.jii.2021.100257

[43] Simonetto, M., Arena, S. & Peron, M. (2022). A methodological framework to integrate motion capture system and virtual reality for assembly system 4.0 workplace design. *Safety Science, 146*, 105561. https://doi.org/10.1016/j.ssci.2021.105561

[44] Molino, M., Cortese, C. G. & Ghislieri, C. (2020). The promotion of technology acceptance and work engagement in industry 4.0: From personal resources to information and training. *International journal of environmental research and public health, 17*(7), 2438. https://doi.org/10.3390/ijerph17072438

[45] Salvadorinho, J. & Teixeira, L. (2023). Happy and Engaged Workforce in Industry 4.0: A New Concept of Digital Tool for HR Based on Theoretical and Practical Trends. *Sustainability, 15*(3), 2781. https://doi.org/10.3390/su15032781

[46] Firescu, V., Gaşpar, M. L., Crucianu, I. & Rotariu, E. (2022). Collaboration between Humans and Robots in Organizations: A Macroergonomic, Emotional, and Spiritual Approach. *Frontiers in Psychology, 13*, 855768. https://doi.org/10.3389/fpsyg.2022.855768

[47] Battini, D., Berti, N., Finco, S., Zennaro, I. & Das, A. (2022). Towards industry 5.0: A multi-objective job rotation model for an inclusive workforce. *International Journal of Production Economics, 250*, 108619. https://doi.org/10.1016/j.ijpe.2022.108619

[48] Trstenjak, M., Hegedić, M., Tošanović, N., Opetuk, T., Đukić, G. & Cajner, H. (2022, October). Key Enablers of Industry 5.0-Transition from 4.0 to the New Digital and Sustainable System. In *Global Conference on Sustainable Manufacturing* (pp. 614-621). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-28839-5_69

Authors' contacts:

Željko Vukelić, PhD, BSc
Marius d.o.o.,
Voćarska 44, 10000 Zagreb, Croatia
zeljko.vukelic@marius.hr

Hrvoje Cajner, Assistant professor PhD, BSc
(Corresponding author)
University of Zagreb,
Faculty of Mechanical Engineering and Naval Architecture,
Ivana Lučića 5, 10002 Zagreb, Croatia
hrvoje.cajner@fsb.unizg.hr

Gordana Barić, MSc
University of Zagreb,
Faculty of Mechanical Engineering and Naval Architecture,
Ivana Lučića 5, 10002 Zagreb, Croatia
gordana.baric@fsb.unizg.hr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240502091759

Original scientific paper
Received: 2024-05-02, Accepted: 2024-05-08

A Combined Discrete Event Simulation and Factorial Design Experiment for the Scheduling Problem in a Hybrid Flow Shop

David Ištoković*, Ivana Lučin, Ariana Pešević, Maja Vlatković

Abstract: Production plants have always been confronted with the problem of scheduling, as this has a direct impact on production time and therefore on production costs. This is especially true in today's world, where it is necessary to produce a quality product at a low price and in a short time, while at the same time responding flexibly to customer demands. Due to the complexity and for economic reasons, testing different variants in a real production environment is insufficient and ineffective. For this reason, this paper proposes a new method that combines discrete event simulation and factorial design experiment to find the optimal schedule in hybrid flow shop. It was tested with the goal of achieving the minimum makespan. The results show that this method makes it possible to find improved solution very quickly. Compared to the original production schedule, the makespan could be reduced by 3 hours and 31 minutes which is reduction of 16.3%. The proposed methodology can be used in many discrete and process production plants.

Keywords: Discrete Event Simulation; factorial design experiment; Hybrid Flow Shop; makespan; scheduling; Tecnomatix Plant Simulation

1 INTRODUCTION

Industrial manufacturing has undergone considerable changes in recent decades. These changes affect both the scale and complexity of manufacturing and the technologies used [1, 2]. In order to be competitive, manufacturers must produce high-quality products at a low price and at the same time respond flexibly to the frequently changing needs of customers [3]. This is precisely why simulation programs are increasingly being used today to assess whether potential changes will have a positive or negative impact. A large number of successful companies and enterprises use a variety of simulation programs for planning and production management. The use of such programs makes it possible to view the entire production process at macro and micro level, whereby the simulation can be stopped and restarted at any time [4]. Among other things, testing solutions in a real environment is extremely costly, which manufacturing companies cannot afford [5]. Therefore, the use of simulation programs has proven to be an excellent tool for improving production performance [6].

To achieve the production goals, one of the main tasks is to set up a schedule for production and allocate all the necessary resources - better known as the scheduling problem. Scheduling is about finding the optimal order of execution of certain jobs within the production system under consideration, with the aim of finding an optimal solution - makespan, flow time, tardiness, manufacturing costs, and others [7].

HFS is characterized by the production of medium or large quantities of several different, technologically similar products, whereby the products move unidirectionally through the production process and are processed at more than one stage [8, 9]. Within each stage, there are one or more machines of the same type - parallel machines. On each machine only one operation can be performed simultaneously, and once the operation has been started on the machine, it can no longer be interrupted [10]. The operation of a work task can only be executed only if the previous operations on this work task have been executed [11]. Considering the real-world environment, Hybrid Flow Shop (HFS) is the preferred choice for process and discrete manufacturing companies that engage in custom manufacturing and mass customization, such as in the manufacture of automobiles, machinery, electronics, and other products [12]. The mentioned problem is an NP-hard problem and there is a growing number of scientific papers dealing with its solution. The authors have proposed various techniques, methods and algorithms to solve this type of problem. In their work, Antonova et al [13] proposed an imitation and heuristic method for solving the assigned scheduling problem in a real project company. They use the proposed method to find a schedule that satisfies the time and resource constraints and minimizes the subcontracted resources cost. Janeš et al. use a modified steady-state Genetic Algorithm (GA) to solve batch sizing and scheduling problems with limited buffers in HFS [14]. To find an optimal schedule with the goal of minimizing production costs, Borojević et al. used an integrated CAD/CAPP platform based on elementary machining features and the GA [15]. To solve the scheduling problem in HFS, Rashid & Mu'tasim [16] used an improved Tiki-Taka algorithm which is a novel sport-inspired algorithm based on a football playing style.

Nowadays, one of the most commonly used tools is Discrete Event Simulation (DES), as it allows the creation of a digital model of the observed production system from a real environment and simple modifications to the model without disturbing the operation of the real production system. It also allows the evaluation of the operational characteristics of the existing production system and the prediction of the operational characteristics of the planned production system, where alternative solutions can be compared [2].

With DES, however, it is not possible to find an optimal or sufficiently good solution. Especially not for scheduling problem, which is considered in this paper. For this reason, the combination of DES with other optimization tools is recognized as a promising solution. The authors Ištoković et al. [17] proposed a combination of DES with a GA to find the minimum production cost. The simulation-optimization approach was also applied by Klanke & Engell [18], who used a tailored Evolutionary Algorithm (EA) as an optimizer and a commercial DES as a schedule builder for solving batching and scheduling problem. Although in these approaches better solutions were obtained, but both approaches require a relatively large amount of real-time to find a sufficiently good solution.

Accordingly, this paper proposes a new method based on the application of the probability of occurrence of a work task, i.e. products in the process, combining DES and factorial design experiment with the aim of finding a improved solution in a small amount of real-time. The method is tested for the case of a minimum makespan. So far, several researchers have applied a combination of DES and factorial design experiment [19-22], but not for the case of solving the scheduling problem in a HFS.

The rest of the paper is organised as follows. Section 2 contains a detailed description of the proposed new method based on the probability of product entry into the process for fixed production quantities, combining DES and factorial design experiment to find an improved solution. The results of the conducted research are presented in Section 3, and finally, concluding considerations and suggestions for future research are made in Section 4.

2 METHODOLOGY

2.1 Notation

This paper considers the scheduling problem in a HFS, where the goal is to determine the optimal schedule to complete the production of a given quantity of products as early as possible. It is important that the result - the improved schedule - is achieved in a short period of time.

The parameters used in this paper are as follows:

- N set of jobs/products,
- S set of stages,
- M_i set of identical parallel machines at each stage i,
- j job/product type ($j = 1, 2, \dots, n$),
- i stage/production phase ($i = 1, 2, \dots, s$),
- k machine ($k = 1, 2, \dots, m$),
- o_{ij} operation of product j at stage i,
- p_{ijk} processing time of product j at stage i on machine k,
- q_j production quantity of product j,

2.2 Problem Description

In this paper, three different types of products are produced in the plant (labelled *A*, *B*, and *C*), and the production quantity of each product to be produced is shown in Tab. 1.

Table 1 Types and quantity of products

j	A	B	C
q_j	40	50	30

The observed HFS consists of five stages, within which there are one or more identical parallel machines. More precisely, there is one machine in the first stage, two identical machines in the second stage, three identical machines in the third stage, one machine in the fourth stage and two identical machines in the fifth stage. The number, sequence of operations and processing times for each of the three products listed are given in Tab. 2. Each operation has a fixed duration and is given in minutes. The values given and the proposed production process were taken from a real-world production plant that manufactures parts for industrial machinery.

Table 2 Sequence of operations and processing times

o_{ij}	A	B	C	M_i
1	5	-	5	1
2	8	7	8.5	2
3	-	10	10.5	3
4	18	-	11.5	1
5	-	5	5	2

In this case, the products move unidirectionally through the production process. The availability of the machines is assumed to be 100%, i.e. failures are not taken into account. The transportation between the workplaces (machines) as well as the transportation from the storage of raw materials (SRM) and the transportation to the storage of finished products (SFP) are neglected. The time needed for the setup of the workplaces was also neglected. Additional conditions are that only one operation can be carried out on each machine, and once it has been started, it must not be interrupted. The sequence of operations must be adhered to. An operation on a product can only be carried out until the previous operations have been carried out on the product.

2.3 Simulation Model

The simulation model for the observed HFS was created using Tecnomatix Plant Simulation software version 14.1. This software enables the modeling, simulation, analysis, visualization, and optimization of production systems and processes. It has proven to be very useful and is used by numerous authors in their research to solve production problems [23-25].

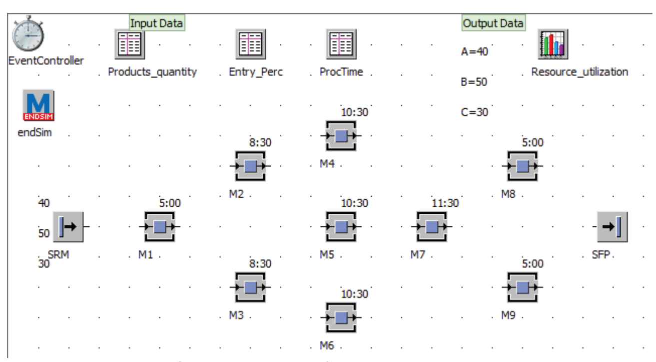


Figure 1 Simulation model of observed production process

The layout of the simulation model with all the necessary elements describing the observed production process is shown in Fig. 1. Every simulation model must be verified and validated to ensure that high-quality and accurate solutions are achieved. According to [26], verification shows whether the created simulation model and all its elements work correctly, while validation shows a satisfactory accuracy range that is consistent with the intended application [27]. For this reason, the necessary elements were included in this simulation model in order to enable simple verification and validation of the simulation model.

The products are initially produced in a cycle, i.e. they are produced in such a way that the entire production quantity of product A is produced first, then the entire production quantity of product B and then the entire production quantity of product C. In such a process, the completion time for the production of these products is 21:33:00. The specified time format is hours:minutes:seconds [hh:mm:ss]. This production strategy was chosen as a starting point because it is the easiest to apply in a real environment. Based on the data obtained from the simulation model, an initial schedule can be seen in Fig. 2.

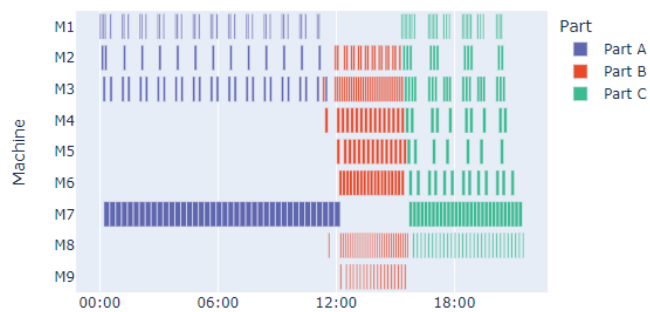


Figure 2 Gantt chart for initial schedule

2.4 Description of the Methodology with Probabilities

A production in which the products are produced in a cycle does not lead to an optimal solution. Therefore, this paper proposes a new method for determining the optimal order of entry of each product into the process based on the probability of each product entering into the process for fixed production quantities. The aim is to minimise the makespan.

Each product type is assigned a specific integer value pn_j according to Eq. (1).

$$pn_j \in [0, 100]. \tag{1}$$

The probability of entry of a certain product type pp_j is then determined according to Eq. (2) as the ratio of the allocation of the pn_j value to the total sum of the allocation values of all product types. In this way, a large search space, i.e. a large number of possible solutions is made possible.

$$pp_j = \frac{pn_j}{pn_1 + pn_2 + \cdots + pn_n}. \tag{2}$$

As an example of a simple case let us assume that the values of pn_j for the three products mentioned (A, B and C) are the same, i.e. $pn_A = pn_B = pn_C = 20$. In this case, the sum of these values is $20 + 20 + 20 = 60$. By calculating according to Eq. (2), we obtain the probabilities for product entry of each product, which are 20/60 = 0.33 or 33%. The given value shows that product A (in this case also B and C) enters the process i.e. it is created at the Storage of Raw Materials (SRM) with a probability of 33%.

Since there is a large search space to find the optimal solution, the optimal probability of entry for a certain product type, the factorial design experiment is applied to find a sufficiently good solution.

2.5 Description of the Factorial Design Experiment and Response Surface Methodology

Factorial experiments are designed in such a way that two or more factors have discrete possible values (levels) so that the effect of each factor on the response value can be studied. It is often used in conjunction with Response Surface Methodology (RSM), which was originally developed to improve manufacturing processes in the chemical industry by optimizing chemical reactions, which was achieved by conducting experiments with different factors. The methodology can be used to optimize any response that is affected by the level of one or more factors. Once the response surface is obtained, it can help to identify the combination of factors considered that can maximize or minimize the target value [28].

In this paper, the specific integer value for products A, B and C were considered in 4 levels, i.e. with values of 20, 40, 60 and 80. This resulted in 4-level 3 factor full factorial design used. The objective was to minimize the production completion time - makespan. Considering the fast execution of the simulations, a full factorial experiment was considered, where the simulation was performed with all possible combinations of stages, leading to 64 results.

The response surface was approximated by a polynomial regression model. The Python module scikit-learn version 0.21.3. was used to implement the regression models. The polynomial orders considered were the second, third and fourth. To validate further the proposed methodology, two additional scenarios with different product quantities were considered (see Tab. 3). In the additional scenarios, the initial production schedule was the same as in the first scenario (first the entire production quantity of product A is produced, then of product B and then of product C).

Table 3 Cases examined for the factorial design experiment with type and quantity of products

Scenario	*A*	*B*	*C*
1	40	50	30
2	175	200	300
3	200	125	145

3 RESULTS

3.1 Simulation Results

The best results obtained with the full factorial approach are shown in Tab. 4, indicating the times and parameters used for the simulation. It can be seen that the full factorial exploration achieves better results in all scenarios.

Table 4 Results of full factorial design experiment

Scenario	1	2	3
Base result [dd:hh:mm:ss]	21:33:00	5:13:53:00	4:06:50:00
Improved result [dd:hh:mm:ss]	18:02:00	4:23:54:00	3:22:01:00
pn_A	40	20/40/40/60/60	60/60/80/80
pn_B	60	40/60/80/80/80	40/80/60/80
pn_C	20	40/80/80/40/80	40/40/60/60
Time reduction [dd:hh:mm:ss]	3:31:00	13:59:00	8:49:00

It should be noted that different combinations of specific integer values pn_j lead to the same results for scenarios 2 and 3. For scenario 2, for example, the two combinations [20,40,40] and [40,60,80], where the numbers are specific integer numbers for product *A*, *B* and *C* respectively, lead to the same production time, i.e. 4:23:54:00. This is due to characteristics of probability calculation, described in more detail in subsection 2.4. Therefore, some of the combinations of variables lead to the same or similar set of parameters.

3.2 Response Surface - Regression Models

For the Scenario 1, the response surfaces for second-, third- and fourth-order regression polynomials are shown in Fig. 3. The figures show areas with minimum values that should be examined more closely in order to possibly obtain better solutions.

Apart from the differences in the shape of the response surface, which can be seen in the figures, the errors of the regression model (Eq. (3)) are calculated as root mean squared error:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}. \tag{3}$$

Where y_i is the actual time of production obtained from the simulation model, $\hat{y}_i$ is the estimated time of production obtained from the regression model, and n is the number of observations, which in this case is 64. The results are shown in Tab. 5. It can be observed that as the polynomial degree increases, the root mean squared error decreases in all cases, which is expected, since greater order polynomial models better fit the data. However, it should be noted that a smaller root mean squared error does not mean a better approximation to the response surface in the optimum area. Additionally, higher order polynomials can lead to overfitting, so the polynomial degree should be chosen according to the problem under consideration.

Table 5 Root mean squared error of regression polynomials

Scenario	Polynomial degree		
	2	3	4
1	19 min	12 min	9 min
2	15 min	13 min	12 min
3	10 min	7 min	6 min

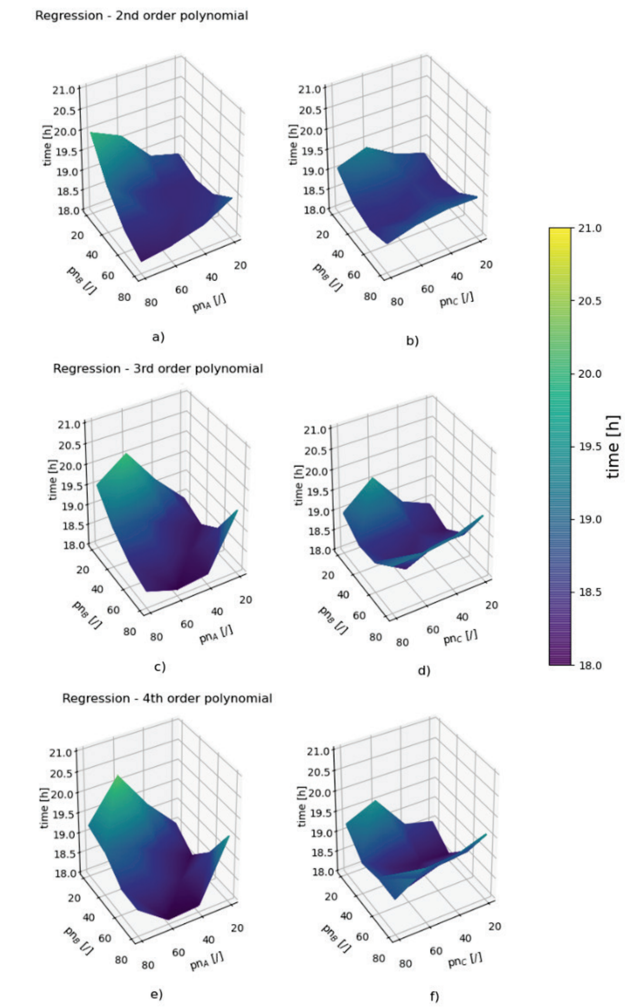


Figure 3 Response surfaces for regression polynomials for: a), b) 2nd order polynomial, c), d) 3rd order polynomial, and e), f) 4th order polynomial

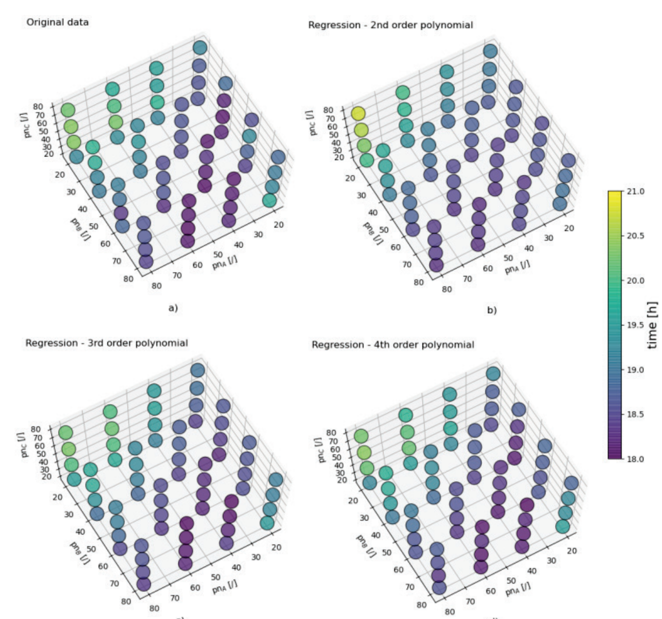


Figure 4 Scatterplot for Scenario 1: Full factorial points for: a) original data, b) 2nd order polynomial, c) 3rd order polynomial, d) 4th order polynomial. The X, Y and Z coordinates represent the specific integer values pn_j. Color represents makespan.

For a better visual representation of the differences in polynomial degree, a scatter plot of the data points is shown in Fig. 4. The colors of the markers indicate the time of the simulation.

3.3 Improved Schedule

Based on the optimal results obtained, a Gantt chart was created for scenario 1, which is shown in Fig. 5 and shows the schedule for the use of production resources (in this case the machines).

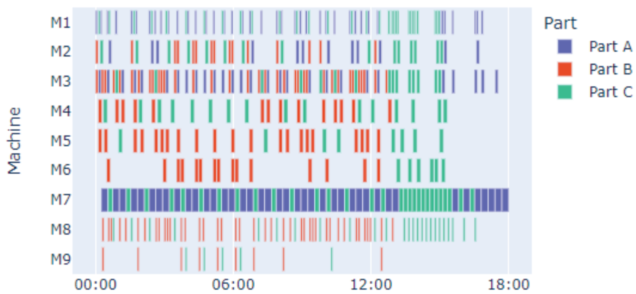


Figure 5 Gantt chart for improved schedule

Compared to Fig. 2, it can be seen that there is no clear pattern for the product entrance in the production process. It can also be observed that the utilisation of the machines is improved, especially for machines M4, M5, M6, M8 and M9, which are mostly not used in the first 12 hours in the initial scenario. It is noted that such a Gantt chart is difficult for the operator to plan without the use of computational strategies such as those used in this paper, as there is no clear pattern for the product entrance. As a result, the production time could be reduced considerably.

4 DISCUSSIONS AND CONCLUSION

One of the most important tasks in ensuring the competitiveness of production companies is to determine production schedules. In contrast to previous achievements in this field, this paper proposes a new method based on the probability of a product entering the production process for fixed production quantities. The proposed method combines DES and factorial design experiment to find the improved (good enough) schedule in HFS with the aim of determining the minimum makespan in this case.

The proposed methodology was successfully applied to the cases investigated. A shorter makespan was achieved in all scenarios compared to the base scenario. In particular, for the observed Scenario 1, the makespan was reduced by 3 hours and 31 minutes, which corresponds to a reduction of 16.3% compared to the original schedule. The results also show that the proposed method provides improved solutions very quickly.

Based on the obtained results of full factorial design response surface was prepared indicating areas for further exploration for obtaining better results. Therefore, considering that only 64 points were used, an optimization procedure of the response surface will be performed in future work to achieve further improvement of results. This can be achieved by increasing the number of levels for the factors or by additional evaluations close to the current best solutions. It should be noted that with longer times required to simulate the production process, the full factorial design could require a significant amount of time. Therefore, a fractional factorial design can also be investigated in future work. This can be particularly advantageous when a quick response to changes in the process is required.

The strategy proposed in this paper was applied to a scenario with 3 parts and 9 machines. In future work, the applicability to scenarios with a larger number of parts will be investigated. It can be assumed that a higher complexity of the problem will not affect the efficiency of the method. Therefore, these studies will help to confirm that the proposed strategies can be used efficiently under conditions that require high production flexibility and allow a quick response to new orders. This will increase the efficiency of production and increase competitiveness, which is becoming increasingly important due to the need for scalability and the complexity of industrial production.

It must be noted that some simplifications have been made in this paper, e.g. the setup of the machines was neglected. When considering the flexibility of the proposed methodology for new orders, this should also be taken into account in order to obtain realistic results. As failures were not considered, it would also be interesting to investigate how failures on different machines could affect the production process and how the proposed strategy could be superior to standard procedures in this case. Another uncertainty that can be considered is the inclusion of the possibility of a defect in the raw material or semi-finished product, which can also cause a stoppage in production. This can be simulated by including these uncertainties in the simulation model, as the software itself allows for this and in this way comes even closer to reality. In this paper, a worker or a robot is not considered for the manipulation of the workpiece/part at the workplace, which should also be considered in future research, since the effective working time of the worker is less than the effective working time of the robot (physiological needs, fatigue, motivation, etc.).

Considering the simplifications mentioned above, which include neglecting transportation between machines, transportation from the SRM and transportation to the SFP, it should be noted that the regression polynomial can be prepared not only with simulated points but with real data from the production process. Moreover, the regression polynomial was used as a simple strategy to investigate the feasibility of the proposed approach. Some other machine learning models could be used to improve the efficiency of the proposed methodology.

Acknowledgements

This paper was financially supported by the University of Rijeka, Croatia, contract number "uniri-tehnic-18-223" and by the University of Rijeka, Croatia, contract number "uniri-mladi-tehnic-22-33". The authors would like to thank the company iTCR d.o.o. for its help.

5 REFERENCES

[1] Jeon, S. M. & Kim, G. (2016). A survey of simulation modeling techniques in production planning and control (PPC). *Production Planning & Control*, *27*(5), 360-377. https://doi.org/10.1080/09537287.2015.1128010

[2] Pešević, A. (2022). *Application of computer simulations in production planning and scheduling*. Master's thesis, University of Rijeka - Faculty of Engineering, Rijeka. https://urn.nsk.hr/urn:nbn:hr:190:522806

[3] Javernik, A., Ojstersek, R. & Buchmeister, B. (2023). The Impact of Collaborative Robot on Production Line Efficiency and Sustainability. *Tehnicki Glasnik*, *17*(2), 237-243. https://doi.org/10.31803/tg-20230425171431

[4] Kliment, M., Trebuna, P., Pekarcikova, M., Straka, M., Trojan, J. & Duda, R. (2020). Production efficiency evaluation and products quality improvement using simulation. *International Journal of Simulation Modeling*, *19*(3), 470-481. https://doi.org./10.2507/IJSIMM19-3-528

[5] Yang, S. L., Wang, J. Y., Xin, L. M. & Xu, Z. G. (2022). Verification of intelligent scheduling based on deep reinforcement learning for distributed workshops via discrete event simulation. *Advances in Production Engineering & Management*, *17*(4), 401-412. https://doi.org/10.14743/apem2022.4.444

[6] Stevanov, B., Sremcev, N., Lazarevic, M., Anderla, A., Sladojevic, S. & Vidicki, P. (2022). Optimization of the subassembly production process using simulation. *International Journal of Simulation Modeling*, *21*(4), 663-674. https://doi.org/10.2507/IJSIMM21-4-633

[7] Xu, W., Sun, H. Y., Awaga, A. L., Yan, Y. & Cui, Y. J. (2022). Optimization approaches for solving production scheduling problem: A brief overview and a case study for hybrid flow shop using genetic algorithms. *Advances in Production Engineering & Management*, *17*(1), 45-56. https://doi.org/10.14743/apem2022.1.420

[8] Neufeld, J. S., Schulz, S. & Buscher, U. (2023). A systematic review of multi-objective hybrid flow shop scheduling. *European Journal of Operational Research*, *309*(1), 1-23. https://doi.org/10.1016/j.ejor.2022.08.009

[9] Ruiz, R. & Vázquez-Rodríguez, J. A. (2010). The hybrid flow shop scheduling problem. *European Journal of Operational Research*, *205*(1), 1-18. https://doi.org/10.1016/j.ejor.2009.09.024

[10] Cai, J., Lei, D., Wang, J. & Wang. L. (2023). A novel shuffled frog-leaping algorithm with reinforcement learning for distributed assembly hybrid flow shop scheduling. *International Journal of Production Research*, *61*(4), 1233-1251. https://doi.org/10.1080/00207543.2022.2031331

[11] Kuo, Y., Yang, T., Cho, C. & Tseng, Y. C. (2008). Using simulation and multi-criteria methods to provide robust solutions to dispatching problems in a flow shop with multiple processors. *Mathematics and Computers in Simulation*, *7*(1), 40-56. https://doi.org/10.1016/j.matcom.2007.06.002

[12] Ribas, I., Leisten, R. & Framiñan, J. M. (2010). Review and classification of hybrid flow shop scheduling problems from a production system and a solutions procedure perspective. *Computers & Operations Research*, *37*(8), 1439-1454. https://doi.org/10.1016/j.cor.2009.11.001

[13] Antonova, A., Aksyonov, K. & Aksyonova, O. (2021). An Imitation and Heuristic Method for Scheduling with Subcontracted Resources. *Mathematics*, *9*(17), 2098. https://doi.org/10.3390/math9172098

[14] Janeš, G., Ištoković, D., Jurković, Z. & Perinić M. (2022). Application of Modified Steady-State Genetic Algorithm for Batch Sizing and Scheduling Problem with Limited Buffers. *Applied Sciences*, *12*(22), 11512. https://doi.org/10.3390/app122211512

[15] Borojević, S., Matić, D. & Dragić, M. (2022). An Integrated Intelligent CAD/CAPP Platform: Part II - Operation Sequencing Based on Genetic Algorithm. *Tehnicki Vjesnik*, *29*(5), 1686-1695. https://doi.org/10.17559/TV-20211012084632

[16] Rashid, M. F. F. A. & Mu'tasim, M. A. N. (2024). Cost-based hybrid flow shop scheduling with uniform machine optimization using an improved tiki-taka algorithm. *Journal of Industrial and Production Engineering*, *41*(2), 182-199. https://doi.org/10.1080/21681015.2023.2276108

[17] Ištoković, D., Perinić, M., Vlatković, M. & Brezočnik, M. (2020). Minimizing Total Production Cost in a Hybrid Flow Shop: A Simulation-Optimization Approach. *International Journal of Simulation Modeling*, *19*(4), 559-570. https://doi.org/10.2507/IJSIMM19-4-525

[18] Klanke, C. & Engell, S. (2022). Scheduling and batching with evolutionary algorithms in simulation–optimization of an industrial formulation plant. *Computers & Industrial Engineering*, *174*, 108760. https://doi.org/10.1016/j.cie.2022.108760

[19] Sajadi, S. M., Seyed Esfahani, M. M. & Sörensen, K. (2011). Production control in a failure-prone manufacturing network using discrete event simulation and automated response surface methodology. *The International Journal of Advanced Manufacturing Technology*, *53*, 35-46. https://doi.org/10.1007/s00170-010-2814-0

[20] Baril, C., Gascon, V. & Vadeboncoeur, D. (2019). Discrete-event simulation and design of experiments to study ambulatory patient waiting time in an emergency department. *Journal of the Operational Research Society*, *70*(12), 2019-2038. https://doi.org/10.1080/01605682.2018.1510805

[21] Li, G., Yang, S., Xu, Z., Wang, J., Ren, Z. & Li, G. (2020). Resource allocation methodology based on object-oriented discrete event simulation: A production logistics system case study. *CIRP Journal of Manufacturing Science and Technology*, *31*, 394-405. https://doi.org/10.1016/j.cirpj.2020.07.001

[22] Fu, J., Zhang, J., Ding, G., Qin, S. & Jiang, H. (2021). Determination of vehicle requirements of AGV system based on discrete event simulation and response surface methodology. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, *235*(9), 1425-1436. https://doi.org/10.1177/0954405421995602

[23] Rolf, B., Reggelin, T., Nahhas, A., Lang, S. & Müller, M. (2020). Assigning dispatching rules using a genetic algorithm to solve a hybrid flow shop scheduling problem. *Procedia Manufacturing*, *42*, 442-449. https://doi.org/10.1016/j.promfg.2020.02.051

[24] Malega, P., Gazda, V. & Rudy, V. (2021). Optimization of production system in plant simulation. *Simulation*, *98*(4), 295-306. https://doi.org/10.1177/00375497211038908

[25] Pekarcikova, M., Trebuna, P., Kliment, M. & Dic, M. (2021). Solution of Bottlenecks in the Logistics Flow by Applying the Kanban Module in the Tecnomatix Plant Simulation Software. *Sustainability*, *13*(14), 7989. https://doi.org/10.3390/su13147989

[26] Sargent, R. G. (2010). Verification and Validation of Simulation Models. *Proceedings of the Winter Simulation Conference*, 166-183. 05.12.2010-08.12.2010, Baltimore. https://doi.org/10.1109/WSC.2010.5679166

[27] Schlesinger, S. (1979). Terminology for model credibility. *Simulation*, *32*(3) 103-104. https://doi.org/10.1177/003754977903200304

[28] Dean, A., Voss, D. & Draguljić D. (2017). *Design and analysis of experiments*. New York, NY: Springer New York. https://doi.org/10.1007/978-3-319-52250-0

Authors' contacts:

David Ištoković, Assistant Professor, PhD
(Corresponding author)
University of Rijeka, Faculty of Engineering,
Vukovarska 58, 51000 Rijeka, Croatia
e-mail: distokovic@riteh.uniri.hr

Ivana Lučin, Postdoctoral researcher, PhD
University of Rijeka, Faculty of Engineering,
Vukovarska 58, 51000 Rijeka, Croatia
e-mail: ivana.lucin@riteh.uniri.hr

Ariana Pešević, mag. ing. mech.
University of Rijeka, Faculty of Engineering,
Vukovarska 58, 51000 Rijeka, Croatia
e-mail: ariana.0807@hotmail.com

Maja Vlatković, Postdoctoral researcher, PhD
University of Rijeka, Faculty of Engineering,
Vukovarska 58, 51000 Rijeka, Croatia
e-mail: maja.markovic@riteh.uniri.hr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240502092606

Original scientific paper
Received: 2024-05-02, Accepted: 2024-05-08

Injection Moulded Part Analysis by Alignment of Simulated and Referent Part Geometry

Damir Godec*, Filip Panđa, Mislav Tujmer

Abstract: Warpage of the injection moulded parts is one of the most common defects after demoulding, which is intensively present in thin-walled moulded parts or parts with non-uniform wall thickness. The level of the warpage can be reduced by optimizing the mould tempering system in the phase of mould design, by optimizing the injection moulding parameters and recently by optimizing the shape of the mould cavity elements geometry – so-called inverse contouring. Computer simulation of injection moulding process can show the level of moulded part warpage after ejection from the mould, but it will not allow detailed measuring of critical moulded part measures deviations. Detailed analysis and prediction of the critical moulded part dimensions can be performed by aligning of the deformed moulded part design obtained with computer simulation, with moulded part reference CAD geometry. This paper shows the main steps and an example of the geometry aligning process for a specific thermoplastic part.

Keywords: injection moulding; moulded part geometry alignment; numerical simulation

1 INTRODUCTION

The modern development of injection moulded polymer parts is directed towards sustainable production, where one of the key measures is the reducing of the moulded part wall thickness. The reason lies not only in reduction of thermoplastic material consumption, but also in significant reduction of the injection moulding cycle time. However, such thin-walled moulded parts are very challenging from the processing point of view because they require high processing parameters, and in addition, they are more prone to deformation after removal from the mould cavity during cooling, most often in the form of warpage. [1] The level of warpage can be reduced by optimizing the mould tempering system in the phase of its design [2], by optimizing the injection moulding parameters [3] and recently by optimizing the shape of the mould cavity elements geometry – so-called inverse contouring [4]. Computer simulation is a powerful tool for predicting behaviour of observed products/systems [5], and in case of injection moulding it is a tool for predicting the critical occurrences during the injection moulding cycle, as well as for prediction of the level of moulded part warpage after ejection from the mould [6].

On the other hand, injection moulding simulation will not allow detailed measuring of critical moulded part measures deviations. Detailed analysis and prediction of the critical moulded part dimensions can be performed by aligning the deformed moulded part design obtained with computer simulation, with moulded part reference CAD geometry [6]. This paper shows the main steps and an example of the process of aligning deformed moulded part geometry of specific thermoplastic part with reference geometry.

2 INJECTION MOULDED PART WARPAGE

Warpage of injection moulded parts is one of the most common visible defects caused by a few factors, the most frequent of which is non-uniform shrinkage of the moulded part during the cooling phase in the mould. If the shrinkage of the moulded part is uniform in all directions, the moulded part will not warp, and it will not change its shape. There will only be a uniform decrease of its dimensions, which can be compensated by simply increasing the dimensions of the referent moulded part CAD model for mould design [6].

Figure 1 Example of moulded part warpage [1]

There are several significant influencing factors on non-uniform moulded part shrinkage, like molecular and/or reinforcement orientation (in flow direction of polymer melt in the mould cavity) during mould filling phase, cooling process of polymer melt in the mould cavity, non-uniform mould temperature field (Fig. 2) and non-uniform moulded part wall thickness. [7]

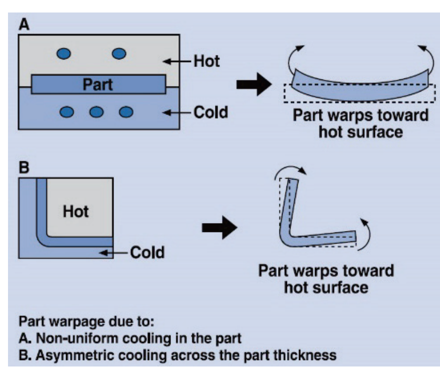


Figure 2 Moulded part warpage due to uneven mould temperature field [7]

Type of the processed polymer material, parameters of injection moulding process, geometry of the moulded part, mould design, mould tempering system as well as type, number, dimensions and position of the gates, are also significant influencing factors on moulded part shrinkage and warpage. Fig. 3 shows the influence of injection moulding parameters on moulded part shrinkage. [8]

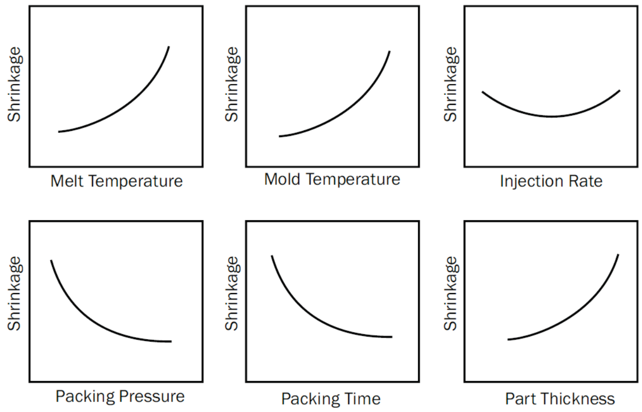


Figure 3 Influence of injection moulding parameters on moulded part shrinkage [8]

In the past, determination of shrinkage and warpage of polymer moulded part was one of the most complicated tasks in mould design process, but today, thanks to the detailed and precise numerical computer simulations, it is possible to estimate final shape and dimensions of the moulded part, even before production of the mould.

Software for simulation of injection moulding, mostly based on *Finite Element Method* (FEM) have been developed to simulate the processing of the polymer material from a melt phase at the start of injection phase to a solidified product at the time of ejection. The aim of these applications is to predict and understand the causes of shrinkage and warpage, which might occur during injection moulding.

Although numerical simulation of the injection moulding process can give an overall insight into moulded part warpage, when more precise measurements of dimensional deviation from referent part are necessary, raw simulation results are not suitable for estimating detailed deviations of critical dimensions and/or shape of the moulded part necessary for redesigning moulded part and mould cavity elements.

3 MINIMIZING INJECTION MOULDED PART WARPAGE

This paper shows optimisation of moulded part design for minimizing moulded part warpage after ejection from the mould cavity. Several numerical simulations of injection moulding process were performed, and warpage of injection moulded part depending on moulded part design was observed as a result. Software package *Autodesk Moldflow Insight* was used for that purpose.

Referent moulded part was a component of automotive connector – a handle for reducing the assembly force necessary for connector plug-in (Fig. 4). Polymer material for production of the handle is thermoplastic polybutylene terephthalate (PBT) reinforced with 20 % of glass fibres.

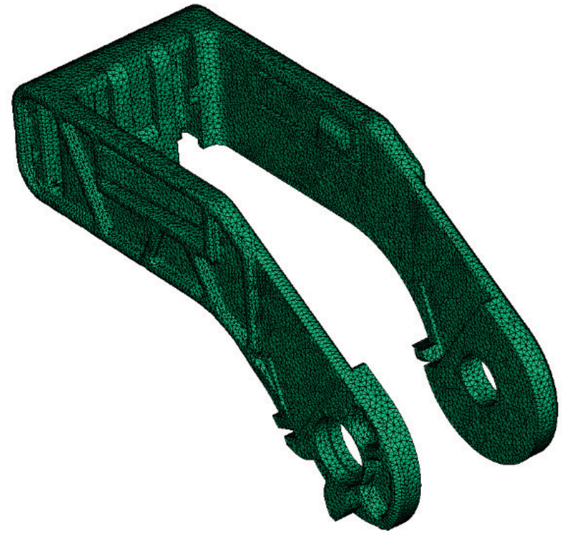


Figure 4 3D mesh model of a handle [6]

3.1 Determination of the Gate Position

The first step in optimization process of handle is determination of optimal gate position. Figure 5 shows optimal area (blue colour) determined by software analysis *Gate Location*.

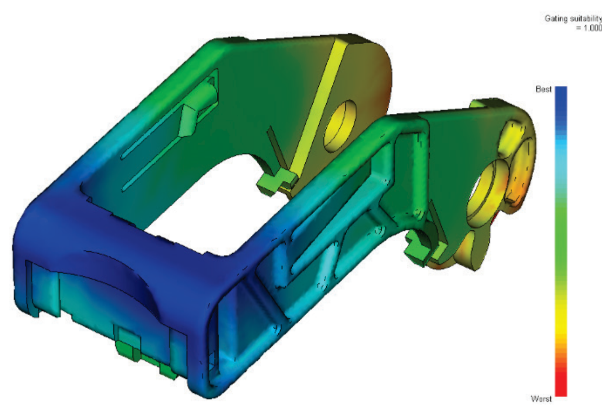


Figure 5 Results of *Gate Location* analysis [6]

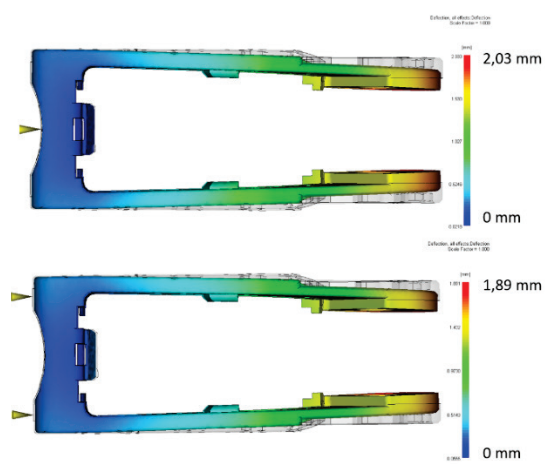


Figure 6 Warpage differences – single and two gates cavity [6]

Analysis of the influence of application more gates was performed within the next optimization step. Analysis sequence *Fill* + *Pack* + *Warp*, which encompasses analysis of mould cavity filling, phase of acting packing pressure and warpage analysis based on previous two analysis, was used for this purpose. Based on results shown on Fig. 6, application of two gates per cavity resulted with reduction of maximum warpage for 0,15 mm (for 7,3 %), and therefore a simulation model with two gates was used for further analyses.

3.2 Initial Injection Moulding Analysis

Within simulation model, moulded part is set into a steel mould with dimensions 75 × 135 × 100 mm. The mould has integrated 6 mm diameter cooling channels, and for all inlets of cooling media (water) the temperature is set to 85 °C. Initial analysis of injection moulding was performed with this model for detailed determination of warpage value of observed moulded part. Analysis sequence *Cool* + *Fill* + *Pack* + *Warp* was used because it includes influence of mould cooling system to the moulded part warpage. Figures 7 and 8 are showing average mould temperature fields in vertical and horizontal mould plane cross-sections.

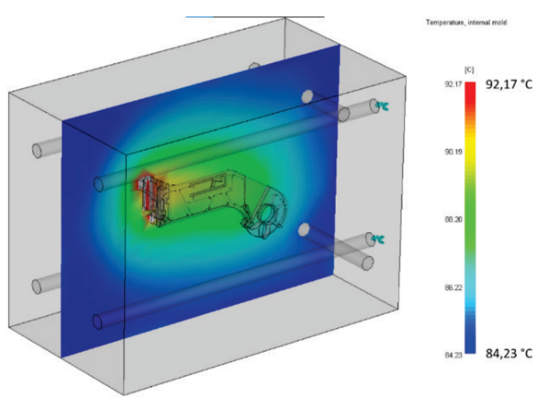


Figure 7 Cool analysis – vertical mould plane cross-section [6]

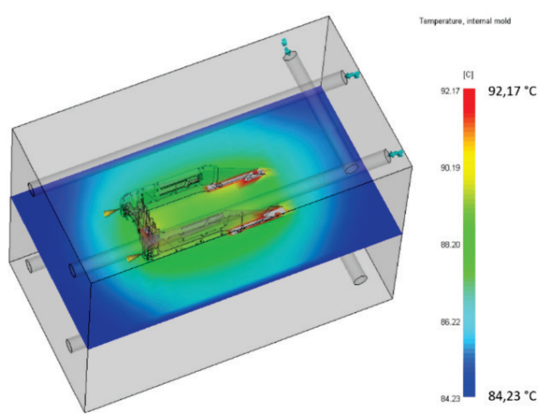


Figure 8 Cool analysis – horizontal mould plane cross-section [6]

Analyses of the filling phase and the phase of acting packing pressure (*Fill* + *Pack*) are performed after *Cool* analysis. Tab. 1 shows some of the results obtained with this analysis which provide important information for running injection moulding process as well as for evaluation of suitability of processing on specific injection moulding machine.

Table 1 Simulated injection moulding parameters [6]

	Value	Unit
Filling time	0,63	s
Max. injection pressure	27,6	MPa
Mould clamping force	18,93	kN
Moulded part cooling time	2,95	s

Last analysis in this sequence is warpage analysis (*Warp*). This analysis provides estimation of moulded part shrinkage and warpage. The results are shown as displacements of the nodes within finite element mesh, compared with initial position, for amounts in x, y and z directions (Cartesian coordinate system). This enables to observe results depending on each axis separately, or as a total displacement, where displacement of each mesh node is calculated according to Pythagorean Theorem.

Fig. 9 shows total displacements of nodes within mesh of the moulded part – top and isometric view [6].

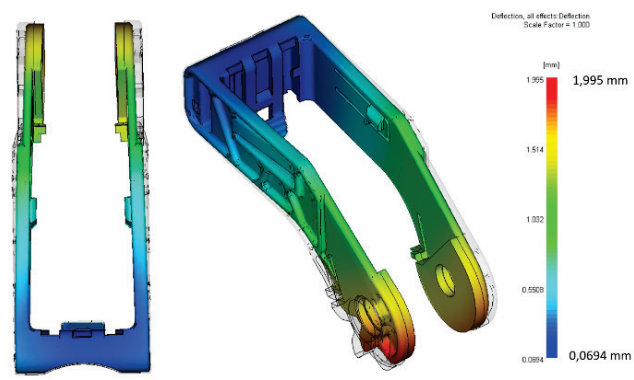


Figure 9 Warpage analysis [6]

Here, it is important to emphasise that presented displacements cannot be used for more precise measuring of deviations from referent geometries until warped moulded part is aligned with referent geometry.

3.3 Alignment of Warped Moulded Part with Referent Geometry

Software package Autodesk Moldflow enables exporting of deformed moulded part geometry in form of STL file. Resulting STL file is a mesh of connected triangles, which represents surface geometry of the moulded part. Deformed geometry in STL file consists of the same number of elements like initially generated *Dual Domain* mesh, which determine moulded part surface geometry. The location of nodes in this mesh is different compared to their initial locations due to shrinkage and warpage of the moulded part. For alignment of warped moulded part geometry with referent, *GOM Inspect* software was used. *GOM Inspect* is a software for non-contact measuring in product development phases, production and quality control. Data for the analysis are mainly provided with 3D scanning or computer tomography (CT scanner), but in this paper, exported geometry from software *Autodesk Moldflow* was used instead of data of scanned part.

GOM Inspect distinguishes two types of geometry – wanted (ideal) geometry in form of CAD body named as Nominal, and deformed mesh named as Actual geometry.

For analysis presented in this paper, in the first step automatic initial alignment (*Prealignment*) was performed, and in second step manual alignment according to generated geometrical element (*Alignment By Geometric Elements*) [6].

Fig. 10 shows initial alignment of moulded part deformed geometry with referent geometry. White colour represents referent CAD geometry, and green colour deformed geometry in STL format exported from *Autodesk Moldflow* analysis. [6]

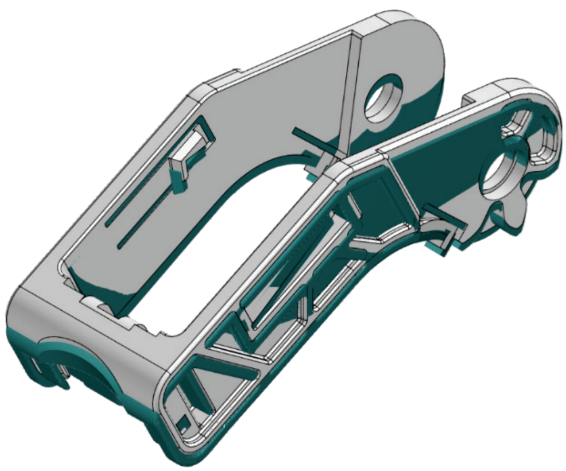


Figure 10 Initial alignment deformed and referent geometry (*Prealignment*) [6]

Disadvantages of such alignment are the most visible at rear surface of moulded part. Although the software with automatic search function found nodes that match each other, geometries generally are not aligned well, therefore such aligned geometry cannot be considered for further analysis. Therefore, it is necessary to perform alignment by manually generated geometrical elements, such are nodes, planes, lines, cylinders, etc.

In all analyses in this paper, referent and deformed geometries are aligned by three planes – two symmetric planes generated between side walls, and one on the rear side of the moulded part. One of the alignments is shown on Fig. 11 [6].

The planes should be generated separately on referent and deformed geometry. Planes on referent geometry have no attached measuring principle, since it is not necessary for it, while at deformed geometry surface have assigned *Fitting Element* measuring principle with which selected nodes on the plane are projected at nominal geometry for assuring the most accurate measuring results.

Additional problem while generating the planes on deformed geometry is in difficulties in recognizing of deformed planes because relatively small number of nodes which are present in STL mesh, despite the element size of approximately 0,4 mm. Measuring software have code written for the meshes with large number of polygons generated from point cloud obtained with scanning of real objects, therefore in this research the mesh is refined using the *Refine Mesh* function within *GOM Inspect* software. Two refining iterations were used, which resulted in increasing the number of polygons of a specific surface approximately 6 times. Thus, smaller triangles were generated within existing ones, therefore the software has more nodes for generating planes on deformed geometry [6].

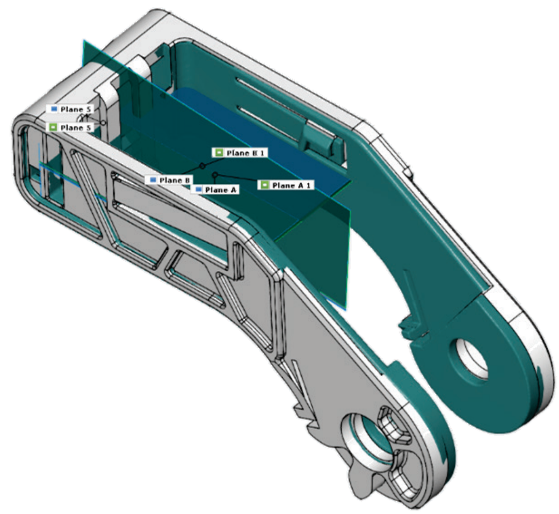


Figure 11 Alignment of deformed and referent geometries by defined geometric elements (*Alignment By Geometric Elements*) – isometric view [6]

From Fig. 11 it is obvious that, after alignment by geometric elements, both geometries are much better aligned at rear side of the moulded part, which is more dimensionally stable, while side walls are extremely deformed inwards [6].

Fig. 12 shows alignment of warped and referent geometry by geometric elements from top view.

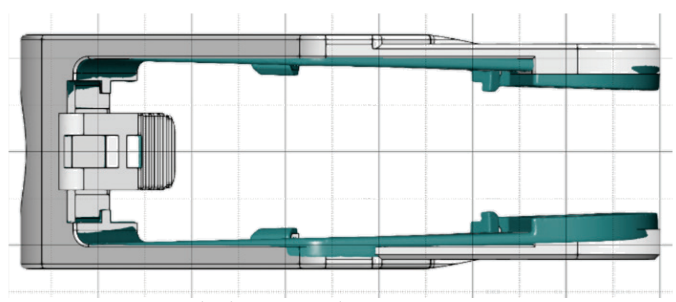


Figure 12 Alignment of deformed and referent geometry by geometric elements (Alignment By Geometric Elements) – top view [6]

3.4 Results of Alignment – Surface Comparison

When all parts are correctly aligned, using the *Surface Comparison on CAD* function it is possible visually present how much deformed (warped) geometry deviates from the referent. Software *GOM Inspect* calculates deviation as an orthogonal distance each node in the mesh from related position in referent CAD geometry and results are shown in different colours at copy of referent CAD geometry.

Figure 13 shows the deviations of nodes on deformed geometry from referent. Green colour at image marks regions where characteristics of deformed geometry matches the best with characteristics of referent CAD geometry. Red and blue colours present zones of maximal deviation of deformed geometry.

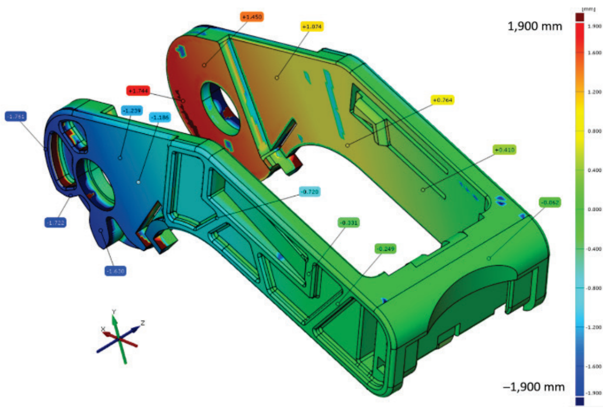


Figure 13 Deviations of nodes at deformed geometry compared with referent geometry (Surface Comparison on CAD - *GOM Inspect*) [6]

From the results it is obvious that maximal deviation of the moulded part is approximately 1,8 mm on side moulded part walls. From Figs. 12 and 13 it can be concluded that side part walls are almost linearly warped inwards, therefore it is possible to calculate the angle of inclination α with determination of length of thinner side part approximately 57,5 mm.

$$\alpha = \sin^{-1}\left(\frac{1{,}8}{57{,}5}\right) = 1{,}79°. \quad (1)$$

3.5 Referent Moulded Part Inverse Contouring

For reduction of moulded part warpage there are few strategies. The simplest approach is optimisation of injection moulding parameters for minimizing moulded part warpage, but this approach often will not result with satisfying warpage reduction or results with to long injection moulding cycles. The second strategy is oriented to optimal mould tempering, however, design mould characteristics often will not allow application of optimal shapes and configurations of mould tempering channels. Within this strategy, application of additive manufacturing allow production of co-called conformal cooling system, but on the other hand, this approach results with increased mould production costs and it is reasonable only in case of large batch or mass production.

Therefore, in this case the third strategy is imposed – correction of the moulded part (and appropriate mould cavity elements) geometry, using which warped geometry after injection moulding will converge to the referent moulded part geometry – inverse contouring. The main goal of inverse contouring is to minimize moulded part warpage by deliberately extracting walls in opposite direction of the warpage, which will result with aimed dimensions and shape of final moulded part. In that case, mould cavity elements are designed in opposite direction of moulded part warpage tendency. This strategy enables easier achieving moulded part dimensional specifications, avoiding mould redesign, expensive corrections and fastest moulded part validation, which all in general result with shorter time from idea to final product ready for the market.

In case of observed handle, the most important task is to achieve flatness of sidewalls. Without inverse contouring, those walls will be notably warped inwards. As the previous analysis shows that side walls will be deformed inwards under the angle of approximately 1,79°, in moulded part process, side walls are extracted outwards under the angle of 1,8° (Fig. 14).

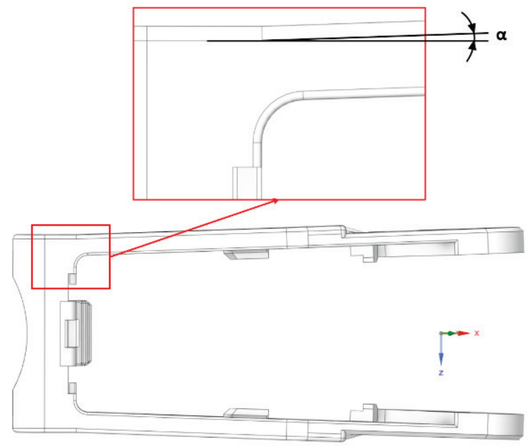


Figure 14 Moulded part geometry after inverse contouring [6]

The process of computer analysis of injection moulding is repeated with such, purposely deformed moulded part geometry. Warpage results obtained in software *Autodesk Moldflow* are again imported in GOM Inspect software in form of STL file, and process of geometry alignment is repeated. Fig. 15 shows results of alignment of moulded part inverse contoured geometry and referent part geometry [6].

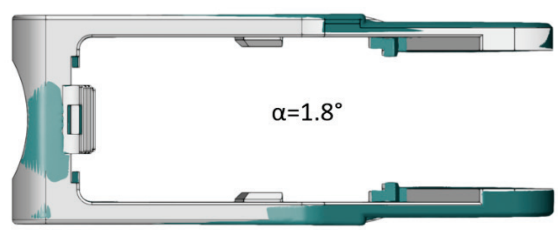


Figure 15 Alignment of warped inverse contoured moulded part geometry with referent geometry [6]

After alignment with function *Surface Comparison on CAD*, *GOM Inspect* software calculates deviations of inversed contour geometry from referent geometry. Results of comparison of deviations of deformed geometries without inverse contouring and with inverse angle 1,8° are shown in Fig. 16.

By application of extraction angle on side walls of handle, maximal deviation of side walls is reduced from approximately 1,8 – 1,9 mm to approximately 0,2 – 0,3 mm, (warpage reduction for more than 80 %).

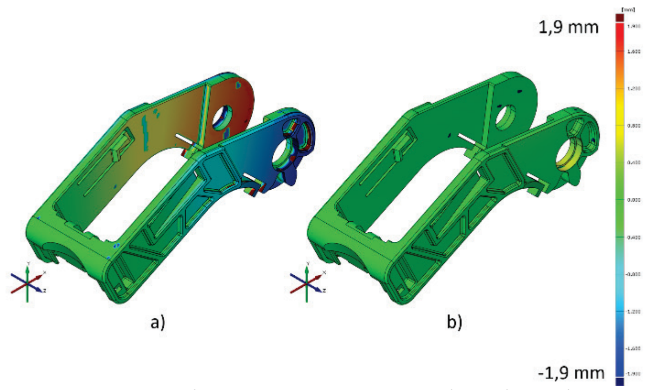


Figure 16 Deviation of warped moulded part geometry from referent: a) without inverse contouring, b) with inverse contouring angle α =1,8° [6]

4 CONCLUSIONS

This paper presents process of injection moulded part geometry optimisation with alignment of deformed geometry obtained with computer simulation of injection moulding process and referent moulded part geometry. Inverse contouring is also presented as one of the tools for optimisation of deformed geometry for achieving minimal warpage values. This moulded part geometry optimisation principle can be useful for the companies whose main business is polymer moulded part development and processing with injection moulding.

Presented example shows decrease of warpage at important moulded part zones by more than 80 % with relatively simple changes in moulded part and mould design. Application of this optimisation principle can result with production of more dimensionally accurate moulded parts and easier validation of production during phase of testing and quality control.

In the mold and die making stage, numerical simulation can be used as a tool for accurate estimation of the zone of moulded part warpage and real deviations of the geometry from referent moulded part geometry can be obtained by analysing warped moulded part within specialised software for metrology. The complete presented process can be performed even before the mould and moulded part are produced.

5 REFERENCES

[1] Jyo, C. (2023, August 7). The Causes and Solutions of Warping in Injection Molding. https://prototool.com/warping-in-injeciton-molding/

[2] Mohd Hanid, M. H., Abd Rahim, S. Z., Gondro, J., Sharif, S., Al Bakri Abdullah, M. M., Zain, A. M., El-hadj Abdellah, A., Mat Saad, M. N., Wysłocki, J. J. & Nabiałek, M. (2021). Warpage Optimisation on the Moulded Part with Straight Drilled and Conformal Cooling Channels Using Response Surface Methodology (RSM), Glowworm Swarm Optimisation (GSO) and Genetic Algorithm (GA) Optimisation Approaches. *Materials, 14*(6), 1326-1326. https://doi.org/10.3390/ma14061326

[3] Hiyane-Nashiro, G., Hernández-Hernández, M., Rojas-García, J., Rodriguez-Resendiz, J. & Álvarez-Alvarado, J. M. (2022). Optimization of the Reduction of Shrinkage and Warpage for Plastic Parts in the Injection Molding Process by Extended Adaptive Weighted Summation Method. *Polymers, 14*(23), 5133. https://doi.org/10.3390/polym14235133

[4] Hopmann, C., Gerads, J. & T. Hohlweck. (2021). Investigation of an inverse thermal injection mould design methodology in dependence of the part geometry. *International Journal of Material Forming, 14*(2), 309-321. https://doi.org/10.1007/s12289-020-01604-6

[5] Daneshjo, N., Mares, A., Malega, P. & Francova, Z. (2022). CAD Model of Rear-View Mirror and Simulation of Its Aerodynamics and Noise. *International Journal of Simulation Modelling, 21*(2), 226-236. https://doi.org/10.2507/ijsimm21-2-598

[6] Pandža, F. (2023). Optimisation of moulding design for minimisation of warpage. *M.Sc. Thesis*, Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb.

[7] Design Guide Performance and Value with Engineering Plastics. (n.d.). Retrieved March 19, 2024, from http://www.hankyan.com/design_guide_with_engineering_plastics_by_0DSM.pdf

[8] Fisher, J. M. (2013). *Handbook of Molded Part Shrinkage and Warpage*. 2nd Edition. William Andrew Publishing, ISBN 978-1-4557-2597-7.

Authors' contacts:

Damir Godec, Full Professor PhD
(Corresponding author)
University of Zagreb,
Faculty of Mechanical Engineering and Naval Architecture,
Ivana Lučića 5, 10000 Zagreb, Croatia
damir.godec@fsb.unizg.hr

Filip Panđa, mag. ing. mech.
YAZAKI-EUROPE Ltd.
Slavonska avenija 26/6, 10000 Zagreb, Croatia
filip.pandja@yazaki-europe.com, filippandja@gmail.com

Mislav Tujmer, mag. ing. mech.
University of Zagreb,
Faculty of Mechanical Engineering and Naval Architecture,
Ivana Lučića 5, 10000 Zagreb, Croatia
mislav.tujmer@fsb.unizg.hr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240503194515

Original scientific paper
Received: 2024-05-03, Accepted: 2024-05-08

Analysis of the Geometric Accuracy of Wax Models Produced Using PolyJet Molds

Mariusz Cygnar*, Tomasz Dziubek, Tomasz Kądziołka, Grzegorz Budzik, Damian Żelechowski, Mateusz Majewski, Paweł Turek

Abstract: The article presents an analysis of the geometric accuracy of wax models produced using PolyJet molds. Photopolymer resin was used to make the molds, and the surface of the model was finished in a gloss mode in the 3D printing process. This allows you to obtain a smooth surface of the molded part model. The geometric accuracy of the models was determined by measurements with the GOM ATOS scanner and analysis of the measurement results in relation to the base model made in the 3D-CAD program. Accuracy analysis was the basis for determining the wear of the mold wear when making subsequent models.

Keywords: 3D printing; geometric accuracy; PolyJet mold; rapid tooling

1 INTRODUCTION

One of the basic concepts used in the area of methods for supporting and rapid production preparation is Rapid Prototyping - RP [1]. This concept refers to the rapid production of models in unit or low-volume production, based on computer-aided capabilities and the use of three-dimensional computer models made in the CAD environment and 3D printing [2, 3]. The benefits of using this type of solutions are based on the possibility of physically presenting the product or visualizing the product to potential buyers. The offered opportunities allow for a significant reduction in investment risk, which in turn is important for the development strategies of companies in the production area [4, 5]. The development of this type of solutions is related to mastering the possibilities offered by additive technologies, leading to a situation in which they can be successfully used to produce technological tools in the Rapid Tooling - RT [6-8] process and ready-made Rapid Manufacturing - RM products [9].

Taking into account the process of introducing new products to the market, it is worth noting that rapid prototyping techniques can be useful at each stage of the process [10]. Direct production of prototypes presenting the visual, geometric or ergonomic properties of products using additive methods offers many benefits [11]. A problem may arise when it is necessary to create a functional model or technological prototype. Such a model should be manufactured using the final selected manufacturing technique and the final selected material, so that the properties of such a model are as close as possible to the final product. Using techniques used in mass production to produce models on a unit scale is not very profitable due to the high cost of preparing production tools [12, 13]. The aviation industry uses proven technologies, especially in the production of aircraft engine parts. In the case of blades of hot parts of aircraft engines, one of the basic manufacturing methods is precision casting from heat-resistant alloys. For this purpose, wax models made in mass production by injection into metal alloy molds are used. In order to test new design solutions for turbine blades, the production of metal injection molds for the production of wax models is very expensive. For this reason, the use of molds based on the Rapid Tooling process from polymer materials using the PolyJet method [14, 15] is an application for the quick and economically effective production of prototype models. Molds of this type enable the production of wax models in numbers ranging from several dozen to even several hundred pieces. However, it is necessary to analyze the wear of the mold cavity, which, due to thermal and mechanical influence, wears out earlier than molds made of metal alloys. One of the methods for assessing mold cavity wear is measuring the geometric accuracy of wax models using optical scanning, which is presented in this article.

2 PREPARATION OF RESEARCH MODELS

The preparation of research models consisted of several stages. It started with the preparation of a 3D-CAD model of the blade, the wax models of which were made of wax. Then, a mold was designed for low-pressure wax injection, which includes a channel for supplying liquid wax, a cylindrical container with a piston for low-pressure injection, and venting channels. A 3D-CAD model of the mold was also created, which was then transformed into an STL model intended for production using additive technology. The MJT – Material Jetting process and PolyJet technology were chosen to produce the molds.

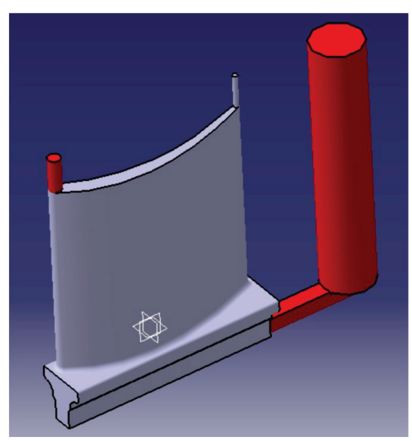


Figure 1 3D-CAD model of blade

In the process of producing the mold model, the Objet EDEN 260 device was used, based on the PolyJet technology solution. The starting material in the printing process was a liquid photopolymer called RGD720 (Stratasys, USA), which guarantees high stability of dimensions and shapes of the created models. The single layer thickness of the printing process was 16 microns and the Glossy finishing mode was used for finishing. With such selected parameters, it was possible to obtain a very smooth surface of the model immediately after printing, which allowed for easier demoulding (removing the cast models from the mold) in the later stages of the experiment. A view of the manufactured mold and the wax model of the blade is shown in Fig. 2.

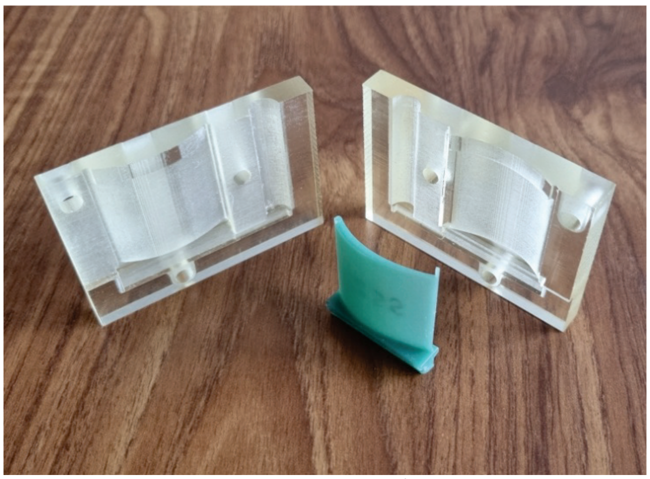


Figure 2 Mold and wax cast of the blade

Figure 3 View of the wax heating furnace

The next step was to cast the blade model from wax using a mold printed in PolyJet technology. One cycle of the casting process took approximately 180 minutes. It consisted of several stages: heating and removing the mold from the oven, pouring wax into the mold, cooling it and demoulding it. Fig. 3 shows a photo of a wax heating furnace.

The mold was filled with KC 6052D casting wax, which was heated in an oven to 100 °C and then poured into the mold to obtain a wax model of the blade. Fig. 4 shows a graph of wax temperature changes during the casting process.

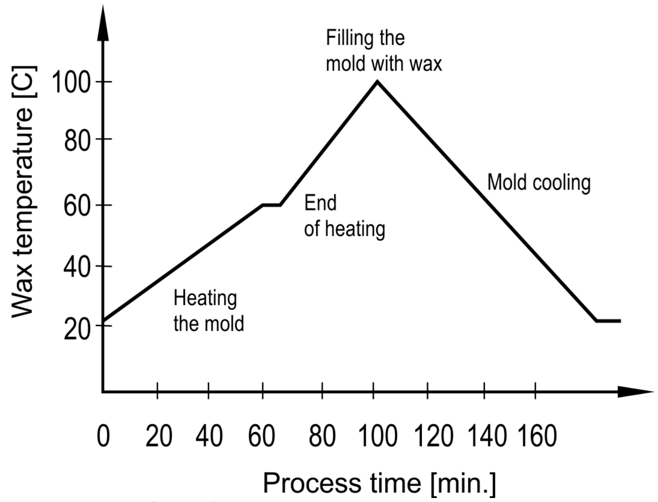


Figure 4 Graph of temperature changes during the casting process

Before the casting process, the two parts of the mold were connected using three screws with nuts, which ensured the stability of the connection of the mold elements during the pouring process. The view of the mold with a filled cavity is shown in Fig. 5.

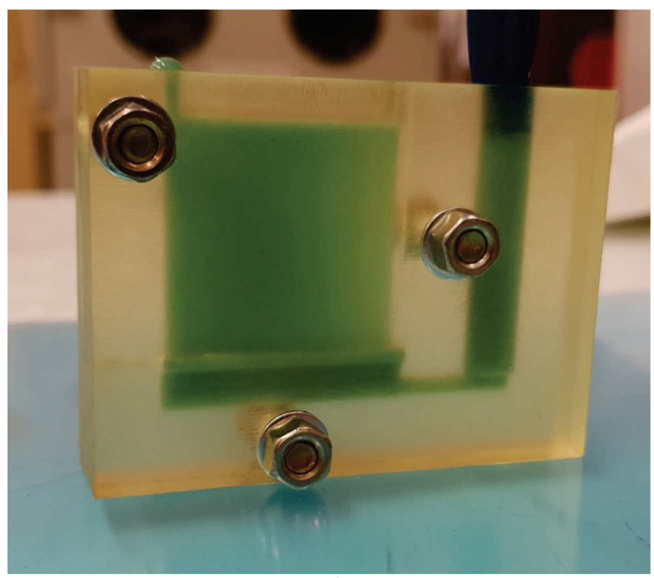


Figure 5 Mold filled with wax

After the mold had cooled down, its parts were separated and the cast wax model was removed. The mold cavity and adjacent surfaces were cleaned before the next casting process. In this way, 100 research models were made to assess the wear of the mold cavity as subsequent pieces of casting were made.

3 RESULTS OF MEASUREMENTS OF THE GEOMETRY OF WAX MODELS OBTAINED BY LOW PRESSURE MOLDING

The shape change was examined using a coordinate measurement technique based on the structured light scanning process. During the research, no measurements of the mold cavity were made because it would have been necessary to scan two parts and then assemble them in software. Such a process would introduce additional errors in the interpretation of the results. A map of deviations from the nominal dimension was made by superimposing the scanned model on the reference model. Using the indicated method, the geometry of the moldings obtained in the low-pressure forming process. Measured was made using the Atos II Triple Scan scanner with structured light illumination (Fig. 6). Measurements of the blades were made from two sides (inner and outer), but due to the volume of the article, the measurement results of only one side are presented.

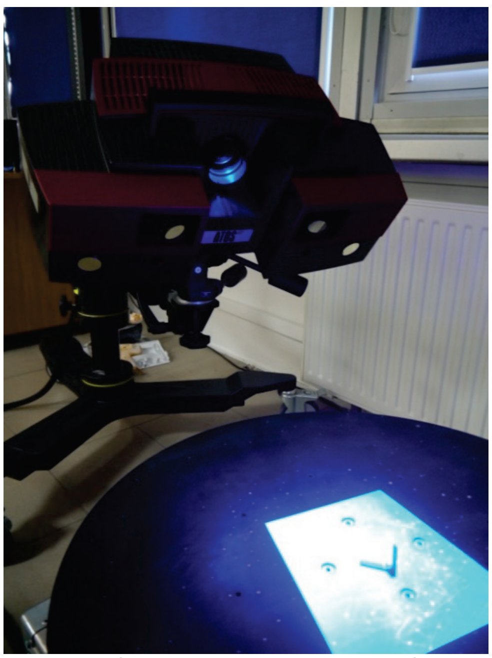


Figure 6 Scanning process using Atos II Triple Scan

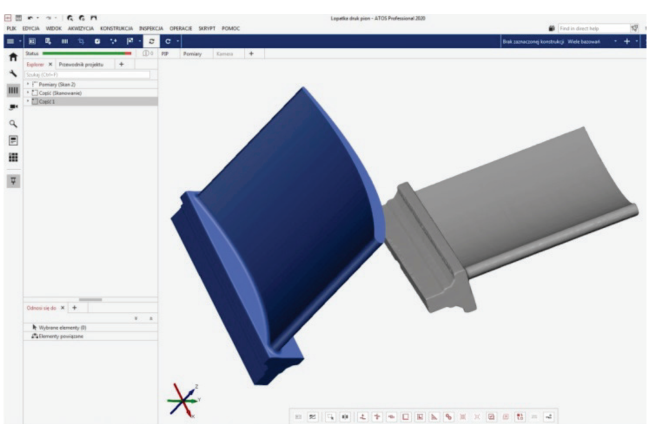


Figure 7 Overlaying models of the measured object on the nominal STL model in Atos Professional

Then, using the Atos Professional software, the mesh of the model obtained as a result of scanning was superimposed on the mesh of the nominal model and an analysis of the deviations of the dimensions of the actual model from the nominal dimension was carried out (Fig. 7).

The following drawings show views of deviation maps for subsequent models obtained in the low-pressure forming process using a mold obtained in the incremental process (Figs. 8 - 17). Measurements were performed on every tenth model for the research cycle of producing 100 wax casts.

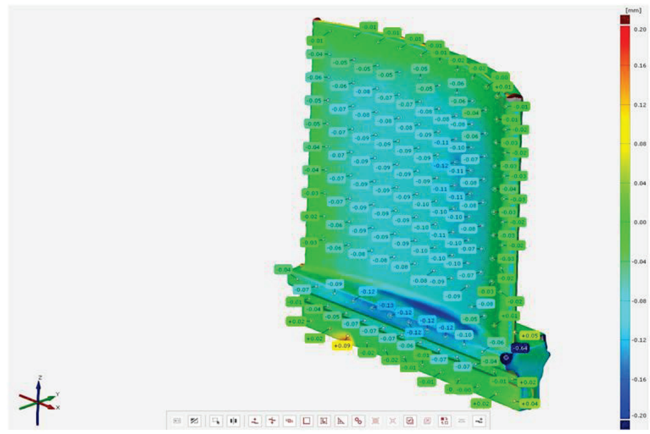


Figure 8 Analysis of the measurement results of the wax model of blade no. 10

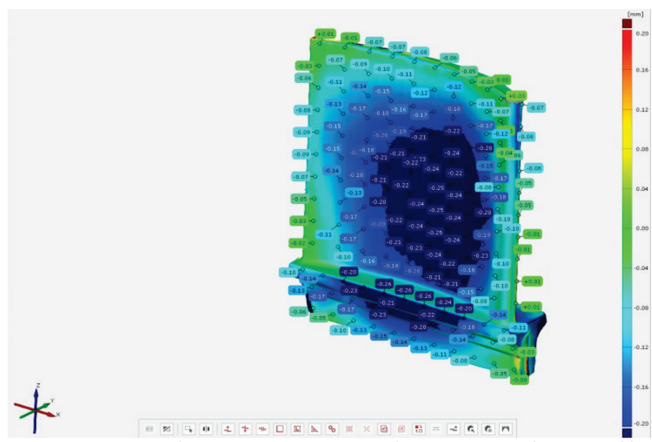


Figure 9 Analysis of the measurement results of the wax model of blade no. 20

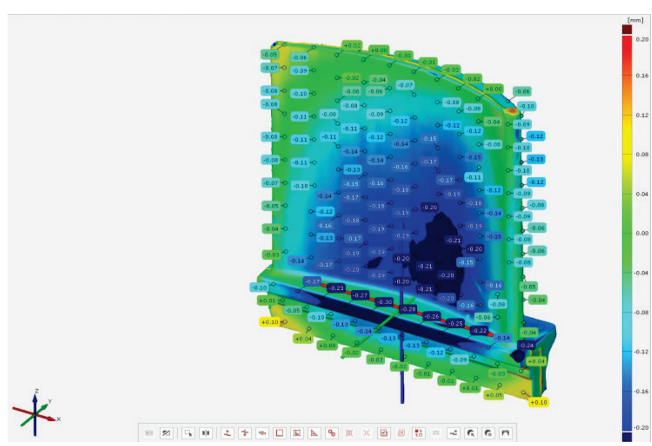


Figure 10 Analysis of the measurement results of the wax model of blade no. 30

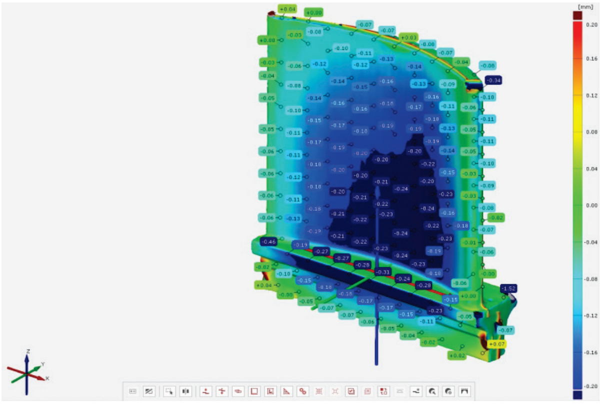


Figure 11 Analysis of the measurement results of the wax model of blade no. 40

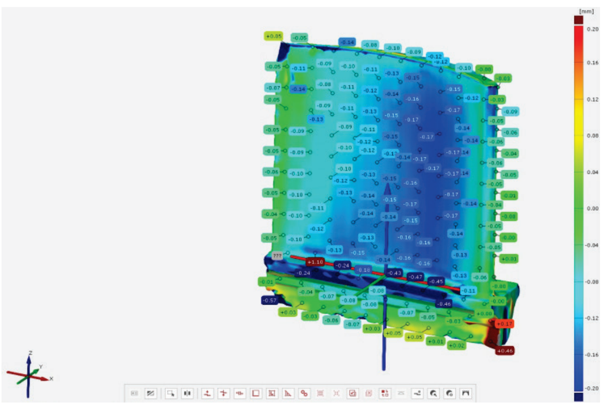


Figure 12 Analysis of the measurement results of the wax model of blade no. 50

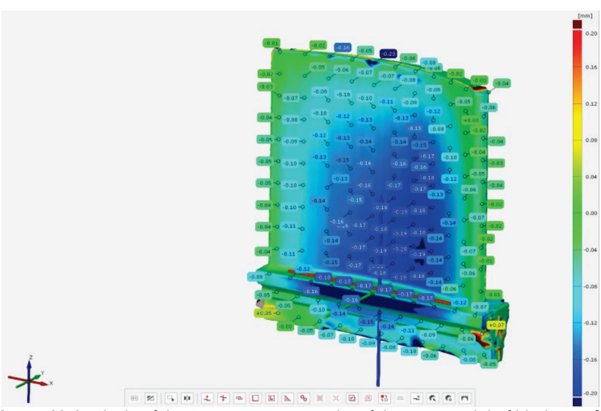


Figure 13 Analysis of the measurement results of the wax model of blade no. 60

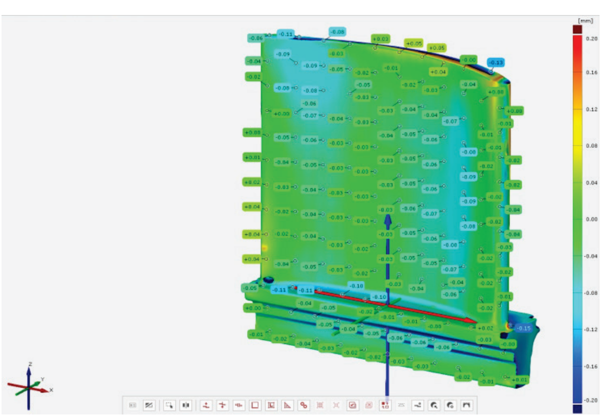


Figure 14 Analysis of the measurement results of the wax model of blade no. 70

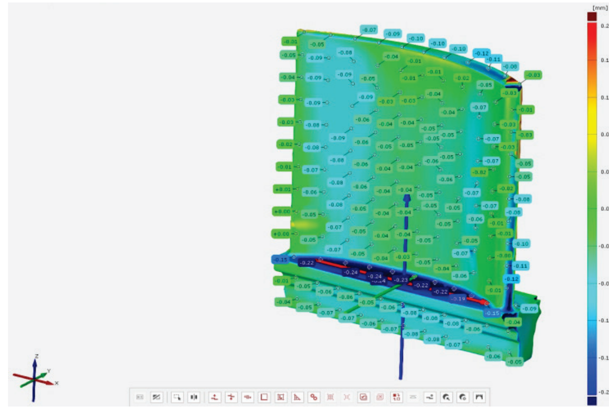


Figure 15 Analysis of the measurement results of the wax model of blade no. 80

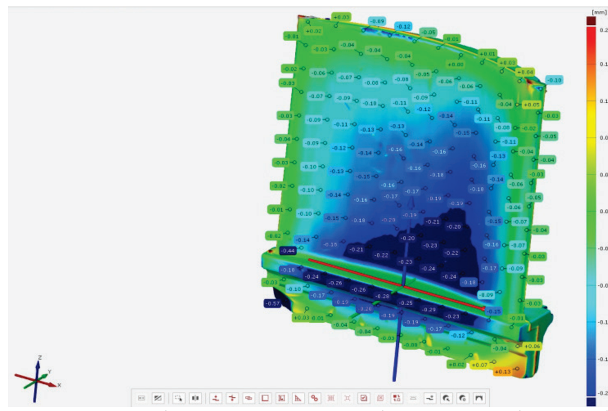


Figure 16 Analysis of the measurement results of the wax model of blade no. 90

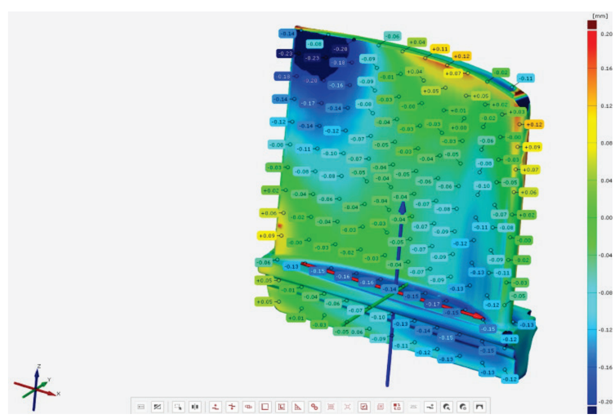


Figure 17 Analysis of the measurement results of the wax model of blade no. 100

4 CONCLUSIONS

During the low-pressure molding process using a mold obtained in the additive process, greater ease in the mold filling process was noticed, despite the use of wax filled with particles of other phases (composite wax). The general analysis of the obtained models allowed us to notice again a significant predominance of negative deviations from the nominal dimension. The shrinkage of the model is probably due to material shrinkage.

In the case of the first ten cast models, analysis of the deviation map allowed for the conclusion of a high regularity of deviations from the nominal dimension: the deviations were one-sided and evenly distributed on both sides of the blades. Deviations on the outer surface were up to -0.1 mm. Such regularity of model inaccuracies can be treated as an advantage, because in such a case, rescaling the mold by a given value would allow for easy compensation of shrinkage and obtaining precise models.

From model 60, positive deviations from the nominal dimension were noticed near the side parts of the model (corners) on both sides of the model. Positive deviations in the indicated places appeared on subsequent models, so they probably result from mold damage. From the 70th model implemented, a negative effect of heating the silicone mold before the molding process was noticed, affecting its deformation. From that moment on, elements stiffening the form were used. Therefore, the dimensional accuracy of the models improved, which can be seen by analyzing the measurement results, mainly for the internal part of samples no. 70 and 80.

In the case of models 70 and 80, a slight improvement in dimensional accuracy was also noticed on their external surfaces, resulting from stiffening the structure with additional elements. In the case of these samples, a negative deviation was noticed on the outer part of the blade in the place where the blade inclination angle changed from negative to positive. In this area, the shrinkage is the greatest, therefore, as the number of molding cycles increases, it is in this area of the surface that the greatest inaccuracy should appear. This shows that the shape of the blade has a significant impact on the direction of deformation resulting from shrinkage. The inaccuracy of the obtained models, even after 100 processes, did not generally exceed –0.2 mm on the curve surfaces (external and internal), which proves that the obtained models are quite accurate compared to other methods enabling obtaining models for precision casting processes. Additionally, the low value of deviations may indicate that the surface of the socket model showed quite good resistance to abrasive wear caused by flowing wax. Damage to the mold was related to small tears during demoulding in mold elements with a narrow cross-section - around the side lines of the blades.

5 REFERENCES

[1] Afonso, D., Pires, L. & Alves de Sousa, R. J. (2017). Direct rapid tooling for polymer processing using sheet metal tools. *Procedia Manufacturing, 13*, 102-108. https://doi.org/10.1016/j.promfg.2017.09.016

[2] Aminabadi, S. S., Tabatabai, P., Steiner, A., Gruber, D. P., Freisenbichler, W., Habersohn, C. & Berger-Weber, G. (2022). Industry 4.0 In-Line AI Quality Control of Plastic Injection Molded Parts. *Polymers, 14*(17), 3551. https://doi.org/10.3390/polym14173551

[3] Bagalkot, A., Pons, D., Clucas, D. & Symons, D. (2019). Methodology for setting the injection moulding process parameters for polymer rapid tooling inserts. *Rapid Prototyping Journal, 25*(9), 1493-1505. https://doi.org/10.1108/RPJ-10-2017-0217

[4] Budzik, G., Woźniak, J., Paszkiewicz, A., Przeszłowski, Ł. & Dziubek T., Dębski M. (2021). Methodology for the quality control process of additive manufacturing products made of polymer materials. *Materials, 14*, 2202. https://doi.org/10.3390/ma14092202

[5] Budzik, G. & Żelechowski, D. (2019). Possibilities of using photopolymer injection molds in small series production in Industry 4.0 structure. *Polimery, 64*(6), 405-410. https://doi.org/10.14314/polimery.2019.6.2

[6] Buswell, R. A., Leal de Silva, W. R., Jones, S. Z. & Dirrenberger, J. (2018). 3D printing using concrete extrusion: a roadmap for research. *Cem. Concr. Res., 112*, 37-49. https://doi.org/10.1016/J.CEMCONRES.2018.05.006

[7] Fuwen, H., Mikolajczyk, T., Pimenov, D. Y. & Gupta, M. K. (2021). Extrusion-based 3D printing of ceramic pastes: Mathematical modelling and in situ shaping retention approach. *Materials, 14*, 1137. https://doi.org/10.3390/MA14051137

[8] Gibson, I., Rosen, D., Stucker, B. & Khorasani, M. (2021). Materials for additive manufacturing. In *Additive Manufacturing Technologies*. Springer: Cham, Switzerland, 379-428. https://doi.org/10.1007/978-3-030-56127-7_14

[9] Ian, C., Olaf, D., Joseph, K., Noah, M. & Terry, W. (2021). Wohlers Report 2019: Additive Manufacturing and 3D Printing State of the Industry, van Rensburg, J., Ed. Wohlers Associates, Inc.: Fort Collins, CO, USA.

[10] ISO/ASTM52921-13. (2019). Standard Terminology for Additive Manufacturing-Coordinate Systemsand Test Methodologies.

[11] ISO/ASTM52915-20. (2020). Specification for Additive Manufacturing File Format (AMF) Version1.2.

[12] ISO/ASTM52904-19. (2020). Additive Manufacturing– Process Characteristic sand Performance: Practice for Metal Powder Bed Fusion Process to Meet Critical Applications.

[13] Wu, H., Li, D. & Guo, N. (2009). Fabrication of integral ceramic mold for investment casting of hollow turbine blade based on stereolithography. *Rapid Prototyping Journal, 15*(4), 232-237. https://doi.org/10.1108/13552540910979749

[14] Turek, P., Budzik, G., Sęp, J., Oleksy, M., Józwik, J., Przeszłowski, Ł., Paszkiewicz, A., Kochmański, Ł. & Żelechowski, D. (2020). An Analysis of the Casting Polymer Mold Wear Manufactured Using PolyJet Method Based on the Measurement of the Surface Topography. *Polymers, 12*, 3029. https://doi.org/10.3390/polym12123029

[15] Zink, B., Szabo, F., Hatos, I., Suplicz, A., Kovacs, N. K., Hargitai, H., Tabi, T. & Kovacs J. G. (2017). Enhanced Injection Molding Simulation of Advanced Injection Molds. *Polymers, 9*(3), 77. https://doi.org/10.3390/polym9020077

Authors' contacts:

Mariusz Cygnar, PhD, DSc. Eng. Associate Prof.
(Corresponding author)
Academy of Applied Sciences in Nowy Sącz,
ul. Staszica 1, 33–300 Nowy Sącz, Poland
mcygnar@ans-ns.edu.pl

Tomasz Dziubek, PhD, DSc. Eng. Associate Prof.
Rzeszów University of Technology,
Al. Powstańców Warszawy 12, 35-959 Rzeszów, Poland
tdziubek@prz.edu.pl

Tomasz Kądziołka, PhD, DSc. Eng.
Academy of Applied Sciences in Nowy Sącz,
ul. Staszica 1, 33–300 Nowy Sącz, Poland
tkadzialka@ans-ns.edu.pl

Grzegorz Budzik, Prof. PhD, DSc, Eng.
Rzeszów University of Technology,
Al. Powstańców Warszawy 12, 35-959 Rzeszów, Poland
gbudzik@prz.edu.pl
Academy of Applied Sciences in Nowy Sącz,
ul. Staszica 1, 33–300 Nowy Sącz, Poland
mcygnar@ans-ns.edu.pl

Damian Żelechowski, PhD, DSc. Eng.
PROSOLUTIONS Majewscy Sp. J.,
Trasa Lubelska 22, 05-462 Majdan, Poland
damian.zelechowski@prosolutions.pl

Mateusz Majewski, Eng.
PROSOLUTIONS Majewscy Sp. J.,
Trasa Lubelska 22, 05-462 Majdan, Poland
mateusz.majewski@prosolutions.pl

Paweł Turek, PhD, DSc. Eng. Assistant Prof.
Rzeszów University of Technology
Al. Powstańców Warszawy 12, 35-959 Rzeszów, Poland
pturek@prz.edu.pl

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240513233046

Original scientific paper
Received: 2024-05-13, Accepted: 2024-05-15

A Maturity Model to Determine the Degree of Utilization of Machine Learning in Production Planning and Control Processes

Jakob Hartl, Jürgen Bock*

Abstract: The presented work introduces a maturity model for evaluating Machine Learning implementations, with a primary focus on Production Planning and Control processes, as well as broader organizational and technical aspects in companies. This model emerges as a response to the research gap identified in the analysis of 14 existing maturity models, which served as foundational bases for the development of this novel approach. By examining success factors and obstacles at different maturity levels, categorized according to defined dimensions and overarching design fields, this model can serve as a catalyst for bridging the research gap between models demanded in practice and the scholary exploration of topics related to Machine Learning in corporate processes. Notably, the structured design of this maturity model ensures accessibility for small and medium-sized enterprises (SMEs).

Keywords: machine learning; maturity model; production planning and control; project success; SME; success factors

1 INTRODUCTION

With the rise of Artificial Intelligence (AI) technologies and applications, our society is witnessing a strong impact in day-to-day tasks through intelligent software solutions. Particularly prediction and generative models become more and more present in private and business environments. However, since most AI technologies require large amounts of training data, application-ready solutions are available mostly in domains where large amounts of (high quality) data is available. While this is clearly the case in internet-based information systems, where consumers voluntarily share and annotate texts, images, audios and videos, the situation is more difficult in industrial environments. Particularly in the realm of production planning and control (PPC) only few companies, are leveraging the potential brought about by AI and Machine Learning (ML). Theoretical research continues to push the boundaries of ML methods, while the industry, particularly small and medium-sized enterprises (SMEs), is primarily focused on implementing fundamental functions with respect to digitalization. Despite this, companies are eager to invest in AI disciplines such as ML due to the perceived importance of future business process enhancements. However, the realization of these benefits through practical implementation remains elusive.

A study of Gupta [1] shows that 78% of AI or ML projects do not reach a stage where they are ready for deployment in a productive environment. Fig. 1 shows the various reasons for early cancellation. In order to overcome such setbacks and improve the success rate of ML implementations, a systematic analysis of the ML-readiness and critical self-assessment and reflection is required by companies. This paper contributes to the state of the art by introducing a maturity model for companies to determine the degree of utilization of ML in PPC processes. To this end, Section 2 introduces the basic concepts required in the scope of this paper. Section 3 provides a systematic analysis of 14 existing maturity models, which results in the novel maturity model introduced in Section 4.

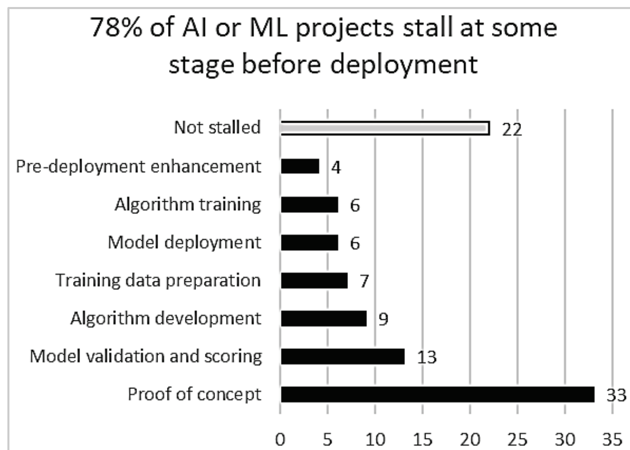


Figure 1 Machine Learning projects failure rate [1]

2 FOUNDATIONS

This section introduces the core concepts of maturity models, Machine Learning, and production planning and control processes, as a basis for the remainder of this paper.

2.1 Maturity Models

Maturity models serve the purpose of assessing the present status of an application and pinpointing areas for potential enhancement [2]. They have emerged as a valuable tool for facilitating the execution of organizational transformation processes, encompassing new projects, within a company, and in a synchronized manner. They assist in designing and implementing transformation projects efficiently, while also providing a comprehensive framework incorporating all required components of the transformation project. Maturity models outline diverse sequences of maturity degrees for various classes of objects, describing anticipated, desired, or typical development paths in discrete rank stages. This approach fosters a holistic procedure and allows for continuous improvement, leading to enhanced performance or quality. Each **maturity level** is characterized by specific traits and the requisite characteristics for advancement. The specific design fields within each maturity level model play a critical role in pinpointing additional elements that need to be addressed, aiding in further streamlining the complexity of the transformation project. A **design field** is characterized by a distinct capability area within the transformation project that effectively represents and organizes the subject matter. Additionally, factors for evaluating the single maturity levels of the model are essential components for further segmenting the design fields [3].

The creation or development of maturity models usually involves three principal phases in the design process [3]:

1) **Problem Definition:** In this initial phase, the focus is on delineating the design scope of the maturity model. This involves identifying the specific need for the maturity model and assessing the maturity of existing areas of application.
2) **Model Design:** Following an iterative approach, the design process unfolds through various steps, culminating in the establishment of the maturity model across distinct yet logically sequential maturity levels. Design areas are specified for examination, with each level linked to these defined areas. Detailed characteristics are then developed or replicated for each design area and its corresponding maturity levels.
3) **Evaluation:** Like any model, a maturity model undergoes evaluation to assess its structural integrity and the efficacy of methods employed in terms of validity, quality, usefulness, reliability, effectiveness, and generalizability. Furthermore, evaluation encompasses ensuring that the model aligns with client requirements and interests, particularly in terms of result provision. Ultimately, the evaluation aims to determine the extent to which the maturity model fulfills its intended benefits.

2.2 Machine Learning

Artificial Intelligence (AI) is a wide discipline of computer science that has been receiving great attention in the recent past. The increasing popularity is due to the success of machine learning technologies and, most recently, large language models, which yield an impact in day-to-day life of people. Success factors for these advancements are (i) the technological developments in the area of compute power, (ii) the availability of large amounts of data, and (iii) progress in research regarding methods and algorithms.

With machine learning (ML) as a sub-discipline of AI, the technologies applied and the tasks to be approached are limited compared to the universal and often exaggerated expectations of general AI. More precisely, ML comprises the technological approach to make a computer program calculate predictions based on a function that is learned from examples. The function that is learned from the training examples is called a (ML) model. Driven by the developments in massively parallel computer hardware, such as GPUs, the specific sub-discipline of deep learning emerged, which is based on artificial neural networks with a large number of neurons and layers, hence often called deep neural networks.

ML can be sub-divided into the fields of supervised learning, unsupervised learning and reinforcement learning. In supervised learning, the training examples are composed of a set of features, also called predictors, and a response, which is a variable dependent on the features. The goal of a supervised learning procedure is to find a function that predicts the response given the features. The learning process can be considered as supervised, since the response is known for the training examples. Once the function is learned, it can be applied to unknown data, i.e., data for which the response is not known, and the prediction will be correct with a certain accuracy. Artificial neural networks (ANNs) is a specific technology to solve supervised learning tasks. In unsupervised learning, there is also a set of features for each training example, but no response variable. Instead of computing predictions, the goal in unsupervised learning is to discover relations between training examples, such as clusters, i.e., subsets of the dataset that might have something in common. Last but not least, reinforcement learning is a sub-discipline of machine learning, where a system being in a given state takes an action to transition into a new state while receiving a reward as a feedback of the action taken. Given this feedback, a reinforcement learning system learns to maximize the rewards, thus learning to take the best action in any given state.

While the theoretical description of any of these ML paradigms is rather generic, it can be applied in a plethora of use cases solving all kinds of difficult problems, where the task can be mapped onto any of the three presented ML formalisms, and a sufficient base of training data of high quality is available. This means on the other hand, that a small data basis or data of bad quality (e.g., biased, flawed, or uncertain data) would lead to unsatisfactory ML models that do not lead to the desired performance in their respective application area.

2.3 PPC Process

Production planning and control processes are pivotal in effectively managing production systems, which encompass interconnected components facilitating information and material flow. These systems are typically divided into the management subsystem and the execution subsystem.

The management system serves to task to control and plan production, which involves defining target values to initiate the transformation process within the execution system. Feedback loops are established to ensure continual alignment between actual performance and desired outcomes. Various factors influencing production, such as delivery bottlenecks, machine malfunctions, or unforeseen disruptions, are identified and addressed through proactive measures [4].

Material Requirement Planning (MRP) and its evolved version, Manufacturing Resource Planning (MRP II), serve as vital tools in optimizing resource allocation and scheduling. They enable organizations to efficiently manage inventory, production orders, and capacity planning, thereby enhancing productivity and reducing operational costs.

Furthermore, the production planning and control system, integrated within the broader Enterprise Resource Planning (ERP) framework, orchestrates the seamless coordination of operational resources across different departments. ERP systems encompass multiple modules, including finance and accounting, human resources, and materials management, all interconnected through a centralized database.

This integrated approach not only streamlines organizational tasks but also facilitates data-driven decision-making, fostering agility and responsiveness to changing market demands. By optimizing production processes and resource utilization, organizations can enhance their competitiveness and achieve sustained growth in today's dynamic business environment [5].

2.4 Factors for Successful Adoption of ML Technologies

IT Infrastructure: A robust IT infrastructure is essential for successful ML implementations, as digitization levels affect ML efficacy. Considerations during planning should encompass scalable high-performance infrastructure to accommodate current and future needs. Incorporating structures like knowledge distillation and Quantum ML can enhance efficiency. Conversely, non-networked, heterogeneous infrastructures hinder data integration. Transforming legacy IT systems into interconnected networks is crucial for comprehensive ML application [12].

ML Competencies: Collaboration among multidisciplinary teams, including domain knowledge, fosters ML competency. Foundational ML knowledge is vital for employees and decision-makers to grasp ML's implications. Difficulties emerge due to the shortage of data scientists and personal inclinations. Financial constraints pose challenges for SMEs in attracting ML experts [24].

Machine Learning: Success factors rooted in scientific principles are essential for enterprise ML adoption. ML applications must be tailored to specific industry needs, emphasizing dimension reduction, visualization, and automation. Challenges include inductive biases, black box nature, and overfitting. Keeping pace with ML's rapid advancements poses difficulty in maintaining an overview [25].

Data: Data quality and availability are crucial for ML success. Challenges such as data silos, imbalanced datasets, and data drift hinder ML training. Implementing operational data governance can mitigate these challenges. Approaches like data lakes and federated ML show promise in overcoming data obstacles [26].

Corporate Culture: Management support and corporate strategy influence a company's readiness for ML adoption. A culture of open innovation encourages employee participation and fosters ML application development. Failure to adapt to changing market demands and immature approaches hinder successful ML integration [27].

Project Management: Effective project management is vital for AI application success. Detailed problem definition and use case development are prerequisites. Agile methodologies and regular progress measurements enhance project outcomes. ML implementation requires a thorough understanding of business processes and a collaborative approach. ML implementation presents technical, organizational, and general challenges. Prerequisites for successful implementation include a dissected structure tailored to specific domains and collaborative approaches. Overcoming challenges requires a collaborative effort and a commitment to continuous improvement [28].

3 ANALYSIS OF EXISTING MATURITY MODELS

Crafting a robust maturity model in the context of PPC requires a thorough review of existing models. Studying and comparing them plays a central role in developing novel maturity models. A particular focus is laid on dissecting design fields and their associated dimensions. This analysis helps identify new relevant design fields to be integrated into a novel model. The analysed models, as outlined in Tab. 1, are categorized based on their connections PPC processes and ML, AI, and data science. Moreover, these models are distinguished by their conceptual origins, as highlighted in Tab. 2. To achieve this, each selected model is briefly described and then their evaluation is illustrated through an example.

Table 1 Selected maturity models for further analysis

No.	Model	Source
1	Artificial Intelligence Maturity Model	[6]
2	Algorithmic Business Maturity Model	[7]
3	AI Maturity Assessment Model	[8]
4	AI Readiness Model	[9]
5	IBM Maturity Framework for Enterprise Applications	[10]
6	Google Cloud's AI Adoption Framework	[11]
7	ML Maturity Framework	[12]
8	ML Operations Maturity Model	[13]
9	The AI Maturity Framework	[14]
10	The Roadmap to ML Maturity	[15]
11	Data Science Maturity Model for Enterprise Assessment	[16]
12	Maturity Model for Digital Analytics & Optimization Maturity Index	[17]
13	Product-Process Framework for Smart PPC	[18]
14	Stage-Based Maturity Model for Industry 4.0 & PPC	[19]

The Artificial Intelligence Maturity Model improves organizational performance, while the **Algorithmic Business Maturity Model** aligns AI technology with clear business scenarios. The **AI Maturity Assessment Model** aids media service and communication providers in avoiding a standard solution approach.

Intel's AI Readiness Model supports organizations at any AI journey stage, aiding decision-makers in prioritizing efforts. **IBM's Maturity Framework for Enterprise Applications** assesses technical and business perspectives of artificial intelligence. **Google Cloud's AI Adoption Framework** facilitates custom transformative AI capabilities.

The **ML Maturity Framework** targets higher machine learning maturity irrespective of the starting point. The **ML Operations Maturity Model** clarifies Machine Learning

Operations (MLOps) principles. The **AI Maturity Framework** assists leaders in prioritizing impactful actions in their unique context.

Table 2 The evaluation matrix
(structurally following Lahrmann et al. [20] & Egeli [21])

Maturity Models		Artificial Intelligence Maturity Model	AI Maturity Assessment Model	AI Readiness Model	IBM Maturity Framework for Enterprise Applications	Google Cloud's AI Adoption Framework	ML Maturity Framework	The AI Maturity Framework	The Roadmap to ML Maturity	Data Science Maturity Model for Enterprise Assessment	Maturity Model for Digital Analytics & Optimization Maturity Index	Product-Process Framework for Smart PPC	Σ
No.		1	3	4	5	6	7	9	10	11	12	13	
Model Basics	Dimensions	4	5	13	7	6	9	5	4	10	6	9	
	Maturity Levels	5	4	3	3	3	5	5	3	5	6	3	
Scope	AI	•	•	•									
	ML				•	•	•	•	•				
	Data Science									•	•		
	PPC											•	
Origin	Academic	•					•					•	
	Practice		•	•	•	•		•	•	•	•		
Design Fields	**Dimensions**												
Production planning & control	Production											•	1
	In-house production planning & control											•	1
	External procurement planning & control												0
ML system properties	Scaling				•	•				•			3
	Automation					•							1
	Comprehensibility				•		•						2
	Reliability			•	•	•							3
	Acceptance & User friendliness			•	•							•	3
Depolyment	Deployment				•				•	•			3
Training	Training						•		•				2
Project management	Framework			•						•		•	3
	ML flow management			•					•				2
	Goal definition						•						1
Data	Data	•	•	•	•	•	•	•	•	•	•	•	11
IT infrastructure	IT infrastructure & Technology	•	•	•	•	•		•			•		7
	Tools			•					•	•			3
Coporate culture	Business culture										•		1
Expertise	Staff & Competences	•		•		•		•	•	•	•		7
Leadership & Strategy	Consumer value				•								1
	Strategy & Leadership	•	•	•		•		•		•	•	•	8
	Governance			•				•	•				3
	Cross-departmental collaboration								•	•	•	•	4

The Roadmap to ML Maturity divides into three levels, aiding in understanding ML sophistication and prioritizing success. The **Data Science Maturity Model for Enterprise Assessment** evaluates dimensions with five maturity levels.

The **Maturity Model for Digital Analytics & Optimization Maturity Index** reflects a company's current capabilities.

Models like the Product-Process Framework for Smart PPC and Stage-Based Maturity Model for Industry 4.0 & PPC offer incremental approaches in the realm of sustainability and Industry 4.0.

In order to precisely identify the design fields and their dimensions, an evaluation of the existing models with respect to methodology was conducted. To address gaps in the scientific understanding of classifying and designing maturity models for PPC processes, a deeper analysis regarding the content of selected models was carried out. This entails highlighting and weighing the various emphases of these models to incorporate their unique aspects into a more comprehensive model. Model no. 8, the "ML Operations Maturity Model" from Microsoft [13], was excluded from further analysis due to its unsuitable and unclear structure. Similarly, the "Algorithmic Business Maturity Model" according to Gentsch [7] (model no. 2) and the "Stage-Based Maturity Model for Industry 4.0 & PPC" according to Busch et al. [19] (model no. 14) were omitted as they deviate from our specified focus. Model no. 2 overly concentrates on business sectors like marketing and services, while model no. 14 primarily addresses Industry 4.0 issues alongside AI and other digital technologies, rather than emphasizing ML within PPC processes. Consequently, there are 11 remaining maturity models that are analysed further with respect to their content. These will be scrutinized according to their designated dimensions. Through this analysis, common topics and aspects shared among these models are identified. These serve as fundamental components (basic reference) for the new model. To avoid unnecessary complexity in evaluation, dimension groups are used to summarize the dimensions of the maturity models under consideration. The overall design fields examined in the evaluation matrix, as illustrated in Tab. 2, were **production planning & control**, **ML system properties**, **development**, **training**, **project management**, **data**, **IT infrastructure**, **corporate culture**, **expertise** and **leadership & strategy**. The maturity models under examination for further evaluation encompass two primary content domains: one emphasizing the implementation and fusion of data science, AI, or ML, and the other focusing on the modeling of PPC processes. Through comparing these models, it was possible to delve deeper into their distinct content emphases. This comparison revealed that dimensions related to data science, AI, ML, and PPC were selected based on specific practical considerations. However, there were also overlapping dimensions consistently present across multiple models in the analysis. For instance, technological aspects, like those in the realm of ML, were predominantly featured in practice-oriented maturity models. Conversely, IT infrastructure, data, or leadership & strategy, as examples for design fields, were addressed across nearly all examined models. An additional notable aspect is that while several maturity models associated with PPC exist, they do not incorporate a comprehensive approach to ML. If at all, such approaches typically pertain solely to internal planning and production unit allocation, overlooking the potential for external procurement planning and control.

Table 3 Overview of the maturity level assessment matrix

Assessment matrix for maturity levels								In-house valuation			
Design fields	Dimensions	Description	Criteria	Defined maturity levels				Maturity level of the current state			
				1	2	3	4	1	2	3	4
Organization	Leadership & Strategy	• Strategic positioning & attitude of/to ML in the company? • Measurement of utilization rate & efficiency?	• Strategy • Investment • Organizational attitude • KPI usage								
	ML competence	• Availability of know-how & competencies regarding ML in the company?	• Training • Personnel recruitment								
	Project management	• Task distribution in the development process? • Development specifications in project management available?	• Task distribution • Use case identification • Problem definition								
	Corporate culture	• Alignment of the in-house work culture with regard to ML? • Acceptance of ML by employees?	• Working culture • Technology acceptance								
Technology	IT infrastructure	• To what extent are IT structures interconnected? • How well do they perform?	• Networking • Performance								
	Data	• What it the importance of data for ML in the enterprise? • Availability of the necessary data present? • How is data security, protection & quality ensured?	• Data availability • Data protection • Data security • Data quality				focused level				
	ML software solutions	• What characteristics of ML based software solutions are desirable?	• Reproducibility • Reliability • Scalability								
PPC process	Production program planning	• Are real time data available for demand forecasting? • Who monitors the processes?	• Sales planning, primary requirements planning & oder processing • Resource rough planning								
	Production requirement planning	• Who determines the disposition parameters? • To what extent is human verification necessary in the process? • How are the data determined for the investigations?	• Gross & net secondary requirements determination • Lead time scheduling, capacity requirements determination & capacity reconciliation								
	In-house production planning & control	• Who selects the calculation method? • To what extent are real time data available? • Are approvals granted autonomously?	• Lot sizing • Detailed scheduling, detailed resource planning & sequencing • Availability check & order release								
	External procurement planning & control	• Who determines the order quantities? • Which processes need to be monitored & reviewed? • What forms are available for placing orders?	• Purchase order invoice • Request for quotation & supplier selection • Purchase order approval								

4 NOVEL MATURITY MODEL

This section presents the novel maturity model taking into account how PPC processes relate to ML, as well as the maturity classification using the example of the "Data" dimension.

4.1 Structure of the New Maturity Model

Thus far, it has been determined that 14 appropriate maturity models have been identified concerning the current subject matter through an assessment of existing maturity models. Through a more thorough analysis, three models were excluded due to their lack of alignment with the research focus, leaving 11 maturity models for further investigation. This underscores the absence of a mature model tailored to the specified research objective: the potential operational integration of ML in PPC, while considering entrepreneurial task areas. The majority of scrutinized maturity models on data science, AI, and ML primarily stem from practical applications. This situation once again highlights the scientific gap in this field, which holds greater significance within the industry.

The new maturity model's structural framework is devised to follow a grid-based format. Within this grid structure, 11 dimensions have been identified and organized into three design fields: organization, technology, and PPC process. In order to ensure clear presentation and enhance practical usability in this article, distinct grid layers are used to particion the individual dimensions of the entire maturity model. Within these layers, the all-embracing design fields are refined, alongside specifying the dimensions to be assessed and the criteria for evaluation. Together, these grid layers form a comprehensive maturity model. A notable benefit of employing this grid-based approach in crafting the maturity model is its flexibility to accommodate the preferences or requirements of the end user. This adaptability allows dimensions to be either omitted or expanded as per user needs. With the provided grid structure, adopting a systematic approach should present no difficulties. The following table offers a summary of the structural framework and allocation of various criteria, derived from the analysis and evaluation of the scrutinized maturity models.

Table 4 Section of the maturity model level for the "Data" dimension

Design field:		Technology
Dimension:		Data
Maturity level	Description	Characteristics
1	Data is not recognized as a strategic asset & as a result, is not enhanced or stored	***Data availability:*** Data assets are not considered strategically relevant & are consequently not stored or utilized, interfaces to extract data are scarce, and real-time data & data exchange are either nonexistent or limited
		Data protection: Country-specific data protection rights are only considered to the extent legally required
		Data security: Data integration occurs without security checks
		Data quality: Data is disregarded in terms of relevance, completeness, timeliness, and validity
2	Data is recognized as a strategic asset at the group or department level but has not yet been made available company-wide	***Data availability:*** Data is consciously recognized as a strategic asset & the necessary data sources are increasingly digitized, real-time data & critical data for individual processes are consistently available
		Data protection: Analogous to level 1
		Data security: Data integration takes place with irregular & unclear security checks
		Data quality: Data is checked to a limited extent for relevance, completeness, timeliness, validity, and format
3	Data is recognized as a strategic asset & made available company-wide	***Data availability:*** Analogous to level 2, in addition, there are digital twins of critical and relevant processes, generators, or facilities & data is made available company-wide
		Data protection: Country-specific & relevant data protection rights are considered beyond the necessary requirements, data that does not fall under data protection regulations or the company's internal governance is separately assessed, if such data does not comply with internal guidelines, it is neither collected, stored, nor further processed
		Data security: Analogous to level 2
		Data quality: Data is regularly checked for relevance, completeness, timeliness, validity, and format
4	Data assets are an integral part of the company's strategic direction & are made accessible across the entire organization	***Data availability:*** Analogous to level 3, in addition, data exchange takes place along the entire value chain, real-time data & critical data can be selected & exchanged with partner companies, service providers, etc., as needed
		Data protection: Analogous to level 3, in addition, further tools such as Federated Machine Learning (Federated ML) are used to generate or optimize ML models while considering the highest privacy guidelines
		Data security: Data integration takes place with proper data security checks, employees are regularly trained on data protection topics, depending on the use case, relevant data remains in secure & distributed environments, allowing only updates to ML models to be exchanged
		Data quality: Data is assessed for relevance, completeness, timeliness, validity, and format using AutoML & domain knowledge

For a maturity model to be practically applicable, it must allow for the individual evaluation of design fields and their subordinate dimensions. Thus, the number of maturity levels is determined based on a grid framework, which varies depending on the objective of the task or purpose at hand. The main objective in defining the number of levels should prioritize an approach that facilitates the seamless transition from applied research to practical implementation. In line with this approach, the number of maturity levels for the current model was set at four. A model with more than four maturity levels would not effectively support an approach that is practice-oriented anymore, as it would lack clarity regarding the dimensions described individually. With a subpartition of four levels, the design fields along with their dimensions can be defined explicitly without sacrificing comprehensibility. This allows the user to easily grasp all pertinent information regarding the concerned design fields and assess the current state of the process. Conversely, a too coarse division of maturity levels fails to provide a detailed categorization of the current state, typically indicating only the presence of a deficit without specifying its extent or the sequence of the process.

4.2 The Maturity Level of the Dimension

Tab. 3 (assessment matrix for maturity levels) provides an overview of individual design fields, dimensions, and possible criteria. In the following sections, we delve into the respective maturity levels of each dimension, providing a detailed description through the lens of an illustrative example. For a more comprehensive exploration of these maturity levels, readers are directed to Hartl [22].

As mentioned earlier, the dimensions of this maturity model consist of four maturity levels, with each dimension characterized by its specific criteria. The current maturity level can be determined based on these criteria. To put it in simpler terms, one could say that the maturity level increases with the awareness or automation of the interrelated elements in the affected dimensions. A prime example is the dimension "Data", as every ML concept relies on data (see Tab. 4).

Specific criteria within this dimension include **data availability**, **data protection**, **data security**, and **data quality**. In this example, all four maturity levels are described based on these criteria. In the first maturity level, generated data is not yet perceived as a strategic element in the company, and therefore, it is not further categorized concerning availability, protection, security, and quality. Moving to the second maturity level, data is recognized as an essential strategic element, but it is only collected, assessed, and/or shared at the group level or within individual departments. It is only from the third maturity level onwards that data is made available company-wide and assessed based on specific criteria and needs. In the final maturity level (level 4), data is considered an integral part of the company and is accordingly examined, evaluated, and protected, as outlined in the criteria described in Tab. 4. This structured approach has been applied to all 11 dimensions [22].

5 CONCLUSION

After reviewing literature and evaluating models, it was evident that change processes and related projects should transcend beyond the process level alone. By employing the three specified design fields—organization, technology, and PPC process—we underscore the pivotal role of organizational alignment in strategies, outcome measurement, project management, investment planning, work culture, and more. In this context, projects often face hurdles stemming from inadequate commitment from senior management or insufficient definitions of project goals. The incorporation of new technologies, like Machine Learning, is frequently underestimated or completely overlooked among employees. Similarly, the "Technology" design field emphasizes integrating innovative technologies to secure the company's long-term competitiveness. Particularly noteworthy is the Technology design field's significance in terms of system integration, IT process compatibility, and data-related advantages.

The integration of ML is poised to shape competitiveness across various forthcoming application domains and is increasingly perceived as a disruptive technology by many companies [23]. Nonetheless, the current ML research focus leans heavily towards technical aspects like learning algorithms, with relatively scant attention given to the organizational context and success factors tied to ML implementations, and the design principles of ML software solutions. This underscores the need to establish a comprehensive scientific foundation for the practical application of ML.

The aim of the model presented herein is to furnish companies, especially SMEs, with a structured approach to integrate ML technologies into PPC processes. Company-specific features and solutions must be devised within each company, as these dimensions are context-dependent. Given time and space constraints, not all aspects can be exhaustively covered. There are additional follow-up projects pending in this regard. Further exploration is warranted to evaluate and implement the mentioned dimensions alongside their maturity levels. Moreover, the model can be broadened to encompass additional company-specific processes through new design fields. Delving into solutions that are rarely used or not at all in current practice, yet are deemed relevant, would also be a promising extension in terms of abstraction, validity, consistency, and benefits.

6 REFERENCES

[1] Gupta, S. C. (2022). *Machine Learning for Developers.* https://www.ml4devs.com/newsletter/008-why-machine-learning-projects-fail/ (Accessed on July 09, 2023)

[2] Kamprath, N. (2011). *Einsatz von Reifegradmodellen im Prozessmanagement.* Gabler Verlag, 93-102. (in German) https://doi.org/10.1007/BF03340648

[3] Becker, J., Knackstedt, R. & Pöppelbus, J. (2009). *Entwicklung von Reifegradmodellen für das IT-Management: Vorgehensmodell und praktische Anwendung.* (pp. 218, 249-260) Gabler Verlag. (in German) https://doi.org/10.1007/s11576-009-0167-9

[4] Dyckhoff, H. & Spengler, T. S. (2010). *Produktionswirtschaft: Eine Einführung*. (3rd ed.) (pp. 6-7). Springer. (in German) https://doi.org/10.1007/978-3-642-13684-9

[5] Schuh, G. (2014). *Produktionsmanagement, Vol. 5* (2nd ed.). (pp. 281-286). Springer. (in German)

[6] Alsheiabni, S., Cheung Y. & Messom, C. (2019). Towards An Artificial Intelligence Maturity Model: From Science Fiction to Business Facts. *Proceedings of the Pacific Asia Conference on Information Systems (PACIS2019)*. https://api.semanticscholar.org/CorpusID:211165308 (Accessed on August 23, 2023)

[7] Gentsch, P. (2019). *Künstliche Intelligenz für Sales, Marketing und Service: Mit AI und Bots zu einem Algorithmic Business - Konzept und Best Practices*. 2. Auflage. Springer Gabler. (in German) https://doi.org/10.1007/978-3-658-25376-9

[8] Pringle T. & Zoller, E. (2018, June 14). *How to Achieve AI Maturity and Why It Matters*. https://www.amdocs.com/sites/default/files/filefield_paths/ai-maturity-model-whitepaper.pdf. (Accessed on August 23, 2023)

[9] Intel Corporation (2023), *White Paper: The AI Readiness Model*. https://www.intel.com/content/dam/www/public/us/en/documents/white-papers/ai-readiness-model-whitepaper.pdf. (Accessed on August 23, 2023)

[10] Vaish, R., Agrawal, A., Kapoor S. & Parkin, R. (2023). *AI maturity framework for enterprise applications.* https://www.ibm.com/watson/supply-chain/resources/ai-maturity/ (Accessed on August 23, 2023)

[11] Google Inc. (2023). Google Cloud's AI Adoption Framework. https://services.google.com/fh/files/misc/ai_adoption_framework_whitepaper.pdf. (Accessed on August 23, 2023)

[12] Akkiraju, R., Sinha, V., Xu, A., Mahmud, J., Gundecha, P., Liu, Z., Liu, X. & Schumacher, J. (2020). *Characterizing Machine Learning Processes: A Maturity Framework*. In: Fahland, D., Ghidini, C., Becker, J. & Dumas, M. (eds) *Business Process Management. BPM 2020. Lecture Notes in Computer Science, vol 12168*. Springer, Cham. https://doi.org/10.1007/978-3-030-58666-9_2

[13] Microsoft. (2023). *Machine Learning operations maturity model.* https://learn.microsoft.com/en-us/azure/architecture/example-scenario/mlops/mlops-maturity-model. (Accessed on August 23, 2023)

[14] Gagnon, G., Khan, J., Guay, M., Guegan, M., Hudson S. & Lin, W.-W. (2023). *The AI Maturity Framework.* https://s3.amazonaws.com/external_clips/3430107/AI-Maturity-Framework_White-Paper_EN.pdf?1589551996. (Accessed on August 23, 2023)

[15] Algorithmia Research (2019). *The Roadmap to Machine Learning Maturity.* https://em360tech.com/sites/default/files/2020-11/The%20Roadmap%20to%20Machine%20Learning%20Maturity%20v1.0%20Final.pdf (Accessed on August 23, 2023)

[16] Hornick, M. (June 16, 2020). *A Data Science Maturity Model for Enterprise Assessment.* https://www.oracle.com/a/devo/docs/data-science-maturity-model.pdf (Accessed on August 23, 2023)

[17] Termer, D. F. (2018). *Reifegradmodell zum Digital Analytics & Optimization Maturity Index (DAOMI).* https://www.bitkom.org/Bitkom/Publikationen/Reifegradmodell-zum-Digital-Analytics-Optimization-Maturity-Index-DAOMI.html (Accessed on August 23, 2023) (in German)

[18] Oluyisola, O. E., Sgarbossa, F. & Strandhagen, J. O. (2020). Smart Production Planning and Control: Concept, Use-Cases and Sustainability Implications. *Sustainability, 12*(9), 3791. https://doi.org/10.3390/su12093791

[19] Busch, M., Schuh, G., Kelzenberg, C. & de Lange, J. (2019). Development of Production Planning and Control through the Empowerment of Artificial Intelligence. *The Second IEEE International Conference on Artificial Intelligence for Industries (AI4I)*, Laguna Hills, CA, USA, 115-118. https://doi.org/10.1109/AI4I46381.2019.00037

[20] Lahrmann, G., Marx, F., Winter, R. & Wortmann, F. (2010). *Business Intelligence Maturity Models: An Overview*. https://www.alexandria.unisg.ch/Publikationen/72444

[21] Egeli, M. (2016). *Erfolgsfaktoren von Mobile Business: Ein Reifegradmodell zur digitalen Transformation von Unternehmen durch Mobile IT*. Springer, 12-15, 34-41. (in German) https://doi.org/10.1007/978-3-658-12771-8

[22] Hartl, J. J. L. (2023). A maturity model to determine the degree of utilization of machine learning in production planning and control processes. *Master's thesis*, Technische Hochschule Ingolstadt.

[23] Sachin (April 04, 2021). Machine Learning: 5 Industries that will See a Staggering Adoption of ML. https://www.rapyder.com/blogs/machine-learning-trends-2021/ (Accessed on September 30, 2023)

[24] Hartmann, E. A. (2020). Kompetzenzentwicklung für Maschinelles Lernen zur Konstituierung der digitalen Souveränität. In: Hartmann, E. A. (eds) *Digitalisierung souverän gestalten*. Springer Vieweg, Berlin, Heidelberg, 31-43. (in German) https://doi.org/10.1007/978-3-662-62377-0_3

[25] Gadekar, B. P. & Hiwarkar, T. (2023). From Business Objectives to Analytics and Machine Learning Solutions: A Framework for Conceptual Modeling. *International Journal of Advanced Research in Science, Communication and Technology (IJARSCT), 3*(1), 277-285. https://doi.org/10.48175/IJARSCT-7876

[26] Schacht, S. & Lanquillon, C., eds. (2019). *Blockchain und maschinelles Lernen: Wie das maschinelle Lernen und die Distributed-Ledger-Technologie voneinander profitieren.* Springer. (in German) https://doi.org/10.1007/978-3-662-60408-3

[27] Herden, T. T., Nitsche, B. & Gerlach, B. (2020). Overcoming Barriers in Supply Chain Analytics - Investigationg Measures in LSCM Organizations. *Logistics, 4*(1), 5. https://doi.org/10.3390/logistics4010005

[28] Pumplun, L., Tauchert, Ch. & Heidt, M. (2019). A New Organizational Chassis for Artificial Intelligence - Exploring Organizational Readiness Factors. *European Conference on Information Systems*, 2-8. https://api.semanticscholar.org/CorpusID:174773897

Authors' contacts:

Jakob Hartl
Technische Hochschule Ingolstadt,
Esplanade 10, D-85049 Ingolstadt, Germany

Jürgen Bock, Prof. Dr.
(Corresponding author)
Technische Hochschule Ingolstadt,
Esplanade 10, D-85049 Ingolstadt, Germany
-49 841 9348-3506
juergen.bock@thi.de

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240514150731

Original scientific paper
Received: 2024-05-14, Accepted: 2024-05-22

Multicriteria Optimisation of Machining Operations Using a Spreadsheet Model

Borut Buchmeister*, Iztok Palcic, Robert Ojstersek, Klemen Kovic, Aljaz Javernik

Abstract: The rapid development of new materials and tools demonstrates the need for efficient and feasible machining processes. Modern production systems must guarantee sustainable, flexible, productive and high-quality production at low cost. Therefore, the combination of technological data with advanced software solutions is very important, especially when using complex CNC machining systems, where the reliability of the technological information is crucial, as simple measures can make a very positive contribution to the productivity achieved. When optimising the machining parameters, costs and machining time must be taken into account. The article deals with the optimisation of the turning process with a large number of influencing variables (machine, workpiece material, tool, cutting parameters, costs, etc.) and two objectives: the fastest possible machining (minimisation of machining time) and the lowest possible machining costs. The model is designed in an Excel spreadsheet and the multi-criteria optimisation is carried out using the approximation method. Step by step, we can find the optimal processing regime for each selected case.

Keywords: cost; cutting parameters; machining; optimisation; time

1 INTRODUCTION

Technological progress is one of the most important development issues. The process that was called the "industrial revolution" in the 19th century because of the key role of industry was called the "scientific-technological revolution" in the 20th century because of the decisive role of science and the development of technology. In the 21st century, automation and computer programmes are gaining ground and replacing traditional machines. This is why we speak of an ongoing "information revolution" that has catalysed the evolution from Industry 4.0 to the emerging frontiers of Industry 5.0, reshaping the global economy.

Accelerated advances in computer, telecoms and information processing technology are enabling real economies of scale in service offerings and technology transfer. There are more and more new suppliers on the world market, the age of globalisation has virtually eliminated the importance of geographical distances, and response times are very short due to customer requirements.

Industrial processes are carried out in computer-integrated and flexible production systems that prioritise the quality of human resources and are designed to be sustainable and adaptable to new conditions at low cost. Rapid advances in the life sciences are also changing the demographic profile of the population and increasing human performance in all age groups.

The latest knowledge in the field of machining techniques opens up opportunities for the industry to successfully adapt to new conditions at a time when much knowledge is needed to reduce the gap with the most developed economies. Compliance with international standards is the key to the global market, so we must give this aspect our full attention. In order to be able to machine the increasingly complex workpieces economically, all process parameters must be adapted to the task at hand [1]. Those who master these challenges will remain competitive on the global market.

All processing companies face the same challenge: processing raw materials into an end product in the right quality, quantity and within the right time frame. Near net shape raw materials are crucial as they minimise waste and streamline production, optimising efficiency and reducing costs. Sustainability and environmental aspects must also be taken into account without compromising competitiveness and profitability. A perfect example of today's process improvement efforts is Industry 4.0, which includes cutting-edge technology for capturing, storing and sharing manufacturing process data [2]. For companies with smaller capacities compared to global giants, increasing productivity often seems out of reach. However, simple, low-cost analyses and measures can have a very positive impact on the productivity of SMEs.

Machining plays an important role in the manufacturing industry and the optimisation of cutting parameters is a key component in the planning of machining processes. This is also confirmed by the abundance of research publications [3]. The cutting tool is directly involved in the optimisation process and is the first element of renewal when the technology changes. The tool must be correctly selected according to the machine's limit capacities in order to achieve the required productivity, quality and low costs. Optimal solutions are only possible if all factors influencing the cutting process are taken into account, creating the most suitable cutting conditions. It is necessary to ensure the best possible combination of process parameters for stable operation of the process within the product requirements.

The profitability of cutting processes is based on the production costs per product [4]. There is a non-linear cost relationship between chip volume per unit of time and tool life. A higher chip volume per unit of time generally increases productivity more, as the machining costs increase. Optimisation begins with the choice of cutting speed, which can be determined according to the tool manufacturer's guidelines, with in-house tests (which are associated with high costs) or with the recalculation of the durability/cutting speed curves for the combinations of machined material – tool – cutting material. In our daily work, we often do not have the capacity to analyse production processes and optimise them for even greater efficiency. We also often do not have enough time to adapt new cutting materials, tool geometries or process technologies to individual machining tasks.

Manufacturers today face a number of new challenges in terms of ensuring sustainability and environmental protection, which are being solved with new technologies and processes [5, 6]. Dry processing, for example, helps to reduce the use of coolants, which has less impact on the environment and lowers costs. The increasing use of lead-free raw materials etc. also avoids the impact of hazardous metals on the environment. Optimised processing parameters also mean energy savings [7].

2 PROCESS OPTIMISATION

The general budgeting approach in manufacturing companies is to procure all necessary elements at the lowest possible price. However, low price alone is not the best basis for choosing a tool, as the end result (product quality, processing time) should also be taken into account. We are also bound by practical constraints, such as the strength and stability of the machine, dimensional tolerance requirements, surface roughness, etc. It should be emphasised that it is not necessary to exceed the requirements in the product specifications.

Depending on the number of tool geometries, their sizes and materials, the number of possible configurations of cutting tools is practically infinite. In the workshops, decisions are usually made on the basis of single operations, mainly one tool for each shape (detail) on the workpiece.

Manufacturers try to reduce processing times (with increased cutting parameters) but neglect other activities that increase product lead times (downtime due to tool failure, poor quality, inadequate chips, etc.). Factors affecting total processing costs include tools or tooling systems, workpiece materials, processes, personnel, organisation, maintenance, equipment; there are also many random influences.

2.1 Selection of Machining Parameters

After deciding on the working method and selecting the appropriate tool, the machining parameters must be determined: cutting speed, feed rate and depth of cut [6]. Higher values for these parameters mean higher productivity, higher tool wear and higher tool costs (and vice versa).

Economic aspects are decisive for the cutting speed. In principle, we want to work at the lowest possible cost when there is no rush. In rough machining, we are limited by the power and rigidity of the machine tool, the properties of the cutting tool and the strength of the workpiece. Of course we want to utilise the power, but not at the expense of permanent deformation of the machine, the tool or the workpiece. The aim of roughing is to remove as much material as possible in a unit of time at as little cost as possible, and the cutting parameters chosen are as high as the machine, cutting tool and workpiece will allow [8]. In fine machining, we want to achieve the appropriate dimensions and surface quality of the workpiece [9]; all the machining parameters mentioned above as well as the geometry of the cutting insert, the lubricant, etc. have a significant influence. The machining parameters must be selected so that the required machining accuracy is achieved, the capacity of the machine is utilised and thus the machining costs are reduced. In general, we only maintain relatively high cutting speeds.

Detailed calculations are required for larger batches, as this is economically justifiable. For small batches, detailed analyses are more expensive than the savings or benefits, so in this case the use of general guidelines and experience prevails.

2.2 Modelling of Turning Process

Turning is a continuous cutting process with a single cutting edge and a constant chip cross-section that is mainly used for the manufacture of round products. Turning is categorised as a conventional machining process, which means that we know the geometry of the workpiece and the tool [10]. Longitudinal turning, in which the direction of the feed motion is parallel to the axis of the workpiece, and transverse turning, in which the direction of the feed motion is perpendicular to the direction of the workpiece, are predominant (Fig. 1).

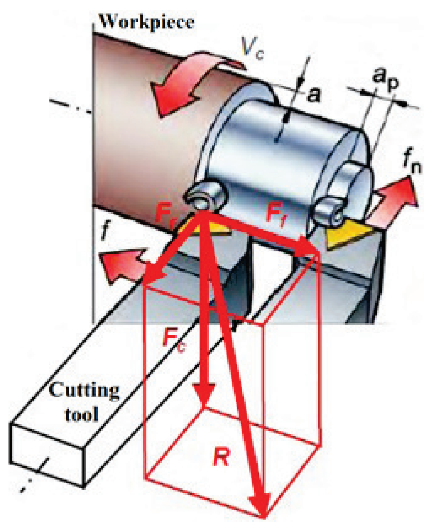


Figure 1 Presentation of longitudinal and face turning process and cutting forces

Optimisation is about finding the optimum cutting parameters in terms of costs and processing time, as this is the only way to achieve the maximum technical and economic effect with the available manpower, machines and tools. If there is sufficient capacity and there is no rush, we look for the cheapest machining; if costs are less important and you are in a hurry or there are bottlenecks, we look for the fastest machining (highest productivity). Another goal of optimisation can be to maximise the profit margin on processing. The processing costs are extremely important as they must be lower than the added value in order to make a profit [11].

The existing optimisation models are based on the selection of the optimum cutting speed and its influence on the optimised variables as shown in Fig. 2.

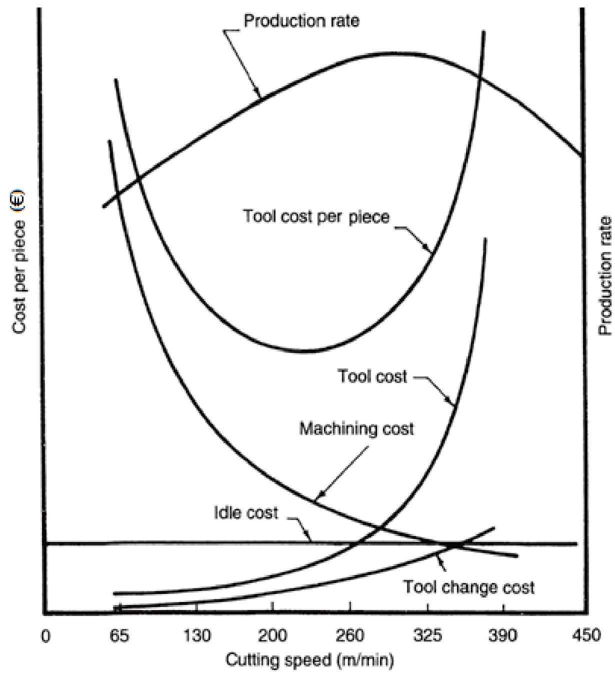


Figure 2 Effect of cutting speed on machining costs and productivity [12]

In this article, we therefore aim to provide an extended model that comprehensively addresses the cutting process in terms of machining parameters and machining economics (the area of interest is circled in Fig. 3).

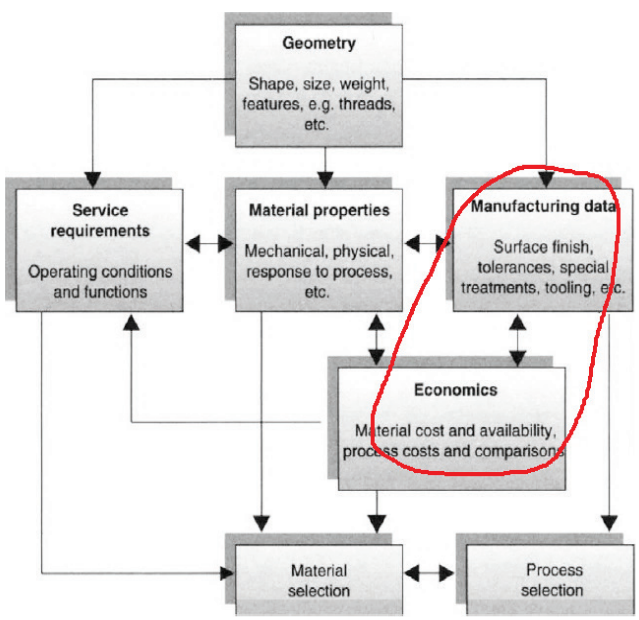


Figure 3 Factors of material selection and technological processes [13]

After choosing the machining method and tool, we need to make an important decision regarding three process parameters: cutting speed, feed rate and depth of cut [14]. To calculate the cutting speed, we use the extended Taylor equation [12] for turning with uncoated turning tools:

$$v_c T_L^n a^q f^p = C_v V_B^m K_{mv} K_\tau , \quad (1)$$

where: v_c – cutting speed (m/min); T_L – tool life (min); n – exponent of tool life; a – depth of cut (mm); q – exponent of depth of cut effect; f – feed rate (mm); p – exponent of feed effect; C_v – constant, the equivalent of cutting speed with one minute of tool life; V_B – flank wear value (mm); m – exponent of tool wear effect; K_{mv} – coefficient of material hardness; K_τ – coefficient of tool cutting edge angle.

The basic assumptions and limitations of the model are:

- a stable lathe – the rigidity or robustness of the machine is adequate,
- good maintenance condition of the machine,
- the positioning accuracy is less than 0.01 mm,
- clamping of workpieces from one side is possible up to a maximum length : diameter ratio of 2 : 1,
- a steady rest and a tailstock are available for long or thin-walled workpieces,
- the clamping point for the tool is stable, with minimal overhang,
- the edges and radii of the workpiece are properly bevelled, ensuring smooth entry and exit of the tool,
- the durability of the tool corresponds to the manufacturer's specifications,
- the foundation of the machine is stable – to avoid vibrations from other machines,
- the system also has uncontrolled influencing variables (undesirable sources of variability), such as changes in the quality of the material of the workpiece and tools, variations in temperature and humidity in the room (environmental influences), etc.

The model was created in an MS Excel spreadsheet and contains the following variables and valid (generally known) relationships between them:

- Workpiece: rough (initial) diameter, finished diameter, length of step (cut), number of passes (cuts), depth of cut, feed, Taylor constant and exponents, batch size, yield, manipulation time.
- Machine tool: operating cost per hour, operator's salary, effective power, machine set-up time, unplanned downtime.
- Cutting tool: toolholder cost, insert cost, number of cutting edges, maximum allowed wear (VB), maximum tool life (carbide quality), tool set-up time, insert index time.

The **steps** of our simple optimisation process are:

1) Initial calculation of the parameters to check the suitability of the machine tool.
2) Multi-stage calculation of the possible number of passes, taking into account the insert constraints and the performance of the machine tool.
3) Selection of the best result and definition of the final cutting speed with approximation to the limit values.

The results of the model calculation include: maximum cutting speed due to power limits, optimal cutting speed, tool life, machining cost and total production time per batch.

The optimisation of the cutting parameters is carried out using the example of longitudinal turning of a cylindrical step. The aim of the optimisation is to find machining conditions in which minimum machining costs and minimum machining time (maximum productivity) are achieved at almost the same cutting speed. We tolerate up to 2.5 % deviation from the optimum.

3 APPLYING THE MODEL TO THE SELECTED CASES

3.1 Case 1 – Turning on Conventional Universal Lathe

The workpiece material is a special low-carbon structural steel EN 6CrMo15-5 (Ravne CT194), which is normally used for gears, crankshafts, etc. Other related data for the example: Taylor constant and exponents (C_v, m, n, p, q) are 331, 0.46, 0.20, 0.25 and 0.1; K_{mv} is 1.13, K_τ is 0.80.

Toolholder cost is 130 €, ISO code: DCLNR3225P16; insert cost is 9 €, ISO code: CNMG160608, double sided, P25 grade, 4 cutting edges, maximum allowed wear (V_B) is 0.1 mm, maximum tool life (T_L) is 30 minutes.

Machining with longitudinal turning of a cylindrical step with a diameter of 150 mm to 80 mm and a length of 250 mm is required. The batch size is 100 parts. The yield is 95 %.

Machine tool hourly cost is 150 €/h. Operator salary cost is 16 €/h. Machine effective power is 18 kW. Conventional universal lathe is used.

The operator needs 4.2 minutes to manipulate the workpiece, the set-up time for the machine is 75 minutes. The set-up time for the tool is 16 minutes and insert index time is 2 minutes. The proportion of unplanned downtime is 15 %.

From the tool manufacturer we have the following guidelines for the use of the turning tool: the cutting depth should be in the range: 1.5 to 8 mm, the feed rate should be between 0.2 and 0.7 mm.

To test the model and the machine tool suitability, we start with the values from the centre of the intervals: the initial depth of cut (a) is 5 mm (7 passes) and the feed rate (f) is 0.45 mm. We obtain the following results of the model:

- The maximum cutting speed (v_{max}) due to machine power limits is 254 m/min.
- The minimum machining costs per part (74.6 €) is achieved at a cutting speed of 195 m/min and a total production time of 2150.0 minutes per batch (2.938 parts/h).
- The maximum productivity of 3.135 parts/h (2014.3 minutes per batch) is achieved at a cutting speed of 295 m/min with machining costs of 80.3 € per part. Due to limited cutting speed, this is not achievable.

The cost difference is 7.6 % and the productivity difference is 6.7 %, both more than the 2.5 % we had previously determined.

By running the model for different combinations of machining parameters, making appropriate use of the effective power of the machine (the average of the cutting speeds must be within the speed limit) and the tooling guidelines, we obtain a range of solutions (see the data in Tab. 1 and the graphic presentation in Fig. 4).

Table 1 Cutting speeds for minimum machining costs and maximum productivity at different machining parameters (case 1)

No. of passes	a (mm)	f (mm)	v_{cost} (m/min)	Cost (€)	v_{prod} (m/min)	Total prod. time (min)
5	7	0.27	215	79.4	320	2131.4
6	5.83	0.37	205	75.0	305	2025.6
7	5	0.48	195	71.9	290	1948.1
8	4.38	0.58	190	70.7	280	1919.3
9	**3.89**	**0.70**	**180**	**68.8**	**270**	**1874.4**
10	3.50	0.70	185	74.0	270	1999.9
11	3.18	0.70	185	79.1	275	2124.0

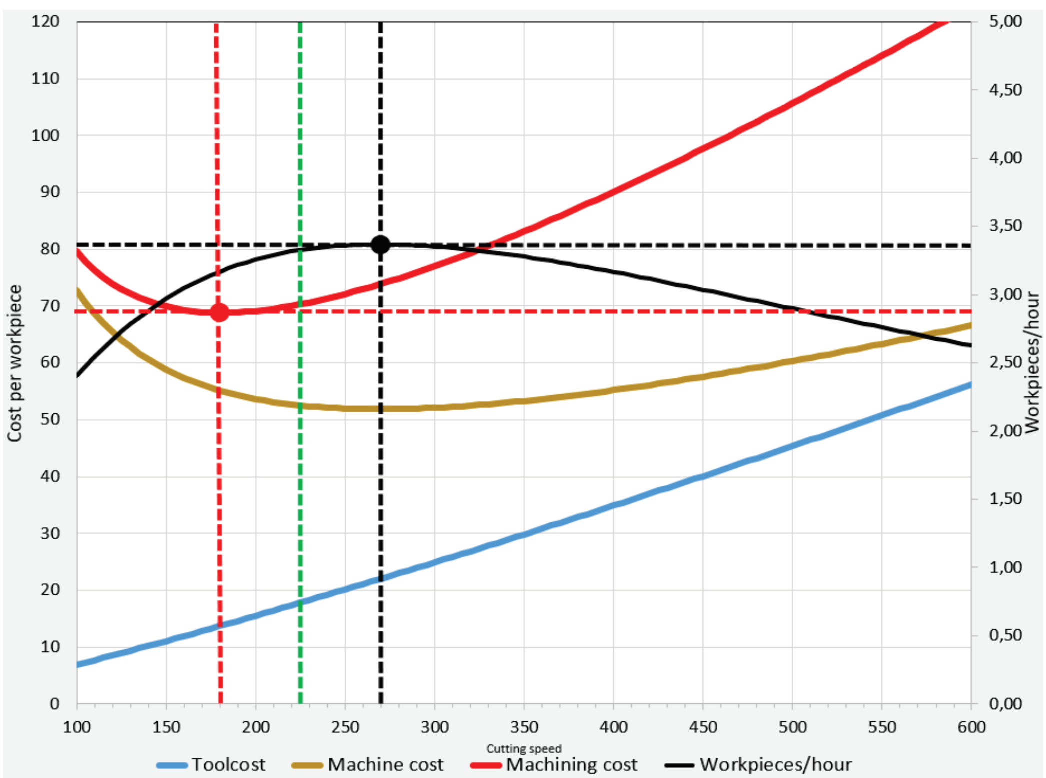


Figure 4 Cost per workpiece and the number of products per hour versus cutting speed

The best results are achieved with 9 passes. The simulation of the costs per workpiece and the productivity achieved as a function of the cutting speed is shown in Fig. 4. Assuming that the two target variables have the same weight, we can consider the average cutting speed as optimal; thus we obtain the optimal feasible solution with the cutting parameters: cutting speed **225 m/min** (green line in Fig. 4), cutting depth 3.89 mm and feed 0.7 mm, which results in a total production time of 1898.0 min per batch (3.328 parts/h) at a cost per part of 70.3 €. The deviation from the ideal productivity is 1.2 % (lower) and from the minimum machining costs 2.2 % (higher). Both within 2.5 %. The maximum cutting speed is 230 m/min (due to the power limitation).

3.2 Case 2 – Turning on CNC Lathe

The workpiece material is a gray cast iron EN-GJL-250 (SL 25) with the following data: Taylor constant and exponents (C_v, m, n, p, q) are 533, 0.46, 0.25, 0.25 and 0.1; K_{mv} is 0.66, K_τ is 0.80. Material hardness is 200 HB.

Toolholder cost is 240 €, ISO code: A20R-PCLNR09, min. internal diameter of 25 mm, max. cutting length is 200 mm; insert cost is 8.9 €, ISO code: CNMG090308-M3, TP2501, double sided, 4 cutting edges, maximum allowed wear (V_B) is 0.1 mm, maximum tool life (T_L) is 30 minutes.

Machining with longitudinal internal turning with a diameter of 32 mm to 84 mm and a length of cut of 170 mm is required. The batch size is 60 parts. The yield is 95 %.

Machine tool hourly cost is 170 €/h. Operator salary cost is 25 €/h. Machine effective power is 7.35 kW. CNC lathe has max. spindle speed of 3000 rpm and 8 turret tools.

The operator needs 1.1 minutes for all workpiece manipulations, the set-up time for the machine is 24 minutes. The set-up time for the tool is 12 minutes and insert index time is 0.8 minute. The proportion of unplanned downtime is 14 %.

From the tool manufacturer we have the following guidelines for the use of the internal turning tool: the cutting depth should be in the range: 0.5 to 3.5 mm, the feed rate should be between 0.12 and 0.36 mm.

We start with the values from the centre of the intervals: the initial depth of cut (a) is 2 mm (13 passes) and the feed rate (f) is 0.24 mm. We obtain the following results of the model:

- The maximum theoretical cutting speed (v_{max}) due to machine power limits is 426 m/min.
- The minimum machining costs per part (77.0 €) is achieved at a cutting speed of 250 m/min and a total production time of 1022.1 minutes per batch (3.707 parts/h).
- The maximum productivity of 4.173 parts/h (908.1 minutes per batch) is achieved at a cutting speed of 410 m/min with machining costs of 87.9 € per part.

The cost difference is 14.2 % and the productivity difference is 12.6 %.

Running the model for different combinations of machining parameters (the average of the cutting speeds must be within the speed limit) gives the results shown in Tab. 2.

Table 2 Cutting speeds for minimum machining costs and maximum productivity at different machining parameters (case 2)

No. of passes	a (mm)	f (mm)	v_{cost} (m/min)	Cost (€)	v_{prod} (m/min)	Total prod. time (min)
8	3.25	0.17	260	65.5	425	779.4
9	2.89	0.20	250	64.6	415	769.0
10	2.60	0.25	240	60.5	395	723.6
11	2.36	0.29	235	59.2	385	708.4
12	**2.17**	**0.33**	**230**	**58.2**	**375**	**697.6**
13	2	0.36	225	58.5	370	701.5
14	1.86	0.36	225	62.1	375	741.6

The best results are achieved with 12 passes. Again we can consider the average cutting speed as optimal (for equal importance of both criteria); thus we obtain the optimal feasible solution with the cutting parameters: cutting speed **303 m/min**, cutting depth 2.17 mm and feed 0.33 mm, which results in a total production time of 713.6 min per batch (5.312 parts/h) at a cost per part of 60.7 €. The deviation from the ideal productivity is 2.3 % (lower) and from the minimum machining costs 4.3 % (higher). Achieved productivity is within the 2.5 % tolerance. The maximum cutting speed is 305 m/min (due to the power limitation).

By lowering the cutting speed, we can lower the machining costs (at the expense of productivity). With the balanced cutting speed of 290 m/min we can get a total production time of 720.6 min per batch (3.3 % lower productivity) and 60 € cost per part (2.6 % higher).

3.3 Discussion

In the optimisation model, we have taken into account the generally known equations for the turning process and for calculating the machining costs. Since the number of passes is an integer, we have a limited number of combinations of machining conditions. The model simulates all possible combinations of cutting parameters, taking into account the effective power of the machine and the limits of the tool (depth of cut, feed rate). Since the optimal values to achieve the lowest cost and the highest productivity are different, we suggest the middle value of the cutting speed as the optimal choice. We have chosen this because we only have two optimization variables of equal importance. If the importance differs, the value within the cutting speed interval could also be shifted towards a more important value. Based on the two examples shown, we can see that the deviations from the individual optima are less than 5 %.

The optimization idea used is universal and can be applied to other turning cases as well as other types of cutting operations. The method is simple, feasible, practically useful (spreadsheet model) and understandable, as we only have two optimization variables. It is supporting process planner's work.

Compared to other authors, there is little similarity in the methods and models used. Anand et al. [1] consider the machining process from the point of view of friction and heat generation in the cutting zone, which have a significant influence on tool life and the quality of the workpiece surface. Cesén et al. [4] start from the recommendations of the tool manufacturers, determine the Taylor constant experimentally and optimize the cutting speed in order to achieve minimum machining costs. Abbas et al. [6] present a very complex multi-criteria optimisation (minimum consumption of energy, tools, machining time and costs; maximum productivity and surface quality) with different algorithms: Gray Wolf Optimizer, Weighted Value Gray Wolf Optimizer, Multi-Objective Genetic Algorithm and Multi-Objective Pareto Search Algorithm. Agarwal and Khare [7] minimise energy consumption, processing time and costs with a mathematical multi-objective model. Pujiyanto et al. [8] emphasise sustainable production and minimise energy consumption, surface roughness, noise, cost and carbon emissions with a special multi-criteria algorithm in Matlab. Vukelic et al. [9] optimise the surface roughness of workpieces in the turning process using a regression model. Wakjira et al. [10] use a mathematical analysis to obtain optimal cutting parameters for minimum forces and power requirements. Pangestu et al. [11] develop a multi-objective optimisation model to minimise energy consumption, carbon emissions, production time and production cost. Verma and Pradhan [14] use finite element simulation of the turning process and predict temperature, forces and strains to select optimal cutting inserts. Jiang et al. [15] investigate how to minimise the environmental cost and maximise the economic benefits of turning. They use the concept of the multi-objective optimisation algorithm NSGA-II.

A comparison of the articles cited shows that our optimisation method is original, easy to apply in practise and effective.

The future digitalisation of the presented optimisation method for turning is very promising. With advances in computing power, the expansion of simple computational models to more sophisticated systems, machine learning algorithms and data analysis, it is increasingly possible to develop digital solutions for optimising cutting parameters in machining processes such as turning. The implementation steps should include:

Data collection and analysis: The first step is to collect data on various cutting parameters such as cutting speed, feed rate, depth of cut, tool material, workpiece material, etc. This data can be collected from machining operations within the company or from existing databases. Once collected, advanced analysis techniques can be used to analyse the data and identify patterns and correlations between different parameters and machining results (e.g. surface finish, tool wear, machining time).

Machine learning models: Machine learning algorithms can then be trained using the collected data to develop predictive models for optimising the cutting parameters. These models can learn from previous machining experience to recommend the most appropriate combination of parameters for a particular machining task. Techniques such as regression analysis, decision trees or neural networks can be used for this purpose.

Integration into existing systems: The digital optimisation tool can be integrated into existing business systems such as Manufacturing Execution Systems (MES), Computer-Aided Manufacturing (CAM) software or even directly into the control systems of CNC machines. Integration ensures seamless communication between the optimisation tool and the machining environment, enabling real-time adjustments to cutting parameters based on changing conditions or requirements.

User interface: An intuitive user interface should be developed that allows machine operators and engineers to easily interact with the optimisation tool. The interface can provide recommendations for optimal cutting parameters based on user input such as desired machining outcomes (e.g. minimising tool wear or cost, maximising productivity) and constraints (e.g. machine capabilities, material properties).

Feedback loop: Continuous improvements are essential to increase the effectiveness of the optimisation tool. Feedback mechanisms should be implemented to collect data on the actual performance of the recommended cutting parameters during machining operations. This data can then be used to refine the machine learning models to improve their accuracy and relevance over time.

Overall, digitising the optimisation of cutting parameters in machining offers considerable potential for improving the efficiency, quality and cost effectiveness of machining processes.

4 CONCLUSION

Companies are increasingly recognising that large investments, the introduction of high-tech solutions or additional employment are not necessary to solve productivity problems. In a fast, cost-effective and sustainable future, intelligent software paves the way for optimised production and better decision-making capabilities. Digital technology and smart software have already transformed the hardware industry, as actionable data for better results has never been more readily available [16]. The right approach and the right knowledge can solve key problems. Optimising cutting parameters during machining is crucial for low machining costs, high productivity and sustainable production. With properly set parameters, we can achieve exceptional results, which include reducing material waste, optimising machining time and extending tool life [15].

When planning the optimisation of cutting parameters, several factors must be taken into account, including the material of the workpiece, the type of cutting tool, the cutting speed, the cutting depth and the feed rate. The combination of precision machining and advanced cutting tools ensures incredible productivity in metalworking. With a focus on low machining costs, the key is to achieve an optimal balance between cutting speed and tool life, which enables a reduction in production costs without compromising on machining quality. With a suitable traditional model, we can capture the influencing factors and arrive at feasible multi-criteria optimisation solutions. In this article, this is demonstrated using the example of longitudinal turning of the workpiece.

At the same time, however, it is also important to consider the sustainable aspects of production. Lowering energy consumption, reducing waste and extending tool life are key elements of sustainable production that can be achieved through the correct optimisation of cutting parameters. This extension of optimisation will be the goal of our future research.

By constantly monitoring, adjusting and improving cutting parameters, manufacturing companies can achieve exceptional results that combine the above-mentioned low machining costs, high productivity and sustainable production. Careful planning and implementation of cutting parameter optimisation is therefore a key strategy for achieving a competitive advantage in modern production. In the near future, the use of artificial intelligence will most likely completely change the way we approach solving the above problem.

5 REFERENCES

[1] Anand, A., Behera, A. K. & Das, S. R. (2019). An overview on economic machining of hardened steels by hard turning and its process variables. *Manufacturing Review*, *6*, Paper 4. https://doi.org/10.1051/mfreview/2019002

[2] Oztemel, E. & Gursev, S. (2020) Literature review of Industry 4.0 and related technologies. *Journal of Intelligent Manufacturing*, *31*(1), 127-182. https://doi.org/10.1007/s10845-018-1433-8

[3] Prasetyo, H. N., Sarno, R., Wijaya, D. R., Budiraharjo, R. & Waspada, I. (2023). Sampling simulation in process discovery. *International Journal of Simulation Modelling*, *22*(1), 17-28. https://doi.org/10.2507/IJSIMM22-1-619

[4] Cesén, M., Vila, C., Ayabaca, C., Zambrano, I., Valverde, J. & Fuentes, P. (2022). Cost optimization of the AISI-1018 turning process under sustainable manufacturing. *Materials Today – Proceedings*, *49*(Part 1), 58-63. https://doi.org/10.1016/j.matpr.2021.07.473

[5] Moran, T., MacDonald, R. & Zhang, H. (2023). A dynamic simulation model for understanding sustainability of machining operation. *Sustainability*, *15*(1), Paper 152. https://doi.org/10.3390/su15010152

[6] Abbas, A. T., Al-Abduljabbar, A. A., El Rayes, M. M., Benyahia, F., Abdelgaliel, I. H. & Elkaseer, A. (2023). Multi-objective optimization of performance indicators in turning of AISI 1045 under dry cutting conditions. *Metal*s, *13*(1), Paper 96. https://doi.org/10.3390/met13010096

[7] Agarwal, G. & Khare, M. K. (2021). Multiobjective optimisation of cutting parameters in machining – a sustainable approach. *Materials Today – Proceedings*, 46(Part 11), 5535-5543. https://doi.org/10.1016/j.matpr.2020.09.272

[8] Pujiyanto, E., Rosyidi, C. N., Hisjam, M. & Liquddanu, E. (2022). Sustainable multi-objective optimization of a machining parameter model for multi-pass turning processes. *Cogent Engineering*, *9*(1), Paper 2108154. https://doi.org/10.1080/23311916.2022.2108154

[9] Vukelic, D., Prica, M., Ivanov, V., Jovicic, G., Budak, I. & Luzanin, O. (2022). Optimization of surface roughness based on turning parameters and insert geometry. *International Journal of Simulation Modelling*, *21*(3), 417-428. https://doi.org/10.2507/IJSIMM21-3-607

[10] Wakjira, M. W., Altenbach, H. & Ramulu, P. J. (2021). Cutting mechanics analysis in turning process to optimise product sustainability. *Advances in Materials and Processing Technologies*, *7*(3), 446-462. https://doi.org/10.1080/2374068X.2020.1785207

[11] Pangestu, P., Pujiyanto, E. & Rosyidi, C. N. (2021). Multi-objective cutting parameter optimization model of multi-pass turning in CNC machines for sustainable manufacturing. *Heliyon*, *7*(2), Paper e06043. https://doi.org/10.1016/j.heliyon.2021.e06043

[12] Halevi, G. (2003). *Process and operation planning*. Springer Science+Business Media, Dordrecht. https://doi.org/10.1007/978-94-017-0259-1

[13] Scallan, P. (2003). *Process planning*. Butterworth-Heinemann, Amsterdam. https://doi.org/10.1016/B978-075065129-5/50003-9

[14] Verma, M. & Pradhan, S. K. (2020). Experimental and numerical investigations in CNC turning for different combinations of tool inserts and workpiece material. *Materials Today – Proceedings*, *27*(Part 3), 2736-2743. https://doi.org/10.1016/j.matpr.2019.12.193

[15] Jiang, Z. P., Gao, D., Lu, Y., Shang, Z. D. & Kong, L. H. (2020). Optimisation of cutting parameters for minimising carbon emissions and cost in the turning process. *Proceedings of the Institution of Mechanical Engineers Part C – Journal of Mechanical Engineering Science*, *236*(4), 1973-1985. https://doi.org/10.1177/0954406220922872

[16] Ghosh, A. K., Fattahi, S. & Ura, S. (2023). Towards developing big data analytics for machining decision-making. *Journal of Manufacturing and Materials Processing*, *7*(5), Paper 159. https://doi.org/10.3390/jmmp7050159

Authors' contacts:

Borut Buchmeister, Full Prof. Dr. Sc.
(Corresponding author)
University of Maribor, Faculty of Mechanical Engineering
Smetanova 17, 2000 Maribor, Slovenia
Tel. +386-2-2207631, e-mail: borut.buchmeister@um.si

Iztok Palčič, Full Prof. Dr. Sc.
University of Maribor, Faculty of Mechanical Engineering
Smetanova 17, 2000 Maribor, Slovenia
Tel. +386-2-2207636, e-mail: iztok.palcic@um.si

Robert Ojsteršek, Asst. Prof. Dr. Sc.
University of Maribor, Faculty of Mechanical Engineering
Smetanova 17, 2000 Maribor, Slovenia
Tel. +386-2-2207585, e-mail: robert.ojstersek@um.si

Klemen Kovič, M.Sc., Assistant, PhD student
University of Maribor, Faculty of Mechanical Engineering
Smetanova 17, 2000 Maribor, Slovenia
Tel. +386-2-2207514, e-mail: klemen.kovic@um.si

Aljaž Javernik, M.Sc., Assistant, PhD student
University of Maribor, Faculty of Mechanical Engineering
Smetanova 17, 2000 Maribor, Slovenia
Tel. +386-2-2207585, e-mail: aljaz.javernik@um.si

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240516125344

Original scientific paper
Received: 2024-05-16, Accepted: 2024-05-20

Addressing Circularity Strategies by Reconfiguring Smart Products during Their Lifecycle

Yannick Juresa*, Damun Mollahassani, Jens C. Göbel

Abstract: Considering circularity aspects during the engineering of smart products implies a significantly increased complexity of engineering processes. Especially technical reconfiguration of smart products, offered as services in availability-oriented business models, enables the integration of circular economy aspects in sustainable products and lifecycles, through realizing several aspects of the 9R strategies. This paper introduces an analysis of interdependencies between 9R strategies potential reconfiguration options, technical characteristics of smart products, different maturity levels of smart product and abilities for circularity. Partial engineering models managed in different product lifecycle management systems provide a technical basis for systematization and evaluation of circular abilities of reconfigurable smart products in different lifecycle phases. The approach aims to improve circularity-related decision making in systems engineering processes in the early development phases and during the reconfiguration of smart products during utilization phase. An industrial use case considering a microelectronic-centered smart product used in e-mobility solutions validates the approach.

Keywords: 9R Strategies; Product Engineering; Product Lifecycle Management; Reconfiguration; Smart Products; Sustainability Strategies

1 INTRODUCTION

Sustainability has become a key concern in many industries, including product development [1]. Companies are increasingly aware of the importance of incorporating sustainability principles into their product development processes to minimize environmental impact and meet evolving customer expectations [2]. Products and how they are used have changed dramatically in a short time. This includes the sharp and fast rise in new technologies as well as a change in consumer behavior, which is reflected in greater individualization of products, greater demand for quality and an expansion of product functions [3]. Traditional product life cycles are changing, meaning premature product obsolescence occurs in addition to shortening useful product lifetime. As a result, products are pushed into the end-of-life phase early and often unnecessarily. In recent years, there has been a growing emphasis on developing smart products and product service systems that not only provide innovative solutions but also address sustainability concerns [4]. The development of smart products and product service systems presents new challenges in terms of sustainability [5]. These challenges arise from the increasing complexity of smart products, which integrate physical and digital components [6], as well as from the need for circular economy (CE) strategies that extend the lifecycle of products [7]. The attribute of reconfiguring smart products and product service systems in the context of the CE is a topic of growing importance [2]. To address these challenges, companies need to adopt new business models and strategies that enable the reconfiguration of smart products and product service systems in a sustainable manner. One approach is to shift from a traditional ownership model to a product-service system model [8, 9], where customers pay for the service or function provided by the product rather than owning the product itself. This model encourages longer product lifecycles and facilitates reconfiguration and upgrades to extend the usefulness of the product. Another crucial aspect of reconfiguring smart products in a CE is designing for reusability and upgradability [10]. This involves incorporating modular design principles, standardized components, and easy disassembly to facilitate component reuse and replacement [11]. Furthermore, designing products for upgradability allows for easy integration of new technological advancements, extending the product's lifespan and reducing electronic waste. By considering these factors during the product development process, companies can create smart products that are not only innovative and functional, but also sustainable. Approaches and procedures for implementing CE strategies with reconfiguration in SP have not yet been sufficiently described. Therefore, This paper aims to open opportunities and challenges for reconfiguring smart products with the inclusion of sustainable strategies at various points in the product life cycle.

2 STATE OF RESEARCH AND TECHNOLOGY

2.1 Circular Economy, Sustainability Strategies and Circularity

Sustainability attempts to form a balanced approach by shifting the dimensions of economy, ecology and society. The aim is to create value without restricting future generations. To incorporate sustainability and channel it into perspectives, goals and phases, the CE model is a suitable system. CE is a model that aims to derive business opportunities from leveraging circularity. However, CE strategies are challenging to define, as research has no uniform definition. There is also the opinion that more than one definition is achievable [12]. CE strategies can simultaneously improve multiple dimensions of sustainability by integrating different stakeholders' needs and creating value through resource conservation. Several approaches exist to categorize these strategies [13]. One such approach divides them into perspectives of strategies for slowing down, narrowing or closing the material loops [14]. Other definitions of CE strategies have a broader perspective on outcomes and focus on processes and outputs that address ecosystem functions and human wellbeing [15]. CE is not an end in itself [16], and it is not necessarily sustainable [17].

To draw a more general conclusion from the CE strategies and objectives, the following assumptions are relevant to this paper. Circular strategies need new business models that support their commercialization [18, 19]. In addition, lifecycle approaches need to be implemented that shift business models to a service-oriented perspective [14, 18]. Sustainability research results in new requirements (or stricter versions of existing requirements) that must be implemented in engineering. Design strategies such as design for long-life products, design for product-life extension or design for a technological cycle must be implemented and strengthened [14]. On the other hand, a business model must be linked to this to be able to commercialize the additional benefits. Designers and engineers must define new attributes that enable products to fulfill circular functions to meet these strategies [20]. However, established CE models still need to integrate established engineering methodologies and guidelines [21]. CE methods are mainly focused on material loops and their circularity in production.

2.2 Smart Products and Reconfiguration

Smart products represent a profound product design and functionality evolution, including technological advances in connectivity, intelligence and autonomy [6]. These products, characterized by their ability to collect, analyze and utilize data, have revolutionized various industries by offering personalized services, enhancing user experiences, and optimizing their performance. A key feature of smart products is their ability to continuously adapt and improve, using their embedded sensors and data analytics capabilities to gain real-time insights into user behavior, environmental conditions and usage patterns [2, 22]. This phase represents a dynamic period where the product interacts with its environment, adapting and responding to evolving circumstances. Smart products excel in optimizing resource efficiency, minimizing waste, and prolonging functional lifespan, aligning with the CE's principles. Reconfiguration emerges as a critical aspect of smart products, enabling them to adapt to changing requirements and preferences [23]. Defined as modifying existing product configurations to fulfill new needs, reconfiguration can manifest in various forms, ranging from general technical improvements to individual runtime enhancements [24]. This so-called smart reconfiguration combines the capabilities of smart products with the flexibility to adjust and evolve over time [25]. The resulting versatility enables smart products to remain relevant and effective in dynamic environments, catering to diverse user demands and preferences.

When developing smart products, everyone involved, e.g., development engineers, must think about the context of how products will be used throughout their life cycle [26]. This includes planning for potential future requirements starting from the early development phase to ensure the product can react to changing conditions during its use phase. In addition to the technical perspective, economic considerations also play a role in being able to change the business model in a resilient manner. These considerations must be anchored and considered in the development process through methods and guidelines. The factors can be summarized with the term "VUCA world", an acronym introduced in the 1990s by the United States Army which describes the possible challenges in a multilateral world through the terms volatility, uncertainty, complexity and ambiguity. Through reconfiguration, it is possible to react to the VUCA factors later in the product life cycle, adapt products and business models and thus enable products or product life cycle for resilience. The large number and diversity of goals for reconfiguring smart products require a more detailed consideration. The following goals are most frequently mentioned in the literature:

- Adaptability and expandability: Smart products must react to changes in the environment as well as to changing user requirements. Part of this goal can be changes to existing features or the enhancement of functions.
- Customization and personalized services: Smart products can analyze user behavior. On this basis, these products can be customized to individual people. As a result, individualized and personalized smart products can have a higher value for the user.
- Longevity and efficiency: The adaptability of smart products means that, on the one hand, their useful life can be extended or made more efficient and, on the other hand, products can be used for longer, which corresponds to an increase in their service life.

Several aspects have to be addressed during product development for dynamically reconfigurable smart products. Procedures and methods need to incorporate and focus on these properties. However, no established approaches currently cover, for example, design for reconfiguration for dynamic reconfiguration in the use phase [6].

3 SUPPORTING CIRCULAR ECONOMY STRATEGIES BY RECONFIGURATION

The 9R strategies arise from the perspective of the CE business model. R strategies are needed to identify value retention across life cycle phases and actors [20]. These strategies can be subdivided as the objectives can be projected to different product levels or product life cycle phases [27]. The 9R strategies range from R0 Refuse, Rethink, Reduce, Reuse, Repair, Refurbish, Remanufacture, Repurpose, Recycle to R9 Recover [28, 29]. Strategies R0 to R2 aim to minimize or reduce the use of raw materials before the economic cycle or production begins [30]. This can be achieved by removing the need for specific products or components, as their functions are fulfilled elsewhere. In addition, raw material consumption is reduced by increasing production efficiency or intensifying product use. Consequently, an equivalent customer benefit can be provided with fewer raw materials.

Strategies from R3 to R7 focus on keeping product parts in the economic cycle [30]. If the reuse of product components is no longer possible, strategies R8 and R9 attempt to at least keep the raw materials in the cycle or minimize waste. When introducing CE principles during the life cycle of an existing product during its utilization phase, strategies R3 through R7 offer the most significant potential for improvement. Additionally, product reconfiguration can offer benefits such as added value at the functional level. R strategies that can be applied at this level are R0 to R2. For this reason, reconfiguration should offer the opportunity to address R-strategies before the business cycle by changing the product in the usage phase. In this way, additional functionalities that would otherwise have to be provided by other products and intensification of product use can keep the product in the economic cycle or the use phase for longer.

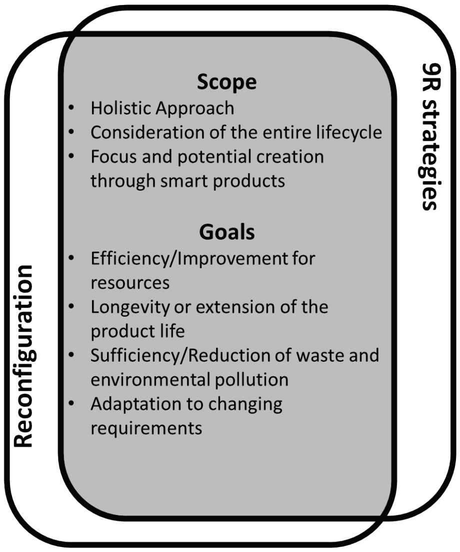


Figure 1 Common scope and goals of 9R strategies and circular product reconfiguration

Suppose the target dimensions of product reconfiguration and the 9R strategies are analyzed for generic overlaps in engineering tasks. In that case, similarities can be identified (Fig. 1). Both the reconfiguration of smart products and the 9R strategies aim to improve energy efficiency. Reconfiguration can help optimize energy consumption by deactivating functions when they are not needed or using more energy-efficient components. Both approaches aim to reduce the amount of waste and environmental impact. Reconfiguration can help extend product life and thus reduce the need for new production and disposal. The reconfiguration of smart products and the 9R strategies aim to optimize the use of resources, be it through the reuse of materials or the more efficient use of energy and other resources. Both approaches address the need to adapt products to changing requirements, whether by adapting functions, adding new features or changing design elements. In summary, the concept of reconfiguration, which also encompasses the objectives of the 9R strategies, can also be described as circular reconfiguration. Overall, there are significant similarities and overlaps between the goals of reconfiguring smart products and the 9R strategies, particularly regarding resource efficiency, waste reduction and adaptability. However, there are also differences in focus and approaches, particularly regarding the product life cycle and material use. While the 9R strategies strongly focus on the reduction, reuse and recycling of materials, the reconfiguration of smart products concentrates more on using existing resources and extending the service life of products. The 9R strategies often refer to early phases in the implementation and the emergence of the technical implementation to activate R3 to R7 in the use phase, for example. Circular reconfiguration can precisely solve this problem by implementing the potential of refurbishment or repair, among other things, in the utilization phase.

Existing approaches either do not consider reconfiguration as a potential for 9R strategies or only marginally [14, 20, 31-33], are mainly concerned with production and the reconfiguration of manufacturing machines, which is not generally transferable to smart products [34-36], or only consider the 9R strategies (R8 and R) for materialities [13, 21, 37, 38].

4 CIRCULAR RECONFIGURATION DESIGN DIMENSIONS

A smart product goes through various phases of the product life cycle. The segmentation is important because it examines the effects of and through the reconfiguration. Fig. 2 shows examples of different stages of the product life cycle. The approach of a circular reconfiguration is explicitly highlighted.

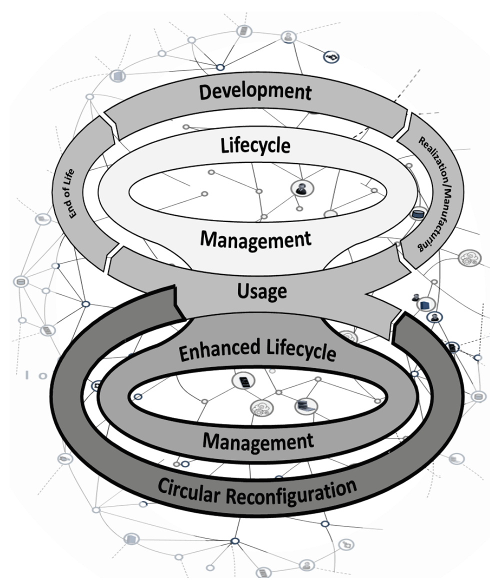


Figure 2 Smart product lifecycle with a focus on circular reconfiguration and enhanced lifecycle management

From a technical perspective, these can be divided into the following phases: strategic planning, development, realization/production, use, end of life cycle. This categorization is very general and can also be subdivided differently. Each phase has characteristics that are either reflected in the smart product or must be considered during its development. To be able to focus on circular reconfiguration, an extension of the life cycle phases is recommended. Either the utilization phase can be extended to include reconfiguration or, as in this case, a separate section can be created. The circular reconfiguration process is initiated in the utilization phase. In the next step, the actual circular reconfiguration process is carried out after

TECHNICAL JOURNAL 18, 3(2024), 452-460

triggering. This includes a separate product development process that covers the requirements and functions right down to the physical parts and sub-products. In the next step, the existing product is reconfigured digitally and/or physically, depending on the type of reconfiguration. This process can be referred to as the remanufacturing phase. After complete and successful reconfiguration, the product is back in the utilization phase. Each of these phases brings with it different challenges, requirements, and opportunities from a technical perspective in connection with reconfiguration. Business models that already consider the economic benefits of reconfiguration must be included as early as the strategic planning stage. The circular reconfiguration of smart products only makes sense for companies and organizations if added value can be created at an economic level. In the following, eight different design dimensions are discussed, which significantly influence circular reconfiguration.

4.1 Continuous vs a Priori Reconfiguration Option Development and Validation

The development and validation of reconfiguration options can be classified as continuous or a priory. Both possibilities have opportunities but also challenges that need to be considered. The a priori reconfiguration option development and validation enables operators to plan the product's behavior in advance, i.e. during the use phase and to take foreseeable conditions into account [10]. It allows precise adjustments to be made to the product right from the start to ensure optimum performance. This influence can increase efficiency through better utilization of resources and smooth (functional) product changes without constantly relying on and reacting to real-time data. Costs can also be better controlled through a priori reconfiguration development and validation, as development and implementation can be planned in advance. Potentially costly changes during operation can thus be avoided. Disadvantages include a lack of flexibility and limited adaptability during the utilization phase. A priori reconfigurations are fixed and offer limited flexibility to adapt to unforeseen or changing conditions. This can lead to sub-optimization if the actual conditions do not match the predicted ones.

Continuous reconfiguration allows product behavior to adapt to changing conditions in real-time, enabling optimal performance even under variable or unforeseen circumstances. The user experience is improved as the product responds dynamically to user needs and preferences. Continuous reconfiguration opens opportunities for ongoing innovation as the product can be continuously improved and optimized based on real-time data and feedback. However, the disadvantages include complexity, data protection and security, and resource requirements [39]. Implementing continuous reconfiguration requires complex algorithms and systems for real-time data processing and customization, which can make the development and maintenance of the product more complicated. It also raises privacy and security concerns, especially when personal or sensitive information is involved. Real-time data processing can also be expected to increase operating costs, energy and computing power.

Overall, both a priori and continuous reconfiguration offer different opportunities to optimize the performance and user experience of smart products. Still, they also come with various challenges that must be carefully considered.

4.2 Engineering Domain Focus

Various disciplines are essential for the reconfiguration of smart products. These can be divided into the development phase and the utilization phase. In the development phase, all disciplines that contribute to the smart product creation are usually involved. This means that the smart product's adaptability must be considered in electrical/electronic and information technology as well as mechanical engineering. Depending on the reconfiguration, these disciplines are involved to varying degrees. The more interdisciplinary the development of reconfiguration options is, the more complex and cost-intensive it becomes [40]. Conversely, reconfiguration options that only require a change or adaptation of the mechanics, for example, can be realized much more easily, quickly and cost-effectively.

4.3 Need-Oriented Triggers for Reconfiguration Options

As mentioned in the description of reconfiguration, the drivers of a reconfiguration can be different. On the one hand, as already described in the VUCA world, external circumstances can change user behavior, user requirements and other environmental conditions. Changing conditions include technological innovations that require rapid adaptation to prevent functional or qualitative obsolescence [41]. New business areas are also possible by seamlessly integrating existing or extended functions into new areas. New functions can be validated and optimized with smart products already on the market. On the user side, new requirements, e.g. for sustainability, or a change or adaptation in use or intensification, which is possible through data analysis of smart products, can be implemented quickly through reconfiguration to counteract obsolescence [42]. In the context of sustainability, smart products can be adapted to extended CE strategies, such as longevity. Technical triggers initiated by the product include faults and malfunctions that can be eliminated through reconfiguration, as well as adapting, improving and optimizing the performance of a function or product.

4.4 Barriers to Introducing Reconfiguration Options

Various barriers to reconfiguration hinder successful implementation or rule out reconfiguration options as early as during the development stage. The information technological foundations for future reconfiguration (including a coherent digital twin and its connection to digital product models) must be considered at the beginning of the development phase. Neglecting this can lead to challenges when accessing the data needed for reconfiguration, such as variant models or field data from the product’s utilization phase. Developing reconfiguration options requires a deep understanding of the product architecture, including hardware, software and interfaces. The complexity of the product architecture can make the development and implementation of reconfiguration options more difficult. Data access and availability also play a crucial role in developing reconfiguration options. However, access to relevant data sources can be challenging, especially if data is proprietary or comes from third-party providers. Integrating reconfiguration options into existing systems or products can be complex and requires careful planning and coordination. Interoperability with other systems and compliance with standards can present additional challenges. User acceptance of proposed reconfiguration options can be challenging, especially if they involve changes to the user experience or workflows. Furthermore, developing reconfiguration options requires time, money and expertise. To overcome barriers, it is essential to develop a comprehensive plan and strategy for the development of reconfiguration options that take into account the above challenges and include appropriate measures to overcome them. This may require collaboration between disciplines, stakeholders and partners to ensure successful development and implementation.

4.5 Ability to Model-Based Mapping and Tracking of New/Extended Product Reconfiguration

Smart products are usually developed based on models [6]. This makes it much easier to trace individual partial models or functions at a later date. Clear and comprehensive modeling of the product architecture, including all components, interfaces and dependencies, is essential. This includes documentation of the current configuration and potential reconfiguration options. A robust tracking system makes it possible to track changes to the product configuration, including new functions, components or interfaces. Using a version control system is essential to track changes to the product models and ensure that different versions of the product configuration are appropriately documented and managed. Automating modeling, mapping and tracking processes can improve efficiency and reduce human error [43]. This includes the use of tools and software to automate repetitive tasks. The capability for model-based mapping and tracking should integrate seamlessly with other development and operational systems to ensure a smooth exchange of information. The model-based reconfiguration tracking system should be flexible and adaptable to support different types of products and reconfigurations, regardless of their complexity or scope. Getting real-time feedback on the status and impact of reconfigurations is essential. This makes it possible to react quickly to problems and make decisions based on up-to-date information. By taking these aspects into account, the ability to model-based mapping and tracking of new and extended product reconfigurations can help to improve efficiency, transparency and control over the reconfiguration process and thus increase the quality and reliability of intelligent products.

4.6 Ability to Assess the Sustainability Contribution of the Reconfiguration Options

Various dimensions must be considered to evaluate reconfiguration options in the context of sustainability. On the one hand, the three sustainability dimensions of ecology, economy and social issues must be integrated. On the other hand, feasibility and implementation must be considered from a technical and economic perspective [44]. Reconfiguration can lead to changes in product components that have a different environmental impact than initially assumed. It is therefore essential to assess whether reconfiguration options can contribute to the resource efficiency or service life extension, for example. Assessments must be based on a life cycle analysis that can estimate features other than material consumption, such as functions. From a CE perspective, there currently need to be suitable tools to guarantee this assessment [45] fully. Nevertheless, various indicators from other areas can be used at least in part, e.g., indicators for longevity of a product [32].

4.7 Ability to Assess the Conflict of Objectives with Traditional Product Development Objectives

When developing smart products to make products reconfigurable, various conflicts of interest can arise with traditional product development. Traditional product development often places great value on the stability and reliability of a product. However, introducing reconfiguration options can increase complexity and compromise the integrity of the product. There is therefore a trade-off between the need to provide a stable product and the flexibility to reconfigure the product as required. However, by striving for continuity and consistency in terms of design, functions and processes, existing and traditional approaches take time and effort to overcome the complex integration required by smart products. However, introducing new reconfiguration options requires innovation and change, which can potentially lead to conflicting goals, especially if users or stakeholders expect a certain degree of continuity. Users often have certain expectations of their products, including ease of use, performance and reliability. Introducing reconfiguration options may not meet these expectations or may be technically challenging, leading to trade-offs between user expectations and technical feasibility [10]. Traditional product development often strives to design and manufacture products cost-effectively while providing adequate benefits to users. However, introducing reconfiguration options can incur additional costs, whether in terms of development, implementation or maintenance, leading to trade-offs between costs and benefits. It is essential to recognize and carefully consider these trade-offs to enable a balanced and successful development of reconfigurable smart products. This requires a thorough analysis, consideration and integration of different requirements, priorities and interests of the parties involved.

4.8 Reconfiguration Option Requirements

To reconfigure a smart product in the usage phase, various key elements are required that are decisive for implementation, depending on the degree of maturity or availability. The product must be designed so that it has a modular structure and different components can be reconfigured independently of each other. This enables flexible adaptation to changing requirements and conditions

[46]. In addition, straightforward interfaces and standards for the communication and integration of different components and systems are required, as well as interoperability with other systems and products and support for standards and protocols to enable seamless integration and communication, facilitating reconfiguration and replacement of components without significant problems. The product should be able to monitor its condition and performance and provide diagnostic data continuously. This makes it possible to detect problems early and identify the need for reconfiguration. The data collection must include robust security mechanisms to prevent unauthorized access and tampering.

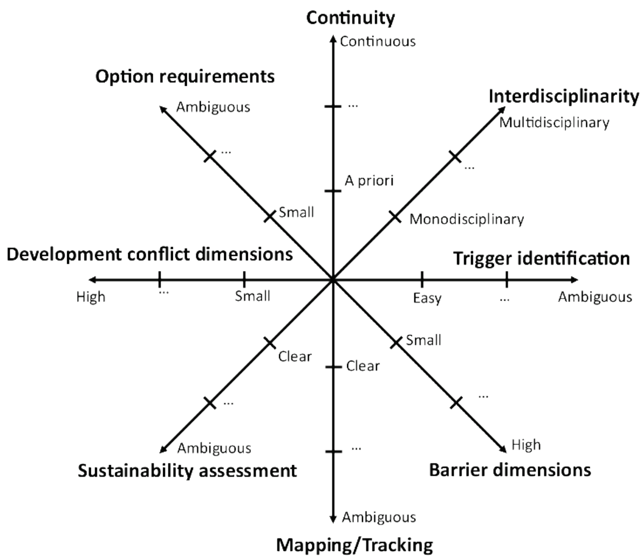


Figure 3 Indicator space for influence on Design Dimensions for reconfiguration options

This is particularly important if the product processes sensitive data or performs critical functions [47]. Additionally, data collection that can be connected to the product's user must comply with the regional data protection laws. To create a highly integrated digital twin of the product, the collected data must be connected to existing product models, especially the configuration and variant models to support reconfiguration and other service processes. Finally, the reconfiguration design dimensions presented can be brought together graphically in the dependency matrix (Fig. 3). By classifying existing products in the indicator matrix, the effort, opportunities, and risks involved in successfully introducing reconfiguration options can be identified. Depending on the degree of maturity and availability or difficulty in the dimension, an effort can be estimated.

5 IDENTIFYING DRIVERS AND GOALS USING THE EXAMPLE OF AN INDUSTRIAL USE CASE

The research consortium, which consisted of 35 partners from industry and academia, came together in 2021 and were motivated by various drivers from the VUCA world. The aim is to establish a European value creation network to manufacture products from the semiconductor industry using new technologies at substrate, device, and module levels. Enormous leaps in innovation have been made in the field of semiconductor development in recent years. The volatility in the demand for semiconductor products can change rapidly, be it due to market trends, technological developments or geopolitical events. This can lead to unpredictable fluctuations in the market, forcing companies to react flexibly to meet demand or adapt to changing market conditions. Similarly, price fluctuations are a characteristic industry feature, often due to supply and demand, competitive pressures and currency fluctuations. Uncertainty in the semiconductor industry relates primarily to technological developments and regulatory changes. Technological trends and innovations are difficult to predict and can influence investment decisions, research and development strategies and product planning. At the same time, unclear regulatory requirements and provisions lead to uncertainty about international trade agreements, data protection regulations and security standards. The complexity of the semiconductor industry is reflected in its sophisticated manufacturing processes and global supply chains. The manufacture of semiconductors requires complex production equipment and methods, which can lead to challenges in terms of quality assurance, productivity optimization and cost management. In addition, the industry's global supply chains are highly complex, making sourcing, logistics and risk management challenging. The ambiguity in the semiconductor industry is reflected in market uncertainty and technological convergence. Market uncertainties such as economic cycles, geopolitical tensions and trade conflicts influence demand, prices and investment decisions in the industry. At the same time, the convergence of different technologies leads to uncertainties about their effects and possible applications for the semiconductor industry. Overall, these challenges show how the semiconductor industry operates in a VUCA world and how companies in the industry must deal with volatile, uncertain, complex and ambiguous conditions to succeed.

The aim of meeting these challenges is to create a purely European value chain that can quickly adapt to market conditions and requirements. The product examples developed in the project aim to integrate new innovations in the field of semiconductor technology, especially about manufacturing processes, into existing products and to be able to offer new functionalities, e.g. bidirectional charging of cars makes it possible to dynamically provide electric car batteries for home use or even grid requirements [48, 49]. This functionality can be created and integrated by reconfiguring and replacing the main inverter on the vehicle side. Another goal is to increase the efficiency of the metal-oxide-semiconductor field-effect transistor (MOSFET) by introducing a modified substrate in production. However, silicon-based semiconductor materials are, in some cases, stretched to their physical limits. For this reason, on the material level, silicon carbide (SiC) has now become the focus of research, as it is not only more efficient but also more minor in design and generates less heat. By changing the material technology, goals such as more prolonged and more intensive product use can be achieved in the context of circular reconfiguration. Due to the VUCA world, the consortium is building a more consistent value chain in Europe to realize more efficient and sufficient product components, which can also be used excellently for circular reconfiguration.

6 FUTURE AREAS OF DEMAND

In summary, the two topics of reconfiguring smart products and sustainability strategies were brought together. Overlaps in scope and objectives were identified, which call for more efficiency, sufficiency, and consistency at the product level but also reveal concrete implications for developing these products. There needs to be more than existing approaches to close the research gap identified. Either there is too rigid a focus on reconfiguration in the context of goals that do not necessarily correspond to sustainability or sustainability strategies have been described too much on a material level that cannot be solved on a functional level or other overlaps with reconfiguration. The design dimensions described show current opportunities and challenges that must be solved at a specific level. Promoting the circular reconfiguration of smart products requires a holistic approach to development processes, tools, and methods. A key factor is a design for reconfigurability, which aims to design products from the outset in such a way that they can be easily disassembled, modified, and reused. This requires the integration of modular design principles and standardized interfaces to facilitate the interchangeability and upgradeability of components. In addition, development processes must consider the product's entire life cycle and factors beyond for extended, modified or second life cycles. Digital tools and technologies like the Internet of Things, big data analytics and artificial intelligence can support product reconfigurability by enabling real-time monitoring, diagnostics and customization throughout the lifecycle. Development methods should aim to maximize resource efficiency and minimize waste by integrating and implementing principles such as 9R strategies. Partnerships and collaboration between different stakeholders are crucial to address the complexity of these challenges and develop joint solutions. Overall, promoting the circular reconfiguration of smart products requires an integrative approach focusing on design innovation, life cycle thinking, digitalization, modular architectures, and CE while fostering partnerships and collaboration between different actors.

7 REFERENCES

[1] Ahmad, S., Wong, K. Y., Tseng, M. L. & Wong, W. P. (2018). Sustainable product design and development: A review of tools, applications and research prospects. *Resources, Conservation and Recycling, 132*, 49-61. https://doi.org/10.1016/j.resconrec.2018.01.020

[2] Li, X., Wang, Z., Chen, C.-H. & Zheng, P. (2021). A data-driven reversible framework for achieving Sustainable Smart product-service systems. *Journal of Cleaner Production, 279*, 123618. https://doi.org/10.1016/j.jclepro.2020.123618

[3] Krause, D. & Heyden, E., Editors (2022). *Design Methodology for Future Products*. Springer International Publishing. https://doi.org/10.1007/978-3-030-78368-6

[4] Sorli, M., Sopelana, A., Salgado, M., Peláez, G. & Ares, E. (2012). Balance between Lean and Sustainability in Product Development. *Key Engineering Materials, 502*, 37-42. https://doi.org/10.4028/www.scientific.net/kem.502.37

[5] Luttropp, C. & Lagerstedt, J. (2006). EcoDesign and The Ten Golden Rules: generic advice for merging environmental aspects into product development. *Journal of Cleaner Production, 14*(15), 1396-1408. https://doi.org/10.1016/j.jclepro.2005.11.022

[6] Tomiyama, T., Lutters, E., Stark, R. & Abramovici, M. (2019). Development capabilities for smart products. *CIRP Annals, 68*(2), 727-750. https://doi.org/10.1016/j.cirp.2019.05.010

[7] Ramani, K., Ramanujan, D., Bernstein, W. Z., Zhao, F., Sutherland, J., Handwerker, C., Choi, J. K., Kim, H. & Thurston, D. (2010). Integrated sustainable life cycle design: A Review. *Journal of Mechanical Design, 132*(9), 910041-9100415. https://doi.org/10.1115/1.4002308

[8] Alcayaga, A., Hansen, E., 2017. Smart-circular systems: a service business model perspective. *PLATE 2017 - Product Lifetimes and the Environment*, Delft, NL. 10-13. https://doi.org/10.3233/978-1-61499-820-4-10

[9] Zheng, P., Lin, T.-J., Chen, C.-H., Xu, X., 2018. A systematic design approach for service innovation of smart product-service systems. *Journal of Cleaner Production, 201*, 657-667. https://doi.org/10.1016/j.jclepro.2018.08.101

[10] Savarino, P., Abramovici, M., Göbel, J. C. & Gebus, P. (2018). Design for reconfiguration as a fundamental aspect of smart products. *Procedia CIRP, 70*, 374-379. https://doi.org/10.1016/j.procir.2018.01.007

[11] Inoue, M., Yamada, S., Miyajima, S., Ishii, K. et al. (2020). A modular design strategy considering sustainability and supplier selection. *Journal of Advanced Mechanical Design, Systems, and Manufacturing, 14*(2). https://doi.org/10.1299/jamdsm.2020jamdsm0023

[12] Korhonen, J., Nuur, C., Feldmann, A. & Birkie, S. E. (2018). Circular economy as an essentially contested concept. *Journal of Cleaner Production, 175*, 544-552. https://doi.org/10.1016/j.jclepro.2017.12.111

[13] Moraga, G., Huysveld, S., Mathieux, F., Blengini, G. A. et al., (2019). Circular economy indicators: What do they measure? *Resources, Conservation and Recycling, 146*, 452-461. https://doi.org/10.1016/j.resconrec.2019.03.045

[14] Bocken, N. M. P., de Pauw, I., Bakker, C. & van der Grinten, B. (2016). Product design and business model strategies for a circular economy. *Journal of Industrial and Production Engineering, 33*(5), 308-320. https://doi.org/10.1080/21681015.2016.1172124

[15] Murray, A., Skene, K. & Haynes, K. (2017). The Circular Economy: An Interdisciplinary Exploration of the Concept and Application in a Global Context. *J Bus Ethics, 140*, 369-380. https://doi.org/10.1007/s10551-015-2693-2

[16] acatech, Circular Economy Initiative Deutschland, SYSTEMIQ, 2021. Circular Economy Roadmap für Deutschland. acatech, Deutsche Akademie der Technikwissenschaften e.V, München. https://www.acatech.de/publikation/circular-economy-roadmap-fuer-deutschland/

[17] Zink, T. & Geyer, R. (2017). Circular Economy Rebound. *Journal of Industrial Ecology, 21*(3), 593-602. https://doi.org/10.1111/jiec.12545

[18] Alcayaga, A., Wiener, M. & Hansen, E. G. (2019). Towards a framework of smart-circular systems: An integrative literature review. *Journal of Cleaner Production, 221*, 622-634. https://doi.org/10.1016/j.jclepro.2019.02.085

[19] Michelini, G., Moraes, R. N., Cunha, R. N., Costa, J. M. et al. (2017). From Linear to Circular Economy: PSS Conducting the Transition. *Procedia CIRP, 64*, 2-6. https://doi.org/10.1016/j.procir.2017.03.012
[20] Diaz Tena, A., Schoeggl, J.-P., Reyes, T. & Baumgartner, R. J. (2021). Exploring sustainable product development processes for a circular economy through morphological analysis. *Proceedings of the Design Society, 1*, 1491-1500. https://doi.org/10.1017/pds.2021.410
[21] Gräßler, I. & Pottebaum, J. (2021). Generic Product Lifecycle Model: A Holistic and Adaptable Approach for Multi-Disciplinary Product–Service Systems. *Appl. Sci., 11*(10), 4516. https://doi.org/10.3390/app11104516
[22] Abramovici, M., Göbel, J. C. & Savarino, P. (2017). Reconfiguration of smart products during their use phase based on virtual product twins. *CIRP Annals, 66*(1), 165-168. https://doi.org/10.1016/j.cirp.2017.04.042
[23] Abramovici, M., Savarino, P., Göbel, J. C. & Adwernat, S. (2018). Systematization of Virtual Product Twin Models in the Context of Smart Product Reconfiguration during the Product Use Phase. *Procedia CIRP, 69*, 734-739. https://doi.org/10.1016/j.procir.2017.11.025
[24] Li, X., Chen, C.-H., Zheng, P., Jiang, Z. & Wang, L. (2021). A context-aware diversity-oriented knowledge recommendation approach for smart engineering solution design. *Knowledge-Based Systems, 215*, 106739. https://doi.org/10.1016/j.knosys.2021.106739
[25] Savarino, P., Abramovici, M. & Göbel, J. C. (2018). A Methodological Approach for the Identification of Context-Specific Reconfiguration Options in the PLM-Context. In: Chiabert, P., Bouras, A., Noël, F. & Ríos, J. (eds) Product Lifecycle Management to Support Industry 4.0. *PLM 2018. IFIP Advances in Information and Communication Technology, vol 540*. Springer, Cham. https://doi.org/10.1007/978-3-030-01614-2_36
[26] Gräßler, I. & Pottebaum, J. (2022). From Agile Strategic Foresight to Sustainable Mechatronic and Cyber-Physical Systems in Circular Economies. In: Krause, D. & Heyden, E. (eds) *Design Methodology for Future Products*. Springer, Cham. https://doi.org/10.1007/978-3-030-78368-6_1
[27] Kara, S., Hauschild, M., Sutherland, J. & McAloone, T. (2022). Closed-Loop Systems to Circular Economy: A Pathway to Environmental Sustainability? *CIRP Annals, 71*(2), 505-528. https://doi.org/10.1016/j.cirp.2022.05.008
[28] Potting, J., Hekkert, M., Worrell, E. & Hanemaaijer, A. (2017). Circular Economy: Measuring Innovation in the Product Chain. The Hague.
[29] Kirchherr, J., Reike, D. & Hekkert, M. (2017). Conceptualizing the circular economy: An analysis of 114 definitions. *Resources, Conservation and Recycling, 127*, 221-232. https://doi.org/10.1016/j.resconrec.2017.09.005
[30] Mast, J., von Unruh, F. & Irrek, W. (2022). R-strategies as guidelines for the Circular Economy. https://prosperkolleg.ruhr/wp-content/uploads/2022/08/rethink_22-03_r-strategien_EN.pdf
[31] Halstenberg, F., Dönmez, J., Mennenga, M., Herrmann, C. & Stark, R. (2021). Knowledge transfer and engineering methods for smart-circular product service systems. *Procedia CIRP, 100*, 379- 384. https://doi.org/10.1016/j.procir.2021.05.088
[32] Franklin-Johnson, E., Figge, F. & Canning, L. (2016). Resource duration as a managerial indicator for Circular Economy performance. *Journal of Cleaner Production, 133*, 589-598. https://doi.org/10.1016/j.jclepro.2016.05.023
[33] Hauschild, M. Z., Jeswiet, J. & Alting, L. (2004). Design for Environment — Do We Get the Focus Right? *CIRP Annals, 53*(1), 1-4. https://doi.org/10.1016/S0007-8506(07)60631-3
[34] Bortolini, M., Accorsi, R., Faccio, M., Galizia, F. G. & Pilati, F. (2019). Toward a Real-Time Reconfiguration of Self-Adaptive Smart Assembly Systems. *Procedia Manufacturing, 39*, 90-97. https://doi.org/10.1016/j.promfg.2020.01.232
[35] Brad, S. & Murar, M. (2015). Employing Smart Units and Servitization towards Reconfigurability of Manufacturing Processes. *Procedia CIRP, 30*, 498-503. https://doi.org/10.1016/j.procir.2015.02.154
[36] Dahmani, A., Benyoucef, L. & Mercantini, J.-M. (2022). Toward Sustainable Reconfigurable Manufacturing Systems (SRMS): Past, Present, and Future. *Procedia Computer Science, 200*, 1605-1614. https://doi.org/10.1016/j.procs.2022.01.361
[37] Benfer, M., Gartner, P., Klenk, F., Wallner, C., Jaspers, M.-C., Peukert, S. & Lanza, G. (2022). A Circular Economy Strategy Selection Approach: Component-based Strategy Assignment using the Example of Electric Motors. In: Herberger, D. & Hübner, M. (Eds.): *Proceedings of the Conference on Production Systems and Logistics: CPSL 2022*. Hannover: publish-Ing., 22-31. https://doi.org/10.15488/12133
[38] Gräßler, I. & Hesse, P. (2023). Considering engineering activities and product characteristics to achieve material circularity by design. *Proceedings of the Design Society, 3*, 1077-1086. https://doi.org/10.1017/pds.2023.108
[39] Kuhlenkötter, B., Wilkens, U., Bender, B., Abramovici, M., Süße, T., Göbel, J., Herzog, M., Hypki, A. & Lenkenhoff, K. (2017). New Perspectives for Generating Smart PSS Solutions – Life Cycle, Methodologies and Transformation. *Procedia CIRP, 64*, 217-222. https://doi.org/10.1016/j.procir.2017.03.036
[40] Persson, J.-G. (2016). Current Trends in Product Development. *Procedia CIRP, 50*, 378-383. https://doi.org/10.1016/j.procir.2016.05.088
[41] Taskan, B., Junça-Silva, A. & Caetano, A. (2022). Clarifying the conceptual map of VUCA: a systematic review. *International Journal of Organizational Analysis, 30*(7), 196-217. https://doi.org/10.1108/IJOA-02-2022-3136
[42] Kristoffersen, E., Blomsma, F., Mikalef, P. & Li, J. (2020). The smart circular economy: A digital-enabled circular strategies framework for manufacturing companies. *Journal of Business Research, 120*, 241-261. https://doi.org/10.1016/j.jbusres.2020.07.044
[43] Zheng, P., Chen, C.-H. & Shang, S. (2019). Towards an automatic engineering change management in smart product-service systems – A DSM-based learning approach. *Advanced Engineering Informatics, 39*, 203-213. https://doi.org/10.1016/j.aei.2019.01.002
[44] Molina, A., Ponce, P., Miranda, J. & Cortés, D. (2021). *Enabling Systems for Intelligent Manufacturing in Industry 4.0 - Sensing, Smart and Sustainable Products (S3 Products)*. Imprint Springer, p. 71. https://doi.org/10.1007/978-3-030-65547-1
[45] Saidani, M., Yannou, B., Leroy, Y., Cluzel, F. & Kendall, A. (2019). A taxonomy of circular economy indicators. *Journal of Cleaner Production, 207*, 542-559. https://doi.org/10.1016/j.jclepro.2018.10.014
[46] Yan, J. & Feng, C. (2014). Sustainable design-oriented product modularity combined with 6R concept: a case study of rotor laboratory bench. *Clean Techn Environ Policy, 16*, 95-109. https://doi.org/10.1007/s10098-013-0597-3
[47] Atlam, H. F. & Wills, G. B. (2020). IoT Security, Privacy, Safety and Ethics. In: Farsi, M., Daneshkhah, A., Hosseinian-

Far, A. & Jahankhani, H. (eds) *Digital Twin Technologies and Smart Cities. Internet of Things*. Springer, Cham. https://doi.org/10.1007/978-3-030-18732-3_8

[48] Fayad, E., Rosas, D. S., Bruyere, A. & Poirier, F. (2023). Automotive Charger Grid-Forming Control Opportunities for G2V and V2X Applications. https://doi.org/10.13140/RG.2.2.31733.96482

[49] Imbruglia, A., Gennaro, F. & di Grazia, P. (2022). The Mobility Scenario vs Green Deal Objectives. https://dev.sic-transform.eu/download/TRANSFORM_Paper%20SSI2022_ST-IT.pdf

Authors' contacts:

Yannick Juresa
(Corresponding author)
RPTU Kaiserslautern-Landau, Institute of Virtual Product Engineering,
Gottlieb-Daimler-Str. 44, 67663 Kaiserslautern, Germany
+49 631 205 2312, yannick.juresa@mv.rptu.de

Damun Mollahassani
RPTU Kaiserslautern-Landau, Institute of Virtual Product Engineering,
Gottlieb-Daimler-Str. 44, 67663 Kaiserslautern, Germany
damun.mollahassani@mv.rptu.de

Jens C. Goebel
RPTU Kaiserslautern-Landau, Institute of Virtual Product Engineering,
Gottlieb-Daimler-Str. 44, 67663 Kaiserslautern, Germany
jens.goebel@mv.rptu.de

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240619083345

Original scientific paper
Received: 2024-06-19, Accepted: 2024-06-24

Investigation of the Influence of NiBSi/NiCrBSi Coatings Applied by Flame Spraying with Simultaneous Fusing on the Substrate Material

Sara Havrlišan*, Štefanija Klarić, Branko Štrbac

Abstract: The aim of this research is to investigate the influence of nickel alloy type from the same group, the parameters of flame spraying, as well as the preparation of the substrate and the heat treatments of the substrate on the microstructure of the coating/substrate system. Due to the possibility of applying nickel alloys in corrective maintenance of tools used on elevated working temperatures, hot work tool steel X38CrMoV5-1 was selected as a substrate material. The investigation of the microstructure of the coating/substrate system was carried out according to the factorial design of experiment, where the input factors were varied on two levels. The factors that were varied are: Ni-based self-fluxing alloys - NiCrBSi and NiBSi; distance of the burner from the workpiece - small (6 mm) and large (20 mm); preparation of the substrate - roughened and non-roughened and the heat treatments of the substrate - soft annealed and tempered condition. Ni-based self-fluxing alloys were applied on samples (12,5 × 25 × 25 mm) by flame spraying with simultaneous fusing process. Analysis of the microstructures of the coating/substrate system was carried out on the Leica DM 2500M light microscope. After the conducted analysis the paper concluded that by spraying the selected coatings onto the X38CrMoV5-1 tool steel base, poor quality coatings are obtained, due to the appearance of cracks (NiCrBSi) or separation of the coating from the substrate (NiBSi). This is attributed to the formation of martensitic structure of the substrate after spraying and the presence of residual stresses.

Keywords: flame spraying with simultaneous fusing; hot work tool steel X38CrMoV5-1; microstructure; Ni-based self-fluxing alloys

1 INTRODUCTION

In order to extend the service life of damaged and worn parts and subsequently aim to reduce maintenance costs, thermal spraying processes can be used to apply different types of coating materials to different substrate materials. Thermal spraying is widely used in almost all branches of industry due to the large number of substrate materials on which coatings can be applied. There is also a large number of coating materials that can be applied, but the most commonly used and researched are coatings based on nickel, iron, and cobalt. Due to their exceptionally good properties, nickel-based coatings are widely used in conditions where good wear resistance, capability to work on elevated temperatures, and corrosion resistance are required. Nickel alloys are suitable for all thermal spraying processes: flame spraying, HVOF, HVAF, a detonation gun, an electric arc, plasma spraying, supersonic plasma spraying, plasma spraying with a transferred arc and laser [1-11], but after the spraying process, a subsequent fusing procedure is often applied, in order to reduce porosity and create a metallurgical connection between the coating and the substrate [1-4].

Due to their wide applicability, flame spraying and nickel-based coatings have been researched in almost all aspects, from resistance to various types of wear and corrosion to researching the influence of external mechanical loads and residual stresses [12-14]. All the aforementioned research was almost always accompanied by microstructure analysis. In almost 90% of microstructure research, the authors analysed mainly the microstructure of the coating. Some authors do not even mention the type of substrate or state it in general (for example, carbon steel or low-carbon steel), which can be understandable when investigating a property that is important only for the coating (i.e., the resistance of the coating to different types of wear). For example, in papers [15-18], where the material of the substrate is mentioned, it is most often low-carbon steel, and the impact of spraying nickel alloys on the material of the substrate itself is not mentioned. Hot work tool steel was chosen as the substrate in a few papers. Hot work tool steels are subjected to different types of wear and thermal fatigue [19, 20], and their protection with different types of coatings has been investigated in order to extend their service life [21-23]. The systematic application of surface engineering procedures for improving the properties of hot work tool steels is presented in the paper [21], in which it is stated that the application of nickel alloys by flame spraying is a possibility. In the paper [23], the results of the investigation of the properties of the NiCrBSi+Ni/MoS2 coating previously applied and then fused using a laser on hot work tool steel substrate (H13, X40CrMoV5-1, Utop Mo2) were presented, but without investigating the effect of the process on the substrate. Difficulties in applying flame spraying for corrective maintenance of molds for high-pressure casting of aluminium castings, which are made of tool steel for working in a hot state, are stated in the paper [24].

Due to the above-mentioned small number of papers investigating the influence of the spraying process, type of coating, spraying parameters, and type of substrate on the microstructure of the coating/substrate system, this paper aims to investigate the effects of the mentioned parameters on the microstructure of the coating/substrate system. The investigation of the microstructure of the coating/substrate system was carried out according to a factorial design of the experiment where four factors were simultaneously changed on two levels: the type of coating (NiBSi and NiCrBSi), the spraying distance during the deposition during flame spraying with simultaneous fusing (small and large), substrate preparation (roughened and non-roughened) and heat treatment of the substrate (soft annealed and tempered condition).

2 EXPERIMENTAL PART

Two types of coatings from the same group of nickel alloys - NiBSi and NiCrBSi (Fig. 1) were applied by flame spraying with simultaneous fusing using Super Jet Euttaloy oxy-acetylene gun (acetylene pressure 50 kPa and oxygen pressure 200 kPa) on samples measuring 12,5 × 25 × 25 mm (substrate - hot work tool steel) [14].

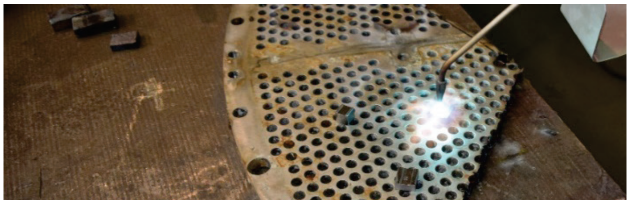


Figure 1 Spraying process using a gas flame with simultaneous fusing

The thickness of the coating is 1 mm and is applied through eight passes. Hot work tool steel X38CrMoV5-1 (Utop Mo1, W300, H11, Č4751, 1.2343) is a high alloy tool steel and one of the most commonly used steels, not only as a hot work tool steel but also as ultra-high strength structural steel. The applied steel has a favourable combination of toughness, hardness, wear resistance, yield resistance with a high limit of elasticity and strength on the elevated temperature and it is air hardening steel [19, 25].

Tab. 1 shows the chemical composition of the substrate material - hot work tool steel X38CrMoV5-1, while Tab. 2 shows the chemical composition of the spraying powder.

Table 1 Chemical composition of the substrate material

Chemical element, % w.t.	C	Si	Mn	P	S	Cr	Ni	Mo	Cu	V	Fe
X38CrMoV5-1	0,36	0,93	0,35	0,018	< 0,001	4,74	0,37	1,00	0,19	0,27	remaining

Table 2 Chemical composition of the spraying powder

Chemical element, % w.t.	C	Cr	Fe	B	Si	Ni
NiCrBSi (*Eutalloy 10009, BoroTec*)	0,7	15	3,5	3,2	4,4	remaining
NiBSi (*Eutalloy 10185, BronzoChrom*)	0,1	0,5	0,5	2,5	3	remaining

The microstructure of the coating/substrate system was investigated according to the factorial design of the experiment, where four factors were changed simultaneously on two levels with two repetitions of combinations of factor levels. The factors changed on two levels are the type of nickel alloy from the same group (type of coating) - NiCrBSi and NiBSi, the spraying distance during the deposition - small (6 mm) and large (20 mm), substrate preparation - roughened and non-roughened and heat treatment of the substrate - soft annealed and tempered condition. Since the spraying process using a gas flame with simultaneous fusing creates a metallurgical connection between the coating and the substrate, the substrate preparation factor is defined on two levels. In addition to the standard preparation of the substrate (level - roughened), as the second level of this factor, it was chosen that this standard preparation will be omitted, and the substrate will be cleaned from impurities and fats only with ethyl alcohol (level - non-roughened). The roughening of the substrate has a significant impact in the so-called cold spraying process using a gas flame, in which a mechanical connection is formed between the coating and the substrate [26]. It can also have a significant impact on the two steps flame spraying process of the investigated alloys - when the alloys are first sprayed with a classic cold process (where the spraying distance from the workpiece during the deposition is up to 220 mm, or even 300 mm [27] and the coating is connected to the substrate by mechanical connection [16, 18, 28], and then subsequent fusing is applied). As the paper investigates the process of spraying using a gas flame with simultaneous fusing (hot process in one step), where the spraying distance from the workpiece is ten and more times less than in the cold process, the coating and the substrate are not connected by mechanical connection, but rather by heating the substrate material to the fusing temperature that allows the diffusion of chemical elements between the coating material in a semi-molten state (between solidus and liquidus temperature) and the substrate material and creation of metallurgical connection. Thus, in research [29, 30], the authors even prepared the substrate to extremely low roughness - mirror finish, with the aim of avoiding sand-sandblasting, which can affect the deterioration of the dynamic properties. The heat treatment of the substrate was chosen as a factor due to the frequent case of applying coatings on heat-treated materials in practice, mainly in corrective maintenance, and due to the observed smaller number of studies related to the impact of the spraying process on the applied substrate, particularly on the pre-heat treated substrate [31, 32]. For this factor (heat treatment), along with the level of tempering of tool steel, the soft annealed state was chosen as the second level. Hot work tool steels are not used in a soft annealed state, but this process allows machining before the tempering. In this paper, however, this condition was chosen as a level, with the aim of proving whether, with the spraying process using a gas flame with simultaneous fusing (the temperature of fusing is similar to the austenitizing temperature of the selected tool steel), a significant change in the structure of the substrate is possible - due to the fact of good quenchability of this type of steel even in case of slow cooling, which is the case after the applied spraying process.

The microstructure of the coating/substrate system was investigated according to the test plan shown in Tab. 3.

2.1 Preparation of Samples for Metallographic Analysis

Fig. 2 shows the dimensions and shape of the samples for testing the microstructure of the coating/substrate (X38CrMoV5-1 steel) system after the spraying process.

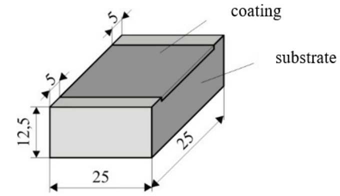


Figure 2 Dimensions and shape of the samples for testing the microstructure of the coating/substrate system

In order to obtain the surface required for metallographic analysis, after the spraying process, the samples were cross-cut on a wire EDM machine (Fig. 3) with intensive water cooling, and then the surface was prepared for metallographic analysis.

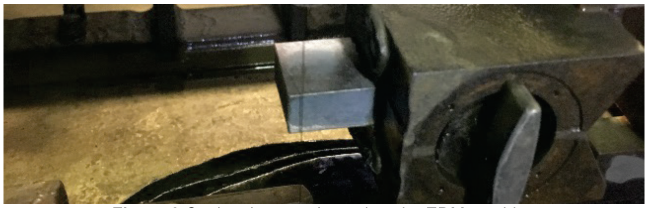


Figure 3 Cutting the sample on the wire EDM machine

All samples were etched with a nital, the aim was to etch the structure of the substrate and not of the coating. Microstructure imaging was carried out on the Leica DM 2500M light microscope according to the test plan shown in Tab. 3, where the Sample column represents the sample labels. For each combination of factor levels, e.g. (substrate material: steel X38CrMoV5-1 - soft annealed; coating type: NiCrBSi; spraying distance: small and substrate preparation: non-roughened) there were two repetitions, which is why the designations are e.g. 1_1-1 and 1_1-2 [14].

The overview of microstructures is given below.

Table 3 Microstructure of coating/substrate microstructure test plan, substrate material - steel X38CrMoV5-1

Substrate material	Coating type	Spraying distance	Substrate preparation	Sample	
Steel X38CrMoV5-1 - soft annealed	NiCrBSi	Small	Non-roughened	1_1-1	1_1-2
	NiCrBSi	Large	Non-roughened	1_2-1	1_2-2
	NiBSi	Small	Non-roughened	1_3-1	1_3-2
	NiBSi	Large	Non-roughened	1_4-1	1_4-2
	NiCrBSi	Small	Roughened	1_1_0-1	1_1_0-2
	NiCrBSi	Large	Roughened	1_2_0-1	1_2_0-2
	NiBSi	Small	Roughened	1_3_0-1	1_3_0-2
	NiBSi	Large	Roughened	1_4_0-1	1_4_0-2
Steel X38CrMoV5-1 - tempered	NiCrBSi	Small	Non-roughened	2_1-1	2_1-2
	NiCrBSi	Large	Non-roughened	2_2-1	2_2-2
	NiBSi	Small	Non-roughened	2_3-1	2_3-2
	NiBSi	Large	Non-roughened	2_4-1	2_4-2
	NiCrBSi	Small	Roughened	2_1_0-1	2_1_0-2
	NiCrBSi	Large	Roughened	2_2_0-1	2_2_0-2
	NiBSi	Small	Roughened	2_3_0-1	2_3_0-2
	NiBSi	Large	Roughened	2_4_0-1	2_4_0-2

3 MICROSTRUCTURE ANALYSIS FOR SUBSTRATE MATERIAL - X38CRMOV5-1 STEEL

In this chapter, microstructures for all combinations of substrate materials - X38CrMoV5-1 steel and coatings are presented and analysed, as follows [14]:

- X38CrMoV5-1 steel, soft annealed - NiCrBSi coating
- X38CrMoV5-1 steel, tempered - NiCrBSi coating
- X38CrMoV5-1 steel, soft annealed - NiBSi coating
- X38CrMoV5-1 steel, tempered - NiBSi coating.

It should be noted that, for the images of the microstructures that follow, the upper part of the image represents the coating, and the lower part of the image represents the substrate. The exceptions are images with cracks (Figs. 5 and 7), where the upper part of the picture shows the substrate and the lower part the coating.

a) X38CrMoV5-1 Steel, Soft Annealed - NiCrBSi Coating

From Fig. 4, it is evident that when spraying on the non-roughened substrate surface, there was no separation of the coating from the substrate and that, as for the roughened substrate, spraying with both smaller and greater spraying resulted in the creation of a metallurgical connection between the coating and the substrate. The white zone is clearly pronounced between the coating and the substrate - the so-called diffusion zone. However, two zones are clearly visible during spraying on the non-roughened substrate surface with a long spraying distance as well as on the roughened surface (with a small and large spraying distance). Below the white zone of the so-called diffusion zone is a zone with austenitic grains, which is wider when spraying with a long spraying distance. In the paper [33], which describes the research results on spraying nickel alloys using flame spraying with subsequent fusing procedure and vibration, the authors also mentioned two zones. They state that the solidification of the coating starts from the coating/substrate boundary, where the substrate has the role of heat dissipation, and that at this boundary, an almost single-phase coating is created - Ni matrix without or with some eutectics. The second zone in the coating is a boron-free zone that has a greater thickness for vibration-subjected samples. The authors [34] also conclude that closer to the substrate, there are fewer precipitates in the coating itself (they cite the reason is the poor fusion that would encourage the growth of precipitates); A similar conclusion is made in the paper [35]. However, in the investigation of the corrosion resistance of NiCrBSi coatings [16], it was concluded that close to the coating/substrate boundary, there is a greater amount of a certain type of precipitates, which, due to the higher mass, are located at the bottom of the coating. In the research of the authors [36, 37], who studied plasma spraying of NiCrBSi coatings with subsequent flame fusing, the coating/substrate boundary was considered, and it was concluded that the coating/substrate boundary is constituted of two zones.

When spraying with a small spraying distance, for both types of preparation of the substrate, the grain boundaries are clearly expressed in the area under the coating and the substrate under the coating is fine grained with distributed carbides along the grain boundaries and inside the grains. Further away from the coating/substrate boundary towards the core the roughening of the austenitic grains is evident. Due to the longer heating time when spraying with larger spraying distance, for both types of the substrate preparation, in the substrate under the coating, the grain coarsening is greater.

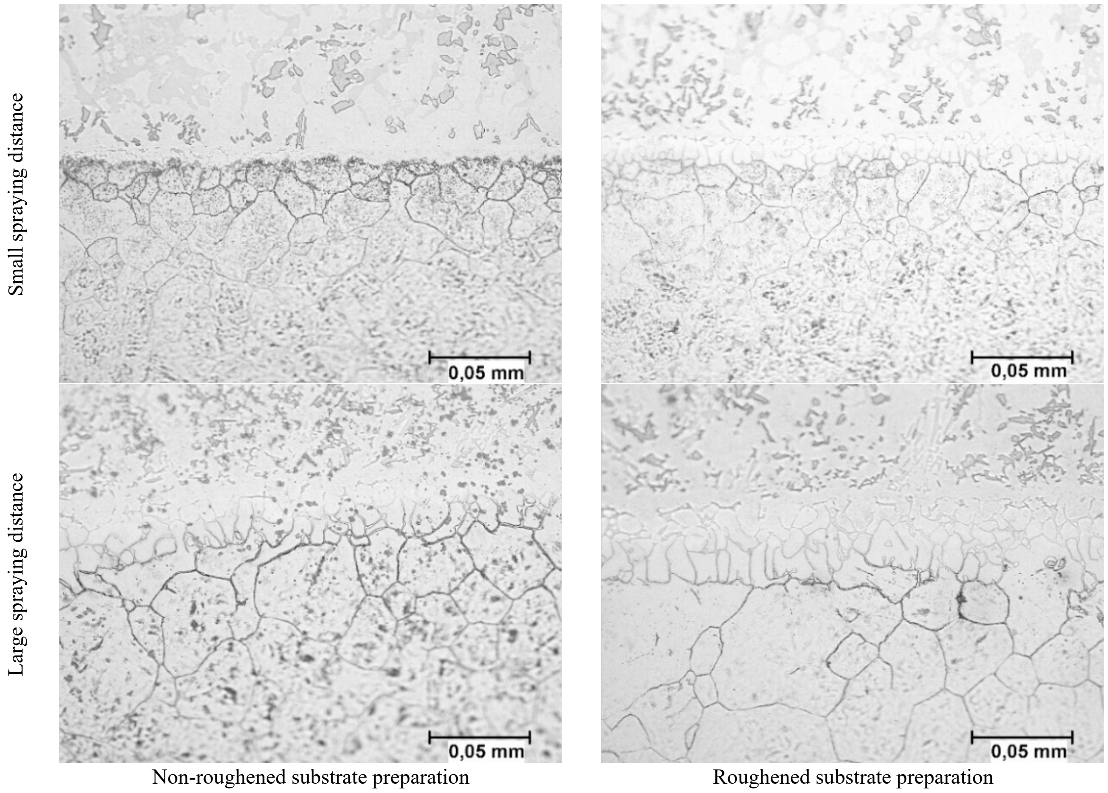


Figure 4 Microstructure system: steel substrate X38CrMoV5-1 (soft annealed) - NiCrBSi coating

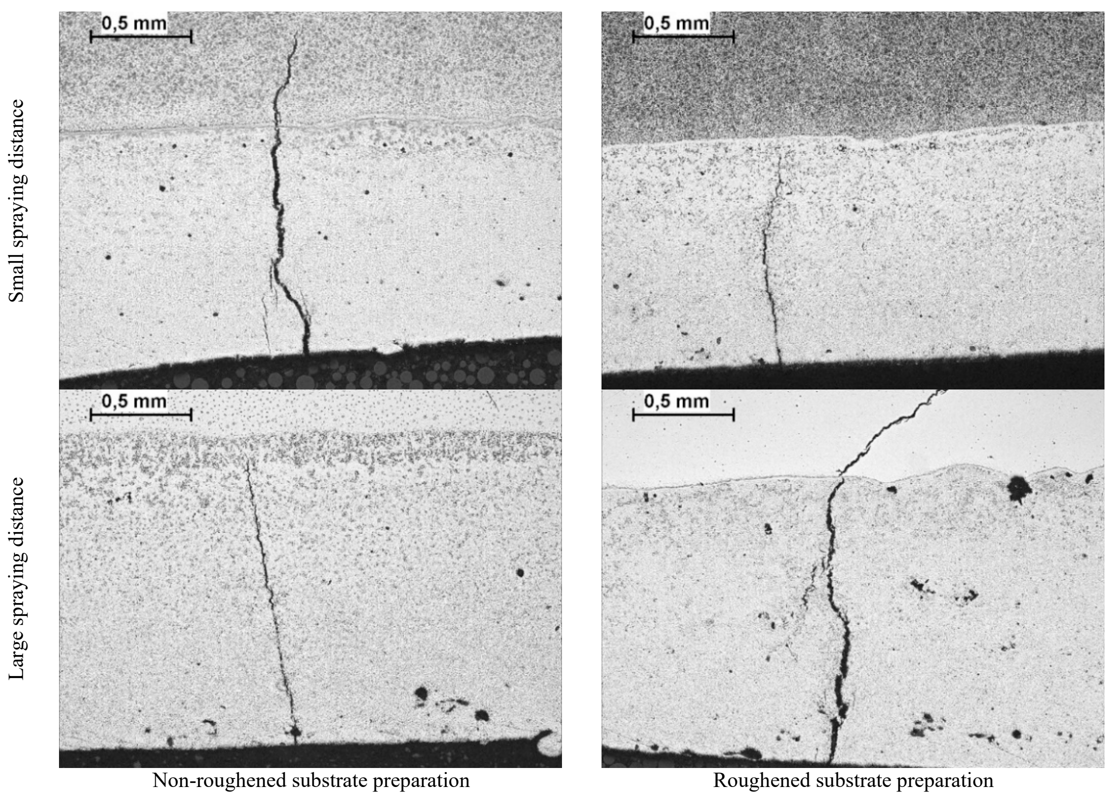


Figure 5 Cracks appearance in a system: steel substrate X38CrMoV5-1 (soft annealed) - NiCrBSi coating

The initial structure of the substrate before the spraying and the fusing process was ferrite with homogeneously distributed spheroidized chromium carbides, with a hardness of 192 HV10. After the process of spraying and fusing at a temperature of about 1040 °C and slow cooling in the air, there was a change of the structure throughout the whole substrate. The austenitizing temperature of X38CrMoV5-1 steel is 1030 °C, which is close to the fusing temperature. At a distance of 2 mm from the edge of the coating, where the cooling is faster, there was a partial transformation, and a bainite-martensitic structure was achieved. By moving away from the edge of the coating to the core due to slower cooling, a bainite structure was achieved. It has been proven that the simultaneous fusing process was actually the process of austenitizing of soft annealed steel X38CrMoV5-1, and cooling in the air still allowed the structure change.

There was an appearance of cracks on all samples, and it is assumed that the cracks were due to residual stresses in the substrate material because the fusing temperature of the NiCrBSi coating (1040 °C) is close to the austenitizing temperature of X38CrMoV5-1 steel (1030 °C). Due to heating the substrate material to this temperature and cooling in the air, a partial transformation into martensite occurred in the area under the coating. After the spraying process, x38CrMoV5-1 steel remained in a quenched state with no tempering following, and due to residual stresses and relaxation of these tensions, cold cracks appeared.

In Fig. 5, cracks are visible in both the substrate and the coating material on the samples that are not ground.

b) X38CrMoV5-1 Steel, Tempered - NiCrBSi Coating

Fig. 6 shows that, as with the combination of soft annealed steel X38CrMoV5-1 - NiCrBSi coating (Fig. 4), there was no separation of the coating from the substrate when spraying it onto an untreated surface. At the coating/substrate border, now, the two zones are also visible for all combinations of spraying distance and surface preparation.

When spraying with a small spraying distance, the diffusion zone (0.0005 mm of thickness) is less pronounced, which could now be measured, while when spraying with a large spraying distance, it is more pronounced (0.001-0.002 mm of thickness). For this spraying distance, there was a coarser grain in the substrate due to the longer heating time for both types of substrate preparation.

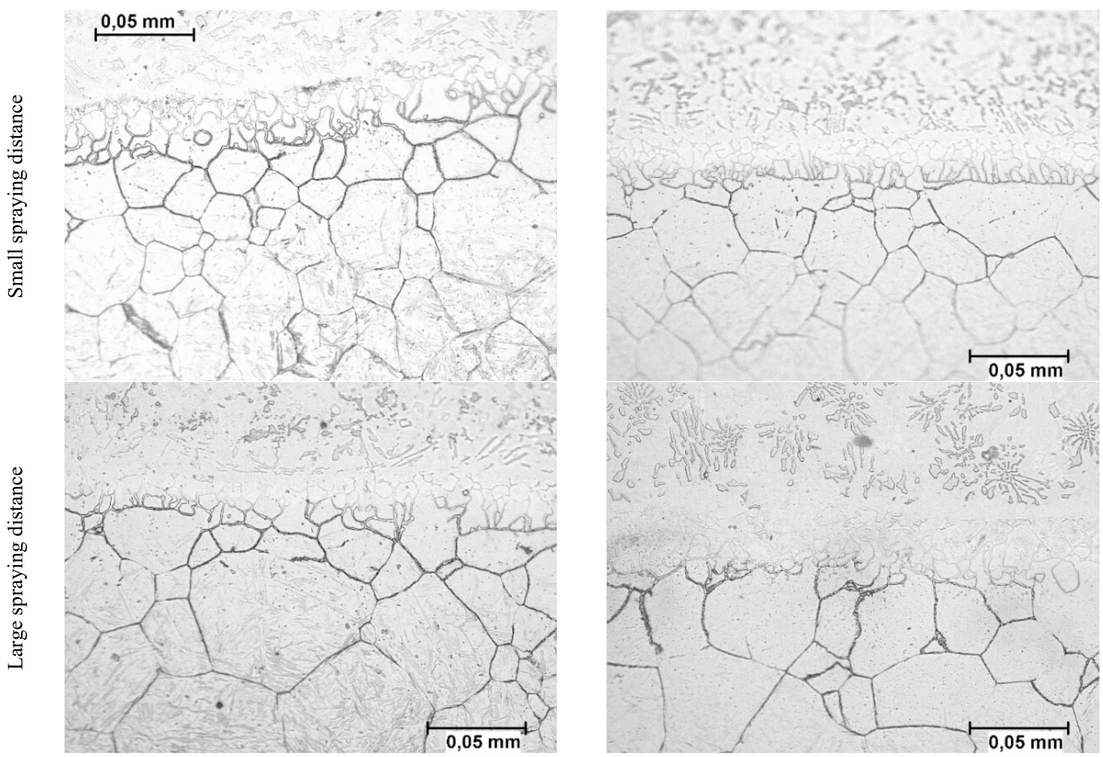


Figure 6 Microstructure system: steel substrate X38CrMoV5-1 (tempered) - NiCrBSi coating

The initial structure of the substrate before the spraying and fusing process was tempered martensite, with a hardness of 52-54 HRC. After the process of spraying and fusing at a temperature of about 1040 °C and cooling in the air, there was a change of structure throughout the substrate sample. During the spraying and fusing process, there was a re-quenching of the substrate material since the austenitizing temperature of X38CrMoV5-1 steel, 1030 °C, is close to the fusing temperature. At 2 mm from the edge of the coating, a martensitic structure with pronounced boundaries of the primary austenitic grain remained [38]. Moving away from the edge of the coating towards the core reduces the size of the primary grain and makes the structure more homogeneous. Due to the cooling in the air up to a distance of 3 mm from the edge of the coating, a complete transformation into martensite was achieved. After this distance, the structure of the substrate is, as before the spraying procedure, tempered martensite. In the paper [39], in which the martensitic stainless steel substrate was also quenched and tempered (the austenitizing temperature of 1020 °C is also similar to the fusing temperature), the authors do not refer to the structure of the substrate after the spraying process.

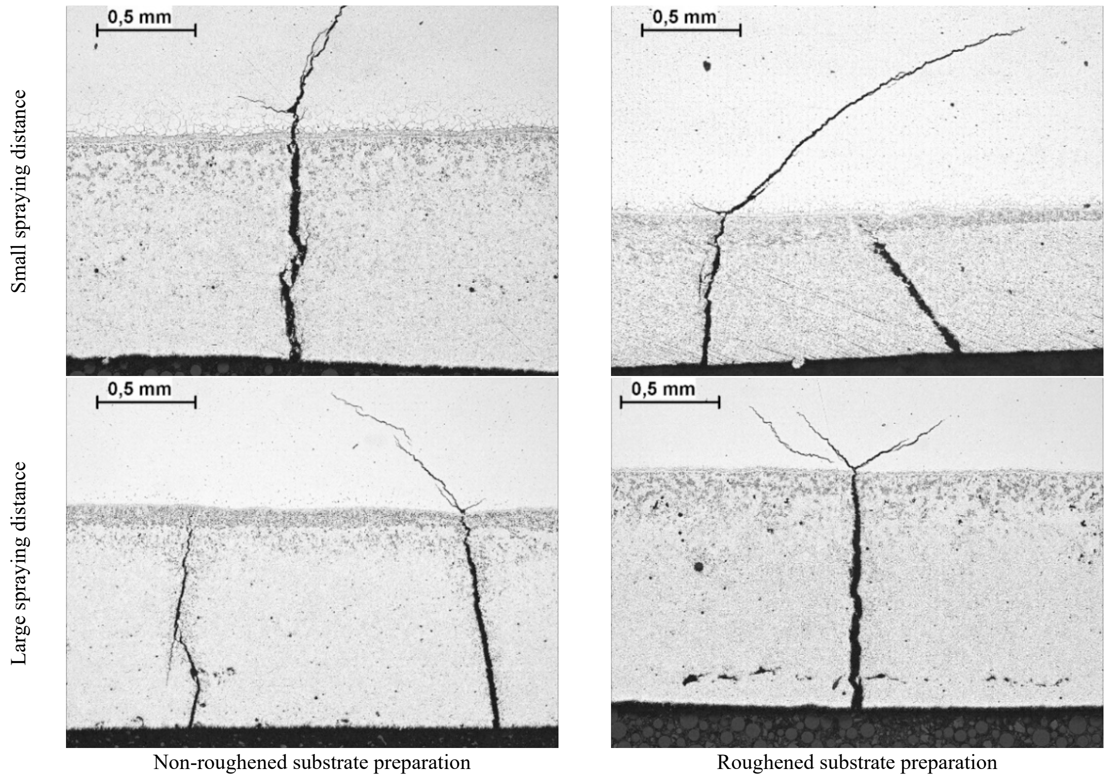


Figure 7 Cracks appearance in a system: steel substrate X38CrMoV5-1 (tempered) - NiCrBSi coating

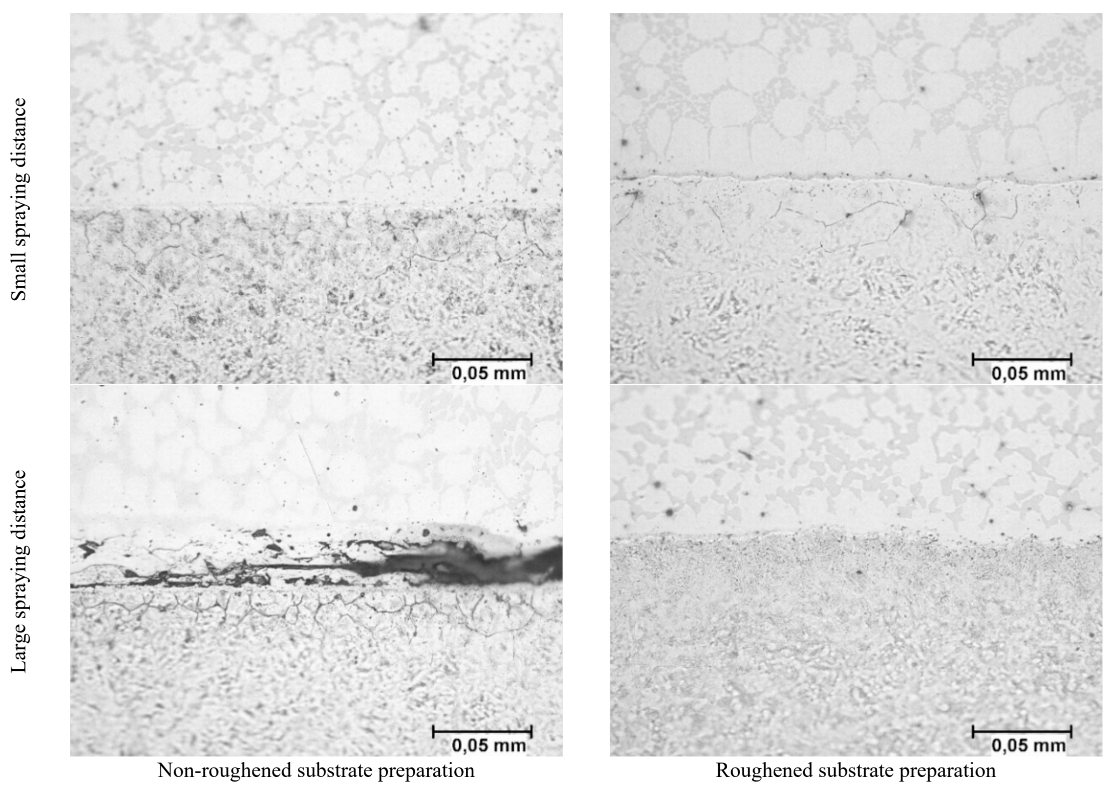


Figure 8 Microstructure system: steel substrate X38CrMoV5-1 (soft annealed) - NiBSi coating

As for the samples with a substrate in a soft annealed state, cracks appeared here in the coating and substrate (Fig. 7), which are now more pronounced and were formed due to residual stresses in the substrate material. As x38CrMoV5-1 steel had already been heat treated before the spraying process, during the spraying and fusing process at a temperature of about 1040 °C, the substrate material was re-quenched since the fusing temperature is close to the austenitizing temperature of this steel.

When heating the substrate material to this temperature and cooling the substrate faster in the area under the coating, there was a complete transformation into martensite up to 3 mm from the edge of the coating. After the spraying process, x38CrMoV5-1 steel remained in a quenched state and no tempering followed. Due to the residual stresses cold cracks appeared.

c) X38CrMoV5-1 Steel, Soft Annealed - NiBSi Coating

From Fig. 8 it is evident that when spraying the NiBSi coating, with a small spraying distance for both types of substrate preparation, as well as for spraying with a large spraying distance on the roughened surface, there is a narrow white zone so-called diffusion zone in which a metallurgical connection between the coating and the substrate is achieved during the fusing at a temperature of about 1070 °C. When spraying with a large spraying distance, and thus with less heat input on non-roughened substrate, there was no complete diffusion between the coating and the substrate material, or there was a subsequent separation of the coating from the substrate, when cooling to the ambient temperature, due to a greater difference in coefficients of thermal elongation between the coating material and the substrate. The coefficient of thermal elongation for the NiBSi alloy is 15.55×10^{-6} K^{-1}. The coefficient of thermal elongation for the substrate material, for the same temperature range, is 14×10^{-6} K^{-1} [40]. For a NiCrBSi alloy, this coefficient is lower than for the NiBSi coating and is 13.62×10^{-6} K^{-1}. A possible cause of partial separation of the coating may also be the local overheating, which can cause damage in the coating or at the coating/substrate boundary [41].

During the metallographic analysis of all the above-mentioned samples, not a single crack through the substrate or coating was observed, which was the case for the NiCrBSi coating on that same substrate.

Based on the structures shown in Fig. 8, it can be concluded that the substrate under the coating has a bainite-martensitic structure, and by moving away from the coating/ substrate boundary towards the core, the hardness of the substrate is lower, and thus the microstructure changes to bainite.

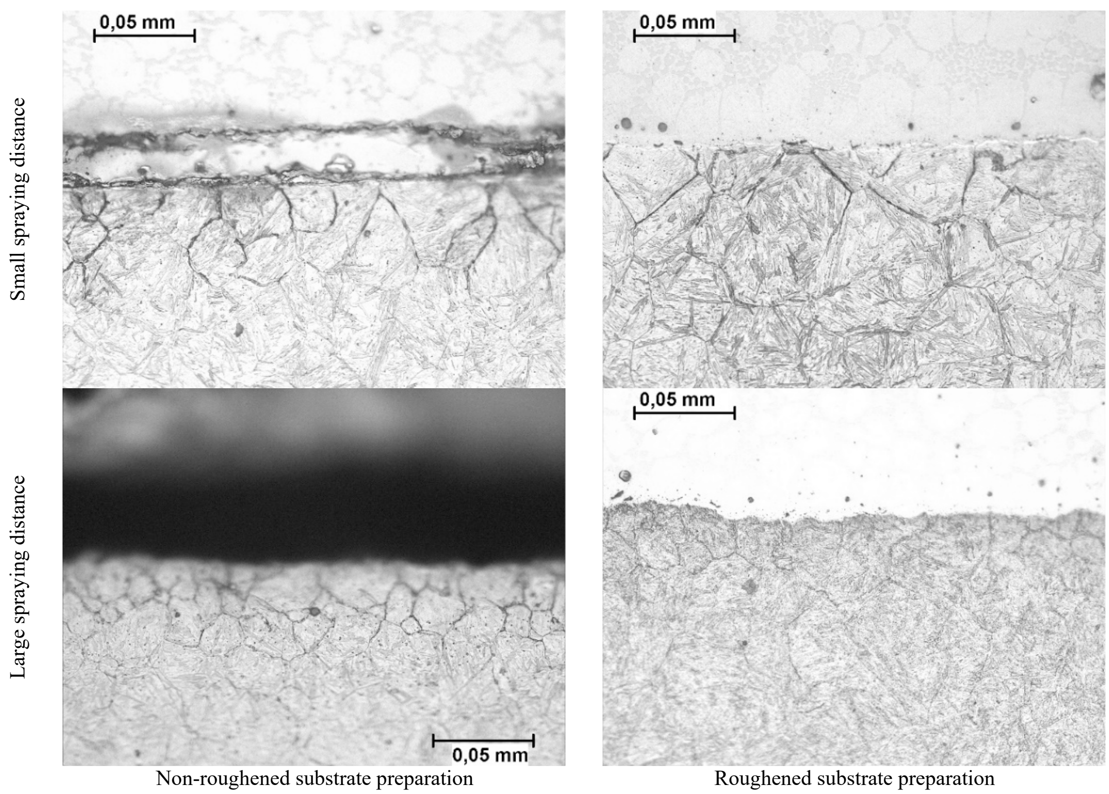


Figure 9 Microstructure system: steel substrate X38CrMoV5-1 (tempered) - NiBSi coating

d) X38CrMoV5-1 Steel, Tempered - NiBSi Coating

The combination of tempered steel substrate and NiBSi coating in Fig. 9 shows that when spraying on a non-roughened substrate surface with both larger and smaller spraying distances, there was a partial or complete separation of the coating from the substrate, while for the roughened substrate surface, a metallurgical connection was achieved between the coating and the substrate material. When spraying with a small spraying distance on the roughened substrate surface, the substrate under the coating does not have a coarse grain structure and a weak distribution of carbides across the grain boundaries is visible. Under the coating, due to faster cooling, a martensite structure with a relatively large primary grain was achieved. In the case of a roughened substrate, with spraying with a large spraying distance, the substrate under the coating has less pronounced grain boundaries with partially distributed carbides. The substrate material retained the loosened martensite structure it had before the spraying and fusing process.

On all samples for this type of coating (NiBSi), according to metallographic analysis, not a single crack was observed through the substrate and coating, for the substrate in tempered or in a soft annealed state. On the more fragile NiCrBSi coating, the relaxation of the tension of the substrate (untempered martensite) led to the formation of cracks (Figs. 5 and 7), while, for this tougher NiBSi coating, energy was absorbed, and no cracks were formed.

4 CONCLUSION

When investigating the microstructure of the coating/substrate system using the factorial design of the experiment, along with the investigated coatings, the spraying distance from the workpiece, the preparation of the substrate-non-roughened and roughened and heat treatment of the substrate-soft annealed and tempered condition were systematically combined. It can be concluded that the application of NiCrBSi coating by spraying on a tool steel substrate would not make sense due to the appearance of cracks caused by the formation of martensitic structure of the substrate after spraying and relaxation of residual stresses because the temperature of the fusing was close to the austenitizing temperature of the specified steel, for which conversion into a martensitic structure was enabled even during slow cooling in the air.

For the NiBSi coating, no cracks were observed, as energy was absorbed for this tougher coating. However, due to the greater difference in coefficients of thermal elongation between X38CrMoV5-1 steel and NiBSi coating and the increased volume of martensitic structure, there was a separation of the mentioned coating on a non-roughened surface.

Acknowledgements

This research was funded by the University of Slavonski Brod, Republic of Croatia (project SmartPRO).

5 REFERENCE

[1] Shuecamlue, S., Taman, A., Khamnantha, P. & Banjongprasert, C. (2024). Influences of flame remelting and WC-Co addition on microstructure, mechanical properties and corrosion behavior of NiCrBSi coatings manufactured via HVOF process. *Surfaces and Interfaces, 24*, 104135. https://doi.org/10.1016/j.surfin.2024.104135

[2] Aliabadi, M., Khodabakhshi, F., Soltani, R. & Gerlich, A. P. (2023). Modification of flame-sprayed NiCrBSi alloy wear-resistant coating by friction stir processing and furnace re-melting treatments. *Surface and Coatings Technology, 455*, 129236. https://doi.org/10.1016/j.surfcoat.2023.129236

[3] Szajna, E., Moskal, G., Tupaj, M., Dresner, J., Dudek, A., Szymański, K., Tomaszewska, A., Trzcionka-Szajna, A., Mikuśkiewicz, M. & Łysiak, K. (2024). The influence of laser remelting on microstructural changes and hardness level of flame-sprayed NiCrBSi coatings with tungsten carbide addition. *Surface and Coatings Technology, 479*, 130403. https://doi.org/10.1016/j.surfcoat.2024.130403

[4] Habib, K. A., Cano, D. L., Heredia Alvaro, J. A., Serrano-Mira, J., Llopis, R., López Moreno, D. & Mohammed, S. S. (2022). Effects of thermal spraying technique on the remelting behavior of NiCrBSi coatings. *Surface and Coatings Technology, 444*, 128669. https://doi.org/10.1016/j.surfcoat.2022.128669

[5] Varis, T., Mäkelä, A., Suhonen, T., Laurila, J. & Vuoristo, P. (2023). Integrity of APS, HVOF and HVAF sprayed NiCr and NiCrBSi coatings based on the tensile stress-strain response. *Surface and Coatings Technology, 452*, 129068. https://doi.org/10.1016/j.surfcoat.2022.129068

[6] Xuan, H.-N., Chen, L.-Y., Li, N., Wang, H., Zhao, C., Bobrov, M., Lu, S. & Zhang, L.-C. (2022). Temperature profile, microstructural evolution, and wear resistance of plasma-sprayed NiCrBSi coatings under different powers in a vertical remelting way. *Materials Chemistry and Physics, 292*, 126773. https://doi.org/10.1016/j.matchemphys.2022.126773

[7] Zhou, J., Guo, W., He, D., Huang, Y., Cai, Z., Zhou, L., Xing, Z. & H. Wang (2022). Study on preparation and wear resistance of NiCrBSi-WC/Co composite coatings by pulsed magnetic field assisted supersonic plasma spraying. *Surface and Coatings Technology, 448*, 128897. https://doi.org/10.1016/j.surfcoat.2022.128897

[8] Hulka, I., Utu, D., Serban, V. A., Negrea, P., Lukáč, F. & Chráska, T. (2020). Effect of Ti addition on microstructure and corrosion properties of laser cladded WC-Co/NiCrBSi(Ti) coatings. *Applied Surface Science, 504*, 144349. https://doi.org/10.1016/j.apsusc.2019.144349

[9] Medabalimi, S. R., Ramesh, M. R. & Kadoli, R. (2021). Developing partially oxidized NiCr coatings using the combined flame spray and plasma spray process for improved wear behaviour at high temperature. *Wear, 478-479*, 203885. https://doi.org/10.1016/j.wear.2021.203885

[10] Zhu, Z., Wei, L., Chen, J., Cheng, J., Chen, W., Zhu, S. & Yang, J. (2024). Tribological behavior of NiCrAlY/ Al2O3/h-BN metal-ceramic composite coatings at wide temperature range by detonation spraying. *Wear, 550-551*, 205421. https://doi.org/10.1016/j.wear.2024.205421

[11] Srichen, A., Linjee, S. & Banjongprasert, C. (2023). Corrosion behavior of heat-treated NiCrMoAl alloy coatings produced via arc spraying. *Surfaces and Interfaces, 39*, 102880. https://doi.org/10.1016/j.surfin.2023.102880

[12] Šimunović, K., Šarić, T. & Šimunović, G. (2014). Different Approaches to the Investigation and Testing of the Ni-Based Self-Fluxing Alloy Coatings - A Review. Part 1: General Facts, Wear and Corrosion Investigations. *Tribology Transactions, 57*(6), 955-979. https://doi.org/10.1080/10402004.2014.927547

[13] Šimunović, K., Šarić, T. & Šimunović, G. (2014). Different Approaches to the Investigation and Testing of the Ni-Based Self-Fluxing Alloy Coatings - A Review. Part 2: Microstructure, Adhesive Strength, Cracking Behavior, and Residual Stresses Investigations. *Tribology Transactions, 57*(6), 980-1000. https://doi.org/10.1080/10402004.2014.927548

[14] Havrlišan, S. (2017). Stohastički pristup modeliranju i optimiranju procesa naštrcavanja pomoću plinskog plamena. *PhD dissertation*, Josip Juraj Strossmayer University of Osijek. (in Croatian)

[15] Kim, K. T. & Kim, Y. S. (2010). Effects of Counterpart Materials on Wear Behavior of Thermally Sprayed Ni-Based Self-Flux Alloy Coatings. *International Journal of Modern Physics B, 24*(15-16), 3023-3028. https://doi.org/10.1142/S0217979210066021

[16] Bergant, Z., Trdan, U. & Grum, J. (2014). Effect of High-temperature Furnace Treatment on the Microstructure and Corrosion Behavior of NiCrBSi Flame-Sprayed Coatings. *Corrosion Science, 88*, 372-386. https://doi.org/10.1016/j.corsci.2014.07.057

[17] Chaliampalias, D., Vourlias, G., Pavlidou, E., Skolianos, S., Chrissafis, K. & Stergioudis, G. (2009). Comparative Examination of the Microstructure and High Temperature Oxidation Performance of NiCrBSi Flame Sprayed and Pack Cementation Coatings. *Applied Surface Science, 255*(6), 3605-3612. https://doi.org/10.1016/j.apsusc.2008.10.006

[18] Bergant, Z. & Grum, J. (2009). Quality Improvement of Flame Sprayed, Heat Treated, and Remelted NiCrBSi Coatings. *Journal of Thermal Spray Technology, 18*(3), 380-391. https://doi.org/10.1007/s11666-009-9304-7

[19] Novosel, M., Cajner, F. & Krumes, D. (1996). *Alatni materijali*. Slavonski Brod, Strojarski fakultet u Slavonskom Brodu. (in Croatian)

[20] Dadić, Z., Živković, D. & Čatipović, N. (2016). Tribological Wear Mechanisms of Molds for High Pressure Die Casting. *Metalurgija, 55*(2), 249-252.

[21] Ahn, D.-G. (2013). Hardfacing Technologies for Improvement of Wear Characteristics of Hot Working Tools: A Review. *International Journal of Precision Engineering and Manufacturing, 14*(7), 1271-1283. https://doi.org/10.1007/s12541-013-0174-z

[22] Tušek, J., Kosec, L., Lešnjak, A. & Muhič, T. (2012). Electrospark Deposition for Die Repair. *Metalurgija, 51*(1), 17-20.

[23] Lei, Y., Sun, R., Tang, Y. & Niu, W. (2015). Microstructure and Phase Transformations in Laser Clad CrxSy/Ni Coating on H13 Steel. *Optics and Lasers in Engineering, 66*, 181-186. https://doi.org/10.1016/j.optlaseng.2014.09.006

[24] Dadić, Z., Živković, D., Čatipović, N., & Bilić, J. (2017). High Pressure Die Casting Mould Repair Technologies. *Proceedings of 7th International Conference "Mechanical Technologies and Structural Materials 2017"*, 23-28.

[25] Krumes, D. (2000). *Toplinska obradba*. Slavonski Brod, Strojarski fakultet u Slavonskom Brodu. (in Croatian)

[26] [Samardžić, I., Marić, M. & Konjatić, P. (2007). Tehnologija naštrcavanja vagonskih osovina postupkom plamena gorivog plina. *Zbornik radova Tehnologična primjena postupaka zavarivanja i zavarivanju srodnih tehnika u izradi zavarenih konstrukcija i proizvoda*, 201-209. (in Croatian)

[27] Karimi, M. R., Salimijazi, H. R. & Golozar, M. A. (2016). Effects of Remelting Processes on Porosity of NiCrBSi Flame Sprayed Coatings. *Surface Engineering, 32*(3), 238-243. https://doi.org/10.1179/1743294415Y.0000000107

[28] Grum, J. & Bergant, Z. (2008). Optimization of the Flame-Spraying Process and Improvement of Properties of a NiCrBSi Coating by Heat Treatment. *Materials Science Forum, 589*, 373-378. https://doi.org/10.4028/www.scientific.net/MSF.589.373

[29] Akebono, H., Komotori, J. & Shimizu, M. (2008). Effect of Coating Microstructure on the Fatigue Properties of Steel Thermally Sprayed with Ni-Based Self-Fluxing Alloy. *International Journal of Fatigue, 30*(5), 814-821. https://doi.org/10.1016/j.ijfatigue.2007.07.003

[30] Akebono, H., Komotori, J. & Suzuki, H. (2006). The Effect of Coating Thickness on Fatigue Properties of Steel Thermally Sprayed with Ni-Based Self-Fluxing Alloy. *International Journal of Modern Physics B, 20*(25-27), 3599-3604. https://doi.org/10.1142/S0217979206040052

[31] Voyer, J. & Kreye, H. (2003). Determination of Cracking Resistance of Thermal Spray Coatings during Four-Point Bend Testing using an Acoustic Emission Technique. *Journal of Thermal Spray Technology, 12*(3), 416-426. https://doi.org/10.1361/105996303770348285

[32] Oliveira, F., Hernandez, L., Berrios, J. A., Villalobos, C., Pertuz, A. & Puchi Cabrera, E. S. (2001). Corrosion-Fatigue Properties of a 4340 Steel Coated with Colmonoy 88 Alloy. Applied by HVOF Thermal Spray, Surface and Coatings Technology, 140(2), 128-135. https://doi.org/10.1016/S0257-8972(01)01015-5

[33] Škamat, J., Valiulis, A. V., Kurzydlowski, K. J., Černašejus, O., Lukauskaite, R. & Zwolinska, M. (2013). NiCrSiB Thermal Sprayed Coatings Refused under Vibratory Treatment. *Materials Science (Medziagotyra), 19*(4), 377-384. https://doi.org/10.5755/j01.ms.19.4.3386

[34] Gonzalez, R., Cadenas, M., Fernandez, R., Cortizo, J. L. & Rodriguez, E. (2007). Wear Behaviour of Flame Sprayed NiCrBSi Coating Remelted by Flame or by Laser. *Wear, 262*(3-49), 301-307. https://doi.org/10.1016/j.wear.2006.05.009

[35] Harsha, S., Dwivedi, D. K. & Agarwal, A. (2008). Influence of CrC Addition in Ni-Cr-Si-B Flame Sprayed Coatings on Microstructure, Microhardness and Wear Behaviour. *International Journal of Advanced Manufacturing Technology, 38*(1-2), 93-101. https://doi.org/10.1007/s00170-007-1072-2

[36] Rudenskaya, N. A., Shveikin, G. P., Kopysov, V. A. & Rudenskaya, M. V. (2011). Fire Polished Coating-Steel Interface Structure. *Doklady Chemistry, 441*(2), 383-386. https://doi.org/10.1134/S0012500811120020

[37] Rudenskaya, N. A., Shveikin, G. P., Kopysov, V. A. & Rudenskaya, M. V. (2013). Specific Features of Interface Formation for Metal-Ceramic Coatings on a Steel Base. *Russian Journal of Applied Chemistry, 86*(4), 475-481. https://doi.org/10.1134/S1070427213040034

[38] Šuman, H. (1981). *Metalografija*, Beograd, Tehnološko-metalurški fakultet Univerziteta u Beogradu. (in Serbian)

[39] Shao, T. & Luo, J. (2005). Response Frequency Spectrum Analysis for Impact Behavior Assessment of Surface Materials. *Surface and Coatings Technology, 192*(2-3), 365-373. https://doi.org/10.1016/j.surfcoat.2004.04.076

[40] SIJ Metal Ravne Steel Selector, v. 5.0, SIJ Metal Ravne d.o.o., 2016.

[41] Thermal spraying - Spraying and fusing of self-fluxing alloys (EN ISO 14920:2015), Zagreb, Hrvatski zavod za norme.

Authors' contacts:

Sara Havrlišan, PhD, Assistant Professor
(Corresponding author)
University of Slavonski Brod,
Mechanical Engineering Faculty in Slavonski Brod,
Trg I. B. Mažuranić 2, 35000 Slavonski Brod, Croatia
E-mail: shavrlisan@unisb.hr

Štefanija Klarić, PhD, Associate Professor
Charles Darwin University
Ellengowan Drive, Casuarina, NT 0810, Australia
E-mail: Stefanija.klaric@cdu.edu.au

Branko Štrbac, PhD, Associate Professor
University of Novi Sad,
Faculty of Technical Sciences,
Trg Dositeja Obradovića, 21000 Novi Sad, Serbia
E-mail: strbacb@uns.ac.rs

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240626221644

Original scientific paper
Received: 2024-06-26, Accepted: 2024-07-03

Comparison of Chromatic Assimilation Effects Depending on Printing Substrate in the Munker-White Model

Katja Milković*, Damir Vusić, Martina Hajdek

Abstract: Understanding the influence of paper surface structures on color perception is crucial for the printing industry. This research investigates the impact of horizontal and vertical parallel lines on chromatic assimilation using the Munker-White grid model. The study employs the perceived ΔE_{00} metric to quantify color differences and determine whether these textures affect perceived color accuracy and consistency differently. Results reveal variations in color perception due to surface textures. The biggest color difference (ΔE_{00}) occurred at 60% RTV coverage for cyan, where the vertical structure showed a significantly lower value than the horizontal. For the yellow primary stimulus, the greatest difference was observed at 20% RTV coverage, with both structures showing high initial values that decrease with increased coverage. These findings provide valuable insights for improving color management practices and enhancing the quality and reliability of color reproduction on various substrates, contributing to advancements in printing technology and color science.

Keywords: chromatic assimilation; color perception; Munker-White model; paper texture; printing surface; printing technology; subtractive synthesis

1 INTRODUCTION

The study of color perception and reproduction is a crucial aspect of the printing industry. One of the significant challenges in this field is understanding how different surface structures of printing substrates influence the perceived color of printed materials. Chromatic assimilation, a phenomenon where the surrounding colors affect the perception of a given color, is particularly relevant in this context. The effects of chromatic induction and assimilation are based on the principle of surface light induction [1]. Visual effects are considered undesirable due to the unwanted shift in the perceived color tone by the observer or consumer, which cannot be detected using instrumental colorimetric methods [2].This article focuses on the impact of paper surface structures, specifically horizontal and vertical lines, on chromatic assimilation in the context of subtractive synthesis.

CMYK subtractive synthesis is the process used in color printing where cyan, magenta, yellow, and black inks are combined to produce all other colors. The quality and accuracy of color reproduction in this process depend significantly on the interaction between the inks and the substrate. Different surface textures can alter the way light is reflected and absorbed by the printed material, leading to variations in color perception. In 2005 Spitzer et al discusses a computational model for chromatic induction that includes mechanisms for both local and remote contrast effects. It provides insights into how different spatial frequencies and orientations affect chromatic assimilation [2]. Trends in graphic reproduction increasingly demand the use of non-standard, structured materials to further enhance the product experience. This can cause problems during the printing reproduction process and result in different color perception compared to materials that do not have structural irregularities. According to the research by Motoyoshi et al., it has been established that increasing the structural irregularity of the material enhances the perception of the brightness of a certain color. It has also been found that adapting to the structure can lead to altered color perception and experience [3, 4].

To investigate these effects, this study employs the Munker-White grid, a well-known visual illusion model. The Munker-White grid is used to illustrate how different background patterns and structures can significantly influence the perceived color of objects [5]. By applying this model, we can better understand the mechanisms of chromatic assimilation and how they are affected by the surface structure of the substrate.

Previous studies have indicated that surface roughness and texture can cause noticeable differences in the final appearance of printed colors. Betz et al. extended White's research by adding horizontal and vertical interferences, i.e., lines, to the Munker-White grid. They discovered that visual adaptation to edges that are positioned parallel to the White grid enhances the brightness of the observed area, while adaptation to horizontal edges has the opposite effect on the observer, resulting in a lower perception of brightness, or even its cancellation [6].

However, there is limited research specifically comparing the effects of horizontal and vertical structures on color perception using the ΔE metric, which quantifies the difference between two colors. To colorimetrically describe and compare the mentioned effects using recognized methods, Milković investigated the color differences using ΔE_{00} [7]. In his work, Helson stated that thinner lines create a narrowing effect that causes assimilation, while thicker lines result in inhibition, leading to greater contrast [8]. Chen et al. investigated the Munker-White effect and chromatic induction, establishing that the brightness of a color can be altered if it is surrounded by different structures within the observer's visual field [9]. This article aims to fill this gap by providing a detailed analysis of how these two distinct surface structures affect the perceived color accuracy and consistency in printed materials.

The objective of this study is to determine whether the horizontal and vertical paper structures differently influence the ΔE values and thus the color perception by observers. Bresan et al. studied the effect of simultaneous contrast on structured and unstructured image backgrounds and concluded that a structured image background enhances this effect [10]. By conducting controlled experiments and thorough analysis using the Munker-White grid model, this research seeks to provide insights that could lead to improved color management practices in the printing industry. The findings of this study will be valuable for printers and material scientists looking to enhance the quality and reliability of color reproduction on different substrates.

The use of the Munker-White grid allows for a controlled comparison [11] of how these surface structures influence color perception under standardized conditions. This approach not only highlights the practical implications for the printing industry but also contributes to the broader understanding of visual perception and color science.

In summary, this article explores the relationship between the surface structure of printing substrates and chromatic assimilation in CMYK printing. By focusing on horizontal and vertical linear textures and utilizing the Munker-White grid model, it aims to provide a deeper understanding of the factors influencing color perception. This study ultimately contributes to advancements in printing technology and color science, offering potential improvements in color management and reproduction quality.

2 EXPERIMENTAL PART

The experimental part of this study determines whether the direction of lines in the lineature structure of the printing surface affects the occurrence of chromatic assimilation effects in the graphic reproduction process. It also investigate of the intensity of psychophysical visual effects of chromatic assimilation on the Munker-White grid under the influence of the following parameters: different printing surface structures (2 regular geometric structures of the printing substrate that differ only in the direction of the lines were observed), different combinations of primary and secondary stimulus colors (combinations of primary colors of subtractive and additive synthesis) and different percentages of printed surface coloring (20%, 40%, 60%, 80% coverage).

2.1 Methodology of the Experiment

A printing substrate was selected for the study that is identical in its psychophysical characteristics and possesses a final surface finish (horizontal "Linea" and vertical "Fili" lines structure). In addition to its surface finish, the printing substrate has identical parameters and properties in its basic technical specifications, including the same weight, brightness, thickness, and stiffness. The chosen paper is a wood-free offset paper made of 100% cellulose, marketed under the artistic name Astroprint, produced by Cordenons, with a weight of 280 g/m². The technical characteristics of the paper, as defined by the manufacturer, are shown in the Tab. 1.

After selecting the printing substrate, test forms were designed and printed on the chosen substrates using identical technique under unchanged conditions. The obtained prints and reference atlases were subjected to spectrophotometric measurement of CIE $L^*a^*b^*$ values, followed by visual binocular matching with a test group of 20 participants.

Table 1 Technical Characteristics of the Paper

Technical Characteristics	Unit of Measurement	Target	Measurement Method
Weight	g/m²	280	MCM-003 (ISO 536)
Thickness	mm	0.355	MCM-004 (ISO 534)
Brightness	%	105	MCM-078 (ISO 2470)
Stiffness (Taber 15°) MD	mN	175	MCM-023 (ISO 2493)
Stiffness (Taber 15°) CD	mN	80	MCM-023 (ISO 2493)

To observe the background effect of chromatic assimilation, a Munker-White line grid was constructed, within which the left and right rectangular elements have the same brightness. This design was used specifically to define the influence of the geometric design of the line grid, which is often used in graphic solutions when it interpolates with a structure that is also geometrically regularly arranged like the Munker-White grid but with different structure direction. The grid has geometrically identical lines and spacings with variations in the observed color (primary stimulus) and the background surrounding that color (secondary stimulus) in the basic colors used in graphic reproduction (Tab 2). The printing substrate also has geometrically identical lines and spacings, but with differing directions.

Table 2 Color Pairs of Test Sample Primary and Secondary Stimuli

Primary Stimulus	Secondary Stimulus Left	Secondary Stimulus Right
Cyan	Blue	Green
Magenta	Red	Blue
Yellow	Red	Green

The primary stimulus was varied in ranges of 20%, 40%, 60%, and 80% coverage, while the secondary stimulus was always 100% coverage (Fig. 1).

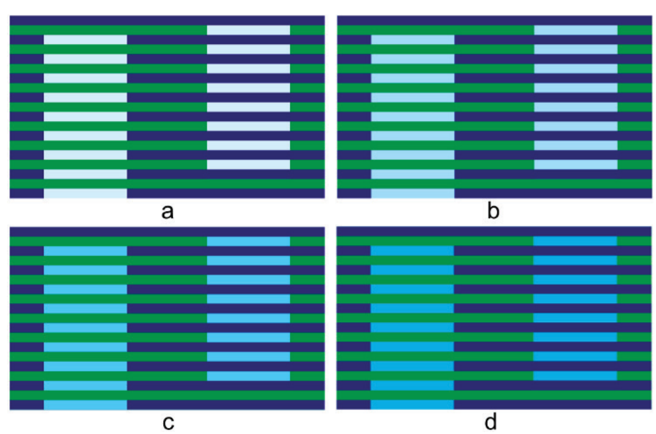


Figure 1 Test Form for Primary Stimulus Cyan (a - 20% RTV, b - 40% RTV, c - 60% RTV, d - 80% RTV)

The size of the test cards was made in accordance with the standard viewing conditions prescribed by ISO 3664:2009 (viewing conditions for graphic technology). The color atlas was constructed such that the RTV range is defined by fields differing by 2% RTV, where the first field has a surface coverage percentage of 2% RTV and the last 100% RTV, which gives a total of 50 fields with a 2% increment step. It was constructed to cover the entire area of perception (Fig. 2).

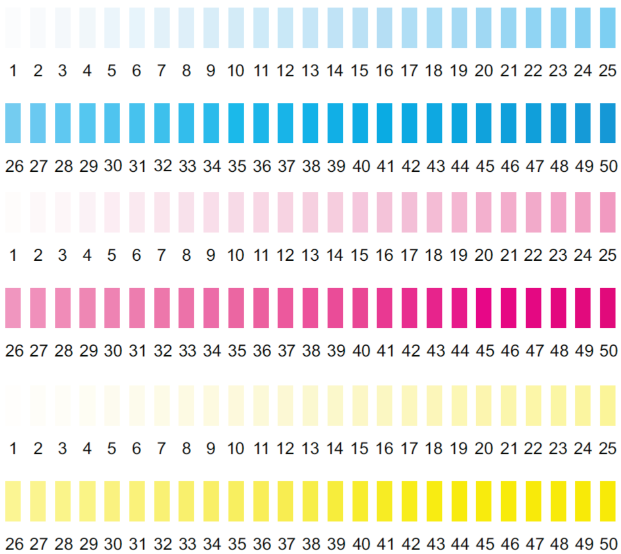


Figure 2 Reference Color Atlas for Primary Stimuli Cyan, Magenta, and Yellow

The described test form was printed on a calibrated digital printing machine, Ricoh C7100 XPRO, with a resolution of 1200 × 4800 dpi without the use of ICC profiles, directly from Adobe Illustrator CS6. Powder toner was used for the printing. Before printing, the paper was conditioned in a room under prescribed standard environmental conditions (temperature of 23 °C and relative humidity of 55%) for 48 hours.

2.2 Instrumental and Visual Analysis

The quantification of reference fields was performed using an Xrite Exact Standard 1 reflective spectrophotometer. The wavelength range used for measurement was from 400 to 700 nm, with a gas light source at a temperature of 2850 K. The measurement step was 10 nm, and the illumination geometry was set to 45°/0°. The measurement accuracy of the device, or the average deviation in terms of reflectance, is up to 0.5% per wavelength step (the calibration reference was measured by the Munsell laboratory with an accuracy of $\Delta E^* = 0.25$, using a D50 light source and a viewing angle of 20 degrees).

To improve the precision of the statistical results, the measurement was repeated 10 times for each type of paper, after which the average values of the measurements for the atlas and test forms were calculated (Tab. 3).

For conducting the visual analysis, 20 participants were selected who underwent the Farnsworth-Munsell 100 Hue Test (FM-100 test). Only the participants who successfully passed the test were included in further analysis. The average age of the participants was 21 years, with an equal representation of male and female participants.

Table 3 Measured $L^*a^*b^*$ Values

	%	"Fili"			"Linea"		
		C	M	Y	C	M	Y
L	20	92,36	90,49	98,33	91,03	89,26	97,48
a		–7,48	10,38	–3,52	–7,73	11,86	3,73
b		–7,06	–2,84	13,37	–7,31	–3,29	16,19
L	40	82,51	76,77	96,81	79,37	77,83	95,49
a		–17,4	30,99	–6,43	–19	28,97	–6,6
b		–18,34	–6,61	37,4	–19,77	–6,35	38,99
L	60	72,85	63,3	94,79	71,94	63,32	93,94
a		–25,67	51,95	–7,24	–25,04	50,91	–7,06
b		–25,3	–6,75	54,96	–25,74	–6,12	56,59
L	80	66,93	57,41	94,74	63,63	53,96	93,33
a		–30,78	61,37	–7,46	–32,64	65,81	–7,78
b		–29,15	–3,94	66,06	–31,51	–3,37	71,59

The psychophysical part of the experiment for visual evaluation was conducted under actual graphic production conditions, allowing participants to evaluate the samples under the following conditions: a viewing angle of 10°, a distance of 50 cm between the participant and the test sample, an environment with a neutral matte gray surface, illumination of 2000 lux with a light temperature of 5000 K in accordance with the ISO 3664:2017 standard.

During the experiment, the test sample and the atlas were placed parallel to each other within the full visual field of the participant. Each participant's task was to identify the field in the atlas that most closely matched the observed line (primary stimulus) located between the two observed secondary stimuli on the left and right sides of the test form. Participants had to carefully observe and compare the samples and select the best match between the test cards and the fields in the atlas to correctly complete their task (Fig. 3).

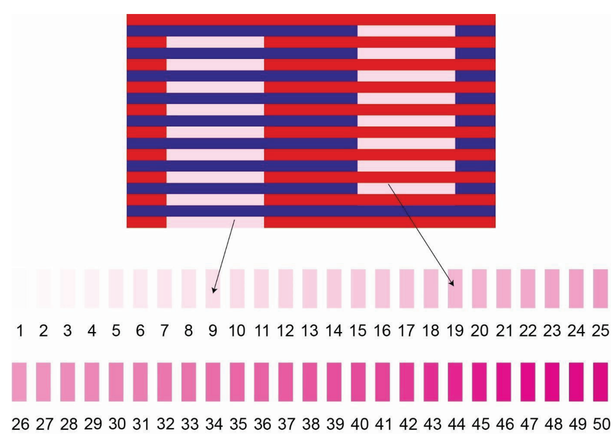


Figure 3 Visual Analysis Based on the Color Atlas

3 RESEARCH RESULT AND DISCUSSION

Statistical analysis of the obtained data was performed using the GraphPad Prism 9.5.1 software package. The results of the statistical analysis were tested using the Mann-Whitney statistical method to examine the impact of the psychophysical visual effect of chromatic assimilation depending on the secondary stimulus with the combination of colors of all primary stimulus at all percentages of coverage (20%, 40%, 60%, 80%) on the observed printing substrates ("Fili" and "Linea").

The following hypotheses were tested using the specified method: H_0: The medians of the two samples are equal, H_1: The medians of the two samples are not equal. The formulas used for the calculations are as follow [12] :

$$U_x = n_x n_y + \frac{n_x(n_x+1)}{2} - R_x \tag{1}$$

$$U_y = n_x n_y + \frac{n_y(n_y+1)}{2} - R_y \tag{2}$$

Where: U_x is the Mann-Whitney U statistic for sample x, U_y is the Mann-Whitney U statistic for sample y, n_x is the sample size of sample x, n_y is the sample size of sample y, R_x is the sum of the ranks for sample x, R_y is the sum of the ranks for sample y. The expected values of U are calculated using the following formulas:

$$\mu_U = \frac{n_x n_y}{2} = \frac{\left(U_x + U_y\right)}{2} \tag{3}$$

$$\sigma_U = \sqrt{\frac{n_x n_y (N+1)}{12}} \tag{4}$$

The total number of samples is denoted by N, where $N = (nX + nY)$. The mean value of the U distribution is denoted by μ_U, while σ_U represents the standard deviation. Through the process of normalizing the U variable, we obtain a new variable Z that follows a standard normal distribution. In other words, after normalization, the Z variable is distributed according to the standard normal distribution.

$$Z = \frac{\left[U - \left(\frac{n_x n_y}{2}\right)\right]}{\sigma_U} \tag{5}$$

At a significance level of $\alpha = 0.05$, the null hypothesis H_0 will be accepted if the statistical measure "z" falls within the interval from –1.959964 to 1.959964. In this context, "z" represents the standardized value of the test statistic used to test the null hypothesis. If the "z" value falls within this interval, it will be considered that the results do not provide sufficient statistical basis to reject the null hypothesis at the 0.05 significance level.

According to the Mann-Whitney test (Tab. 4), among the primary colors of subtractive synthesis, a significant deviation was present only for the yellow primary stimulus. In the tested combinations of primary stimulus colors, surface coverage, and paper structure, the research findings indicate that the smallest effect of simultaneous contrast was observed at the edge values of surface coverage, specifically at 20% and 80% coverage. On the other hand, the greatest differences were observed at the mid-range values of surface coverage, specifically at 40% and 60% coverage.

Table 4 Mann Whitney test results

Structure	%	Primary Stimulus	Secondary Stimulus	Mann Whitney	
				Z	p
"Fili"	20	C	B/G	1.432137	0,1556
		M	R/B	1.089024	0,2825
		Y	R/G	1.181454	0,2431
"Linea"		C	B/G	1.210160	0,2316
		M	R/B	0.397615	0,6991
		Y	R/G	0.367458	0,7251
"Fili"	40	C	B/G	1.618924	0,1074
		M	R/B	0.938259	0,3555
		Y	R/G	2.499181	0,0116
"Linea"		C	B/G	2.049117	0,0402
		M	R/B	1.447711	0,1511
		Y	R/G	2.630522	0,0077
"Fili"	60	C	B/G	0.573402	0,5749
		M	R/B	0.747816	0,4630
		Y	R/G	2.389955	0,0160
"Linea"		C	B/G	0.176866	0,8665
		M	R/B	0.993640	0,3276
		Y	R/G	1.317129	0,1923
"Fili"	80	C	B/G	0.474334	0,6437
		M	R/B	1.955048	0,0507
		Y	R/G	0.108552	0,9199
"Linea"		C	B/G	1.273207	0,2078
		M	R/B	0.951451	0,3488
		Y	R/G	0.379593	0,7123

Table 5 ΔE_{00} Primary Stimulus Cyan Median on Secondary Stimulus Blue

ΔE_{00} Cyan Median / Blue				
%	20	40	60	80
"Fili"	7,7	7,5	3,2	3,1
"Linea"	8,3	8	4,5	4,5

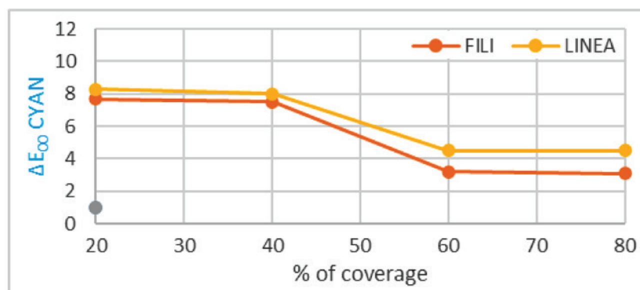


Figure 4 Colorimetric Difference ΔE_{00} for Primary Stimulus Cyan Color and Secondary Stimulus Blue Color on the Left

Based on the obtained results, the calculation of median values was performed to describe and correspond to the perception of individual test cards (left and right) as seen and observed by the CIE standard observer. The deviation in perception caused by the manifestation of the psychophysical visual effect of chromatic assimilation was expressed by the total color difference ΔE_{00} (Tabs. 5-10) and are graphically represented (Figs. 4-9) and commented according to following criteria from the perspective of a standard observer: A ΔE_{00} value of less than 1 indicates a not noticeable difference. A ΔE_{00} value between 1 and 2 indicates a very slight difference, noticeable only to an experienced observer. A ΔE_{00} value between 2 and 3.5 indicates a moderate difference, noticeable even to an inexperienced observer. A ΔE_{00} value between 3.5 and 5 indicates a significant difference. A ΔE_{00} value greater than 6 indicates a very significant difference [13].

Table 6 ΔE_{00} Primary Stimulus Cyan Median on Secondary Stimulus Green

ΔE_{00} Cyan Median / Green				
%	20	40	60	80
"Fili"	6,6	4,3	3,2	3,7
"Linea"	8,3	5,1	4,4	4,5

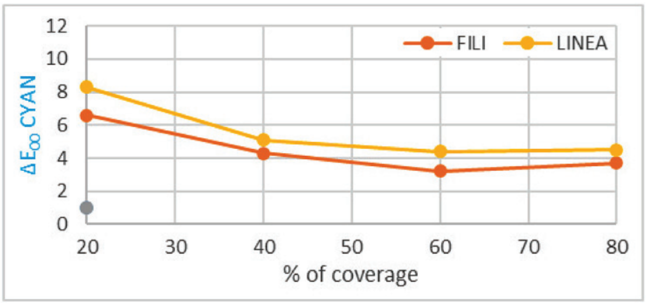


Figure 5 Colorimetric Difference ΔE_{00} for Primary Stimulus Cyan Color and Secondary Stimulus Green Color on the Right

Table 7 ΔE_{00} Primary Stimulus Magenta Median on Secondary Stimulus Red

ΔE_{00} Magenta Median / Red				
%	20	40	60	80
"Fili"	4,8	3,2	4,2	2,7
"Linea"	4,1	4	3,4	3,1

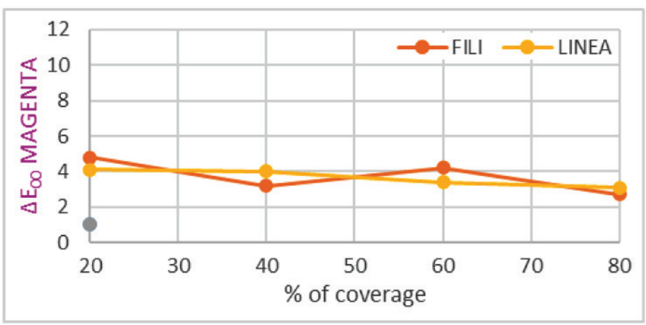


Figure 6 Colorimetric Difference ΔE_{00} for Primary Stimulus Magenta Color and Secondary Stimulus Red Color on the Left

The ΔE_{00} values for the cyan primary stimulus (Tabs. 5-6) show a general decreasing trend as the percentage of coverage increases. At 20% coverage, both "Fili" and "Linea" have similar ΔE_{00} values, approximately 7.7 and 8.3, respectively. As the coverage increases to 40%, the values decrease slightly to around 7.5 for "Fili" and 8.0 for "Linea". A significant drop in ΔE_{00} values is observed at 60% coverage, with "Fili" showing a much lower value compared to "Linea". At 80% coverage, the ΔE_{00} values stabilize around 3.1 for "Fili" and 4.5 for "Linea". The greatest difference in ΔE_{00} values is observed at 60% coverage, where "Fili" demonstrates a notably lower value than "Linea".

Table 8 ΔE_{00} Primary Stimulus Magenta Median on Secondary Stimulus Blue

ΔE_{00} Magenta Median / Blue				
%	20	40	60	80
"Fili"	5	4,3	3,6	4,9
"Linea"	4,9	4,3	5,3	3,8

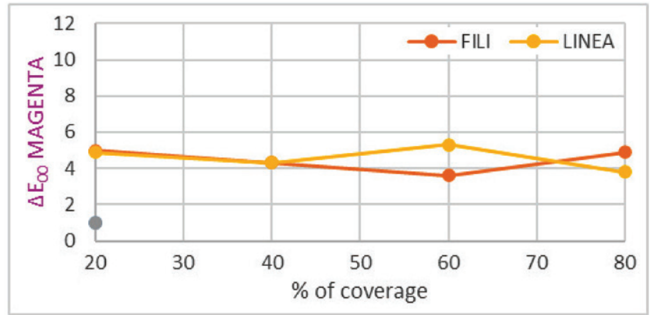


Figure 7 Colorimetric Difference ΔE_{00} for Primary Stimulus Magenta Color and Secondary Stimulus Blue Color on the Right

The ΔE_{00} values for the magenta primary stimulus (Tabs. 7-8) exhibit minimal variation across different coverages. At 20% coverage, both "Fili" and "Linea" show similar ΔE_{00} values around 4.0. This consistency persists across 40% and 60% coverage, with only slight fluctuations observed. By 80% coverage, the values stabilize again for both "Fili" and "Linea". There are minimal differences across all coverage percentages, indicating a stable chromatic assimilation effect for the magenta primary stimulus.

Table 9 ΔE_{00} Primary Stimulus Yellow Median on Secondary Stimulus Red

ΔE_{00} Yellow Median / Red				
%	20	40	60	80
"Fili"	8,2	5,5	3,9	3
"Linea"	9,4	5,7	2,6	2,6

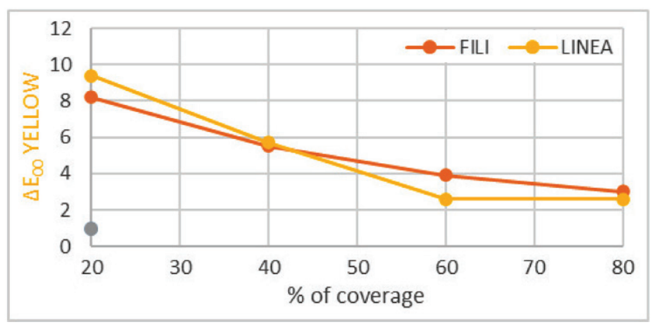


Figure 8 Colorimetric Difference ΔE_{00} for Primary Stimulus Yellow Color and Secondary Stimulus Red Color on the Left

Table 10 ΔE_{00} Primary Stimulus Yellow Median on Secondary Stimulus Gren

ΔE_{00} Yellow Median / Green				
%	20	40	60	80
"Fili"	5,3	3,5	3	3,2
"Linea"	9,4	4,1	2,4	2,8

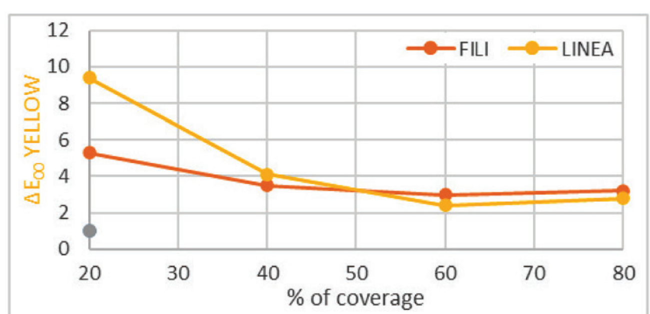


Figure 9 Colorimetric Difference ΔE_{00} for Primary Stimulus Yellow Color and Secondary Stimulus Green Color on the Right

The ΔE_{00} values for the yellow primary stimulus (Tabs. 9-10) show a distinct trend depending on the percentage of RTV coverage. At 20% coverage, there is a notable difference between the two structures, with "Fili" having a ΔE_{00} value of 5.3 and "Linea" having a significantly higher value of 9.4. As the coverage increases to 40%, both structures exhibit a decrease in ΔE_{00} values; "Fili" drops to 3.5, and "Linea" decreases to 4.1. This decreasing trend continues at 60% coverage, where "Fili" shows a ΔE_{00} value of 3.0 and "Linea" further drops to 2.4. At 80% coverage, the values slightly increase for "Fili" to 3.2, while "Linea" shows a slight increase to 2.8. The greatest difference in ΔE_{00} values is observed at 20% RTV coverage, where "Linea" exhibits a significantly higher value compared to "Fili." As the coverage percentage increases, the difference between the two structures diminishes, with "Linea" showing lower ΔE_{00} values than "Fili" at 60% and 80% coverage. The graphs depict the variation in ΔE_{00} values for different primary stimuli (Cyan, Magenta, Yellow) under various coverage percentages for two paper structures, "Fili" and "Linea".

4 CONCLUSION

The key observations are as follows: For the cyan primary stimulus, the greatest ΔE_{00} difference is at 60% RTV coverage, with "Fili" showing a significantly lower value compared to "Linea". For the magenta primary stimulus, minimal variation is observed across all coverage percentages, indicating stable chromatic assimilation effects. For the yellow primary stimulus, the greatest ΔE_{00} difference is at 20% coverage, with both "Fili" and "Linea" showing high initial values that decrease with increased coverage.

These findings highlight the importance of considering paper structure and coverage percentages in printing applications to manage and predict chromatic assimilation effects. The results are significant for optimizing color accuracy and consistency in graphic reproduction, especially for specific color stimuli and coverage levels.

5 REFERENCES

[1] Hong, S. W. & Shevell, S. K. (2004). Brightness contrast and assimilation from patterned inducing backgrounds. *Vision Research, 44*(1), 35-43. https://doi.org/10.1016/j.visres.2003.07.010

[2] Spitzer, H., & Barkan, Y. (2005). Computational adaptation model and its predictions for color induction of first and second orders. *Vision Research, 45*(27), 3323-3342. https://doi.org/10.1016/j.visres.2005.08.002

[3] Motoyoshi, I., Nishida, S., Sharan, L. & Adelson, E. H. (2007). Image statistics and the perception of surface qualities. *Nature, 447*(7141), 206-209. https://doi.org/10.1038/nature05724

[4] Wolfe, J. M. & O'Connell, K. M. (1986). Fatigue and structural change: Two consequences of visual pattern adaptation. *Investigative Ophthalmology & Visual Science, 27*(4), 538-543.

[5] Munker, H. (1970). Farbige Gitter: Abbildung auf der Netzhaut und übertragungstheoretische Beschreibung der Farbwahrnehmung. (in German)

[6] Betz, T., Shapley, R., Wichmann, F. A. & Maertens, M. (2015). Testing the role of luminance edges in White's illusion with contour adaptation. *Journal of Vision, 15*(11), 14-14. https://doi.org/10.1167/15.11.14

[7] Milković, M., Mrvac, N. & Bolanča, S. (2008). Evaluation of the chromatic induction intensity on Munker-White samples. *DAAAM International Scientific Book 2008*, Chapter 41, 485-498.

[8] Helson, H. (1963). Studies of anomalous contrast and assimilation. *JOSA, 53*(1), 179-184. https://doi.org/10.1364/JOSA.53.000179

[9] Chen, C.-C., Chien, S. H.-L. & Lin, Y.-J. (2010). The Munker–White effect and chromatic induction share similar nonlinear response properties. *Seeing and Perceiving, 23*(3), 223-240. https://doi.org/10.1163/187847510X516395

[10] Bressan, P. & Actis-Grosso, R. (2006). Simultaneous lightness contrast on plain and articulated surrounds. *Perception, 35*(4), 445-452. https://doi.org/10.1068/p5247

[11] White, M. (1979). A new effect of pattern on perceived lightness. *Perception, 8*(4), 413-416. https://doi.org/10.1068/p080413

[12] Mann, H. B. & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 50-60. https://doi.org/10.1214/aoms/1177730491

[13] Schläpfer, K. (1993). *Farbmetrik in der Reproduktionstechnik und im Mehrfarbendruck*. UGRA3. (in German)

Authors' contacts:

Katja Milković, PhD
Jurja Dobrila University of Pula,
Zagrebačka 30, 52100 Pula, Croatia
katja.milkovic@unipu.hr

Damir Vusić, PhD, Full Professor
University North,
104. brigade 3, 42000 Varaždin, Croatia
damir.vusic@unin.hr

Martina Hajdek, PhD
Faculty of Graphic Arts, University of Zagreb,
Getaldićeva ulica 2, 10000 Zagreb, Croatia
martina.hajdek@grf.unizg.hr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20230712124048

Preliminary communication
Received: 2023-07-12, Accepted: 2023-08-10

Fuel Consumption of the Tractor-Machine Aggregate Conditioned by the Navigation Systems

Željko Barač*, Denis Ćosić, Tomislav Jurić, Ivan Plaščak, Mislav Jurić, Atila Bezdan, Antonija Kojić

Abstract: This paper presents a study regarding fuel consumption, distance traveled and aggregate operation time during agrotechnical harrowing operation with and without a navigation system in the company. The measurement was carried out on an area of 32.41 ha, which was divided into two equal parts (16.205 ha). The researched area was divided into two equal parts using SMS Advance Software. The tractor-machine aggregate consisted of a Claas Axion 830 tractor as the driving part and a harrow as the working part of it. Fuel consumption was monitored via the tractor's display and the volumetric method. A significant difference in fuel consumption was determined for the volumetric method with navigation (9.97%) and without navigation (6.45%). Higher fuel costs were determined when the unit was operated without a navigation system.

Keywords: agrotechnical operation; fuel consumption; harrow; navigation system; tractor

1 INTRODUCTION

Modern agricultural production, in addition to the sufficiency of food production, strives to achieve its competitiveness on the world market by using seeds with high genetic potential, sophisticated machines with the application of modern technologies and scientific achievements. The reduction of input production costs is conditioned by the specifics of the production of an individual crop, the area where it is grown, climatic conditions, etc.

Among other things, the cost of fuel required for the operation of agricultural aggregates has a significant influence on the size of production costs. Therefore, in order to reduce fuel consumption, an attempt was made to introduce new production technologies that reduced the use of agricultural aggregates (reduced processing, direct sowing, etc.).

Tractor-machine aggregates were operated by operators whose quality of work depended on their work experience, overlapping of passages, and ultimately fuel consumption and utilization of working time. This problem is particularly pronounced during work with tractor-machine aggregates of large scope, where the operator is exposed to greater loads, especially in unfavorable working conditions (fog, night work, dust), notes the author [1].

Aggregating is understood as combining the driving and working machine into a harmonious whole, and tractor aggregates are mostly used in agriculture. The coefficient of utilization of the working capacity of machines that do not work in rows and do not have a marker (harrowing, discing, etc.) ranges from 0.90 to 0.96, according to the authors [2]. Therefore, by applying new technologies in production, the influence of the operator's personal abilities and other factors was minimized.

Regular maintenance of agricultural machinery enable timely and quality work with minimal fuel consumption and favorable working conditions with regard to mechanical vibrations, noise and other factors, according to the authors [3].

The author [4] states that on smaller farms where they use working machines with a working reach of three meters (cultivator, grubber), working hours are reduced by up to 15.7 % with fuel savings of up to 8.66 %, and on farms where they use machines with a working reach of six meters, working time is reduced up to 12.6 % and achieves fuel savings of up to 8.28 %. The reason for the reduction in working time and fuel consumption is the smaller overlap of the passages, which with manual navigation amounts to 6.5 to 9.5 % of the entire work that needs to be done. The authors [5] conducted a survey on 55 out of 100 respondents who used autopilot guidance systems in Turkey. The results indicate time savings (80 %), fuel (80 %) and labor savings (50.09 %). The authors [6] state that by applying an automatic guidance system, fuel consumption savings of 8 % can be achieved, that is, from 3 to 7 % of total energy consumption.

The authors [7] state that the use of navigation systems can contribute to the reduction of fuel consumption during the operation of agricultural machinery.

The use of a navigation system in the guidance of aggregates is one of the possibilities of high-quality performance of work with minimal overlapping of passages, better use of working time, reduction of fuel consumption, less soil compaction and less burden on the operator. Investigating the application of navigation systems and automatic steering systems on North Dakota farms, the authors [8] state that the average respondent saves 65 hours of machine operator time, i.e. reduces fuel consumption by 1.647 liters (which is about 6.3 % of fuel consumption) when using the navigation system. Using the automatic steering system can save an average of 75 hours of machine operator time and 1.866 liters of fuel (5.33 % of fuel consumption). The average fuel savings per farm for the navigation system is $1.305 and for the automatic steering the savings are $1.479.

In a study of the economic profitability of applying the technology of precision agriculture and Controlled Traffic Farming in Denmark in the production of four main crops, it was estimated that the Danish gross domestic product will increase by €34 million with the application of the mentioned technologies, according to the authors [9]. The authors [10] conducted research on a total of 2.200 ha during the application of fertilizers with a navigation system. Potential fuel savings ranged from 0.41 to 0.53 l/ha at a fuel price of 0.86 euros/l.

The aim of this research was to determine the difference in fuel consumption, distance traveled and aggregate operation time during harrowing with and without the use of the tractor navigation system.

2 MATERIALS AND METHODS

The research was carried out through February 23, 2023. in the production areas when closing the winter furrow. The investigated area was 32.41 ha (Fig. 1) and was divided during the measurement into two equal parts (16.205 ha). The investigated area was divided into two equal parts using SMS Advance Software. According to the type, the mentioned plot belongs to coherent soils, the composition of which is the same. A Claas Axion 830 tractor (Fig. 2) with double tires was used during the measurement in combination with a heavy Peck harrow with a working reach of 7 m.

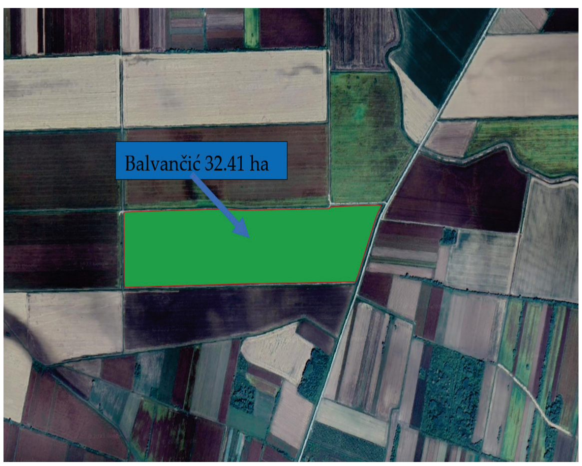


Figure 1 The production area where the research was carried out

Figure 2 The tractor-machined aggregate used to perform the agrotechnical operation of closing the furrow

The closing of the furrow of the first half of the plot was done using Trimble Autopilot navigation with the correction of the RTK of the signal ± 2.5 cm, and the closing of the furrow of the second half of the plot was done without the use of navigation. During the measurement, the following parameters were monitored: working time, distance traveled and total fuel consumption during operation (read on the tractor display and by volumetric method). The volumetric method was carried out by bringing the tractor to a horizontal position and pouring fuel up to a certain mark in the tank before starting work. After that, the tractor with the harrow performed an agrotechnical operation and again brought itself to a horizontal position, and the tank was filled up to the fuel mark. This research is part of a future larger study with regard to fuel consumption, working time and distance traveled by the aggregate when using the navigation system and without using the navigation system.

3 RESULTS

The research results indicate differences in the working time spent (38 minutes) when driving without navigation compared to driving with navigation. Tillage without navigation increases working time consumption by 24.67 % compared to tillage with navigation. It is to be assumed that the greater distance traveled (33.97 km) during operation without navigation is a consequence of the overlap of the passages. The difference in the distance traveled when using navigation compared to working without navigation is (27.51 %), (Fig. 3 and Fig. 4).

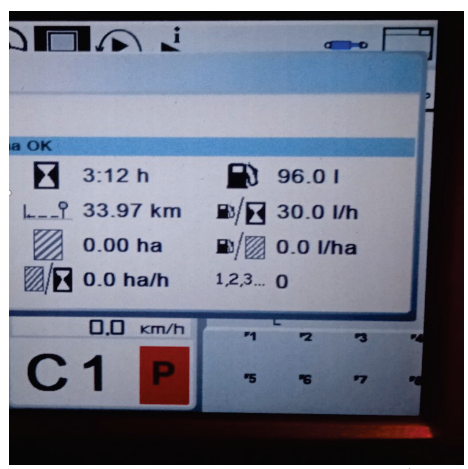


Figure 3 Measured research values - without the use of navigation

The fuel consumption read on the tractor display when working with navigation was 81.2 l, and when working without navigation it was 96.0 l, which is 18.22 % higher consumption than when using navigation. At the same time, during the research, the amount of fuel consumed was checked using the volumetric method, which determined the fuel consumption when working with the navigation system was 89.3 l. Without the navigation system the fuel consumption was 102.2 l, which is 14.44 % higher fuel consumption than when using the navigation system.

Figure 4 Measured research values - with the use of navigation

The results indicate a difference in the consumed fuel by 9.97 % when read on the tractor display and by 6.54 % when calculated by the volumetric method between using the navigation and not using the navigation.

Table 1 Total fuel consumption

Type of work	Fuel cost - display	Fuel cost - volumetric method
With navigation	72.26 €	79.47 €
Without navigation	85.44 €	90.95 €

At the time of the research, the price of blue diesel fuel was 0.89 euros [11], a new regulation on determining the highest retail prices of petroleum products was adopted by the Ministry of Economy and Sustainable Development in Croatia. The fuel costs are shown in Tab. 1. The fuel consumption read on the display per hectare of cultivated area with navigation is 5.01 l and 5.92 l without navigation. Fuel consumption per hectare of cultivated area determined by the volumetric method for working with navigation is 5.51 l and 6.30 l without navigation. The cost of fuel per hectare read on the display with navigation is 4.45 euros and 5.26 euros without navigation. Furthermore, the cost of fuel per hectare determined by the volumetric method when working with navigation is 4.90 euros and 5.60 euros without navigation.

4 DISCUSSION

The higher consumption of working time (24.67 %) during operation without navigation indicates that the aggregate traveled a significantly greater distance (27.51 %) than when using navigation, which, in addition to higher fuel consumption and greater overlap of the passage, will also affect the reduction of the coefficient of utilization of working time, i.e. aggregate effect. All of the above factors depend largely on the operator's work experience and ability. Aggregate management without navigation on large areas is problematic for two reasons: the first is related to the correct measurement of the starting width and the correct first pass of the aggregate, and the second is related to the quality connection of the passes, according to the author [1]. The same problem is pointed out by the authors [12], who state that during the harrowing of 100 ha, the overlap of passages without navigation was found to be 7.04 ha, and 4.94 ha with navigation. Similar results are obtained by the authors [13, 14], when investigating the overlap of passages with and without navigation during spraying, where they determine that the overlap is reduced by using navigation. The authors [15] report fuel and time savings for machines with a working width of 3 m (8.66 % and 15.7 %) and 6 m (8.28 % and 12.6 %) with and without a navigation system, which is similar to the results obtained in this research.

The use of the navigation system resulted in a reduction of fuel consumption by 18.22 % compared to operation without a navigation system, where 5.01 l/ha were used with navigation and 5.92 l/ha without navigation. Fuel costs per hectare of cultivated area are 4.45 €/ha with navigation and 5.26 €/ha without navigation, which is significantly more than 0.93 €/ha reported by the authors [16]. Volumetrically determined fuel consumption with navigation is 5.51 l/ha, and without navigation it is 6.30 l/ha, which gives a cost of 4.91 €/ha and 5.60 €/ha.

A difference in the consumed fuel read on the tractor display was found to be 9.97 % when using navigation and 6.45 % without using navigation, considering the amount of fuel consumed using the volumetric method. These values are significantly higher than the 4.69 % difference stated by the author [17] when measuring the average fuel consumption read from the trip computer with regard to the volumetric method.

Significantly higher fuel consumption determined by the volumetric method in both variants indicates the need for further research into this problem in order to determine the factors that caused it.

5 CONCLUSIONS

Based on the research results, the following conclusions can be drawn:

- higher fuel consumption was found when harrowing without a navigation system,
- a difference was found in the fuel consumption read on the tractor display and the consumption determined by the volumetric method,
- during operation without a navigation system, the aggregate travels a longer distance, which increases the working time,
- higher fuel costs were determined when working without navigation when reading the fuel on the tractor display and when using the volumetric method.

6 REFERENCES

[1] Kostić, M. M. (2021). *Precizna poljoprivreda*. Poljoprivredni fakultet Novi Sad, Novi Sad, Serbia. (in Serbian)

[2] Brkić, D., Vujčić, M., Šumanovac, L., Kiš, D., Jurić, T. & Knežević, D. (2005). *Eksploatacija poljoprivrednih strojeva*. Fakultet agrobiotehničkih znanosti Osijek, Osijek, Croatia. (in Croatian)

[3] Barač, Ž., Plaščak, I., Jurić, T., Baličević, P., Duvnjak, V., Jurišić, M., Heffer, G. & Marković, M. (2019). Influence of an uneven surface on the vibration occurrence affecting the tractor operator. *Tehnički glasnik, 13*(3), 192-196.

https://doi.org/10.31803/tg-20190211142404
[4] Stanjko, D. (2018). The possibility of increasing the efficiency of soil preparation and winter grain sowing using precision agriculture technology. *Journal of Plant Protection, 5*, 20-27. https://doi.org/10.1027/0269-8803.20.4.253
[5] Keskin, M., Sekerli, Y. E., Say, S. M. & Topcueri, M. (2018). The possibility of increasing the efficiency of soil preparation and winter grain sowing using precision agriculture technology. *Journal of Agricultural Faculty of Gaziosmanpasa University, 35*(2), 172-181. https://doi:10.13002/jafag4421
[6] Rodias, E., Berruto, R., Busato, P., Bochtis, D., Sorensen, C. G. & Zhou, K. (2017). Energy Savings from Optimised In-Field Route Planning for Agricultural Machinery. *Sustainability, 9*, 1956. https://doi.org/10.3390/su9111956
[7] Savickas, D., Steponavičius, D., Kliopova, I. & Saldukaitė, L. (2020). Combine Harvester Fuel Consumption and Air Pollution Reduction. *Water, Air, & Soil Pollution, 231*, 95. https://link.springer.com/article/10.1007/s11270-020-4466-5
[8] Bora, G. C., Nowatzki, J. F. & Roberts, D. C. (2012). Energy savings by adopting precision agriculture in rural USA. *Energy, Sustainability and Society, 2*, 22. https://energsustainsoc.biomedcentral.com/articles/10.1186/2192-0567-2-22
[9] Jensen, H. G., Jacobsen, L. B., Pedersen, S. M. & Tavella, E. (2012). Socioeconomic impact of widespread adoption of precision farming and controlled traffic systems in Denmark. *Journal of Agricultural Faculty of Gaziosmanpasa University, 13*, 661-677. https://doi.org/10.1007/s11119-012-9276-3
[10] Shannon, D. K. & Charles, E. E. (2012). Evaluating GPS Guidance Technologies for Energy Savings. *Proceedings 2012 ASABE Annual International Meeting Sponsored by ASABE Hilton Anatole, Dallas, Texas,* 121338140. Retrieved from https://elibrary.asabe.org/abstract.asp?aid=41879
[11] A new Regulation on determining the highest retail prices of petroleum products was adopted. (2023, June 2). Retrieved from https://mingor.gov.hr/vijesti/donesena-nova-uredba-o-utvrdjivanju-najvisih-maloprodajnih-cijena-naftnih-derivata-9149/9149
[12] Jurišić, M., Plaščak, I., Zimmer, D., Barač, Ž. & Rapčan, I. (2018). Modern Navigation System DGPS as a Function of Saving in Agriculture. *Tehnički vjesnik, 4*(25), 1067-1074. https://doi.org/10.17559/TV-20170223094621
[13] Topçueri, M. & Keskin, M. (2019). Effectiveness of GNSS-based tractor auto steering systems in crop spraying. *Mustafa Kemal University Journal of Agricultural Sciences, 24*, 78-90. https://dergipark.org.tr/en/download/article-file/904047
[14] Esau, T. J., MacEachern, C. B., Farooque, A. A. & Zaman, Q. U. (2021). Evaluation of Autosteer in Rough Terrain at Low Ground Speed for Commercial Wild Blueberry Harvesting. *Agronomy, 11*(2), 384. https://doi.org/10.3390/agronomy11020384
[15] Kelc, D., Stajnko, D., Berk, P., Rakun, J., Vindiš, P. & Lakota, M. (2019). Reduction of environmental pollution by using RTK-navigation in soil cultivation. *International Journal of Agricultural and Biological Engineering, 5*, 173-178. https://doi.org/10.25165/j.ijabe.20191205.4932
[16] Marković, D., Pokrajac, S., Simonović, V. & Marković, I. (2013). Economic Evaluation of GPS Technology in Serbian Agriculture. *Business School, 3-4*, 1-11. https://doi.org/10.5937/skolbiz1304001M
[17] Benković, R. (2022). Eksploatacijski pokazatelji sredstava poljoprivredne mehanizacije i tvorba prinosa pri različitim sustavima obrade tla. *PhD Thesis*, Fakultet agrobiotehničkih znanosti Osijek, Osijek, Croatia. (in Croatian)

Authors' contacts:

Željko Barač, PhD, Assistant Professor
(Corresponding author)
Faculty of Agrobiotechnical Sciences Osijek,
Vladimira Preloga 1, 31000 Osijek, Croatia
zbarac@fazos.hr

Denis Ćosić, MSc
Jerković d.o.o.,
Josipa Kozarca 79, 31224 Koška, Croatia
denis.cosic10@gmail.com

Tomislav Jurić, PhD, Full Professor
Faculty of Agrobiotechnical Sciences Osijek,
Vladimira Preloga 1, 31000 Osijek, Croatia
tjuric@fazos.hr

Ivan Plaščak, PhD, Associate Professor
Faculty of Agrobiotechnical Sciences Osijek
Vladimira Preloga 1, 31000 Osijek, Croatia
iplascak@fazos.hr

Mislav Jurić, MSc
Independent researcher
Hrvatskih kraljeva 110, 32100 Vinkovci, Croatia
mislav.juric01@gmail.com

Atila Bezdan, PhD, Associate Professor
Faculty of Agriculture,
Trg Dositeja Obradovića 8, 21000 Novi Sad, Serbia
atila.bezdan@polj.edu.rs

Antonija Kojić, MSc, Assistant
Faculty of Agrobiotechnical Sciences Osijek,
Vladimira Preloga 1, 31000 Osijek, Croatia
akojic@fazos.hr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240430191724

Preliminary communication
Received: 2024-04-30, Accepted: 2024-05-08

Possibilities of Automating the Additive Manufacturing Process of Material Extrusion – MEX

Grzegorz Budzik*, Mateusz Przytuła, Andrzej Paszkiewicz, Mariusz Cygnar, Łukasz Przeszłowski, Tomasz Dziubek

Abstract: The article presents the possibilities of automating production pre-processing and post-processing operations for the Material Extrusion - MEX process based on Fused Filament Fabrication technology. Automation is based on hardware and software solutions. For this purpose, a special research station was developed, equipped with a warehouse of working platforms, a 3D printer and a collaborative robot that integrates individual elements of the manufacturing process. The developed solution allows for increasing the efficiency of the manufacturing cell and reducing the operator's involvement in manual operations at the pre-processing and post-processing stages.

Keywords: 3D Printing; Collaborative Robot; Material Extrusion; Rapid Manufacturing

1 INTRODUCTION

Manufacturing products using additive techniques involves the need to prepare the 3D printer for operation (pre-process operations) and remove the model from the 3D printer and clean the working space (post-process operations) with human participation. For this reason, in the case of large-scale production, a more economical solution is to use traditional manufacturing methods, e.g. material injection. One way to reduce production costs using additive techniques may be to automate pre- and post-process operations, enabling it to be carried out continuously.

Automation of additive manufacturing processes also allows 3D printing to be moved into the era of Industry 4.0/5.0, where one of the main assumptions is to minimize human physical work as well as control and measurement operations [1, 2]. The use of artificial intelligence to monitor the process can significantly reduce the operator's working time. Additionally, the introduction of collaborative robots may result in the physical presence of the operator at the production station being required only to arm it and prepare the machines for the process, and further activities will be performed remotely [3, 4].

Current trends in the development of 3D printing show that the topic of automation of additive manufacturing processes has recently become very popular. Manufacturers of 3D printing machines present both prototypes and serial solutions of scalable production systems. However, these are solutions dedicated to specific devices from a given manufacturer. Currently, both on the market and in the equipment of many companies, you can find a large number of 3D printers with very good production parameters, but not having the characteristics of automated systems. Therefore, it is reasonable to develop a methodology for integrating such machines into networked, automated production units.

The MEX (Material Extrusion) method of layered extrusion from thermoplastic polymer materials involves extruding a plasticized material called a filament, which is most often in the form of a thread wound on a spool. Standard filament diameters are 1.75 and 2.85 mm. The extrusion process takes place in the head, which includes, among others: heating part, extruder and nozzle. The material is delivered to the heating block, where it is plasticized and then forced through the nozzle. The second function of the head is retraction, i.e. the withdrawal of material in order to prevent material leakage during idle movements of the head. A 3D printer using this method usually has one or more heads. Many heads are used, for example, when printing support structures from a material other than the main building material or when using various combinations of construction materials. Depending on the way the print heads move, the following designs can be distinguished: Cartesian, delta, polar or using a robot arm. In the first three cases, it is possible to apply flat layers of material on a leveled work platform. Structures based on a robot arm allow you to build curved layers, eliminating the need to use support structures. Additionally, in solutions of this type, the working platform can be placed on a robotic table with three axes of freedom. There are also solutions in which several robots work on one printout in parallel, shortening the time of the incremental process of a single model [5, 6]. There are also systems in which integrated printing robots move on mobile platforms or are suspended from flying drones. Additive processes are also subject to standardization in terms of terminology, design and production supervision [7, 8].

2 ANALYSIS OF DESIGN SOLUTIONS FOR 3D PRINTERS

The analysis covered 3D printing devices commonly used in scientific research, industrial applications and amateur applications, which are part of the machinery of the Department of Machine Design of the Rzeszów University of Technology. These devices may constitute a set of separate production machines, or it is possible to combine them into one coherent, heterogeneous production environment, which would increase the level of comfort and efficiency of their management. As part of the work carried out, an analysis was undertaken of the possibility of integrating printers into a common network structure with the possibility of automating selected processes. The possibilities of physical automation and integration through network interfaces available in individual devices, communication protocols used and possible ways of connecting to the network were analyzed.

Among the analyzed machines, we can distinguish those that cooperate only with the software of a given manufacturer, using classified communication protocols. An example would be brand devices equipped with a USB interface. They can only be connected to the network structure by providing remote access to a computer connected to the 3D printer via a USB cable. Such solutions are used, among others, in Objet Eden 260 and Objet 350 Connex 3 3D printers. These machines are additionally equipped with a wired LAN network interface enabling connection to a local network [9].

The second group of devices are machines with extensive network integration capabilities, including Stratatys F170 and Ultimaker3 3D printers. Both are equipped with the possibility of wired connection to the network, and Ultimaker3 also has a wireless WiFi interface. In both cases, device manufacturers provided documentation of communication protocols. Stratatys has also prepared a set of programming tools SDK (Software Development Kit) that facilitates the integration of 3D printers with the IT systems of modern, intelligent factories [10].

The third group of 3D printers are devices without factory network interfaces, the integration of which is possible using external print servers. Two types of solutions can be distinguished here: closed and open [11]. The first of them only work with the software of a given supplier (e.g. Repetier or 3DQue). The possibilities of their configuration and customization are very limited. Open solutions, however, allow for a flexible and comprehensive approach to the topic of network integration. This type of solutions includes, for example, Octoprint software. The third group of printers includes the Gence F340 3D device. Although its manufacturer does not provide information about the device's network operation capabilities, as part of the work carried out, a network connection was established using an external Repetier print server. The second example is the Prusa i3 MK3 printers. The Rzeszów University of Technology has several dozen such devices in the Rapid Prototyping Laboratory (Fig. 1). Printers are used for research and educational purposes.

Figure 1 Research and teaching Rapid Prototyping Laboratory

Prusa i3 MK3 3D printers can be extended with network functionality in several ways. One of them, used in the Rapid Prototyping Laboratory, involves performing a hardware modification. To perform it, you need a miniature single-board computer Raspberry Pi Zero and installing the Octoprint software. The assembly elements were made using 3D printing (Fig. 2). The electronic system should be connected directly to the serial port connector on the 3D printer controller board. This solution does not require an additional power supply.

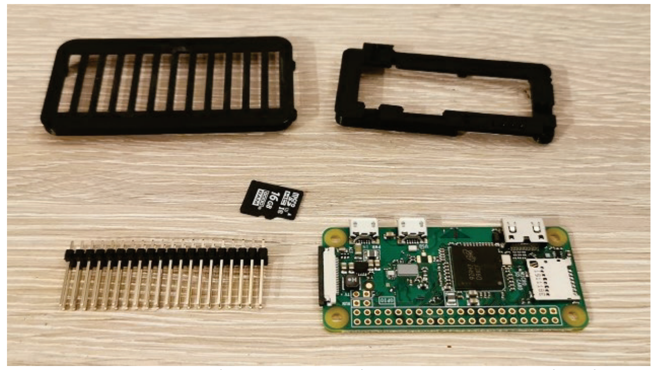


Figure 2 Elements used for hardware modification to extend Prusa i3 MK3 with network functionality

A computer of this class has relatively low computing power. In this case, software developers do not recommend installing more additional plug-ins and extensions. However, during the tests, the OctoEverywhere plug-in was installed, enabling remote access from anywhere, and during printing, no negative impact on the 3D printing processes was found. Octoprint software can also be installed on a much more powerful single-board computer, e.g. Raspberry Pi 4. However, in this case, a USB cable is used to connect to the printer. Additionally, it requires an additional power supply. This method was used during the construction of a laboratory research stand for the presented solution.

Video monitoring is an extremely important element of automatic and network-capable additive manufacturing systems. Its implementation allows for remote viewing of printing processes from anywhere. This saves the time needed for the operator to approach individual production machines to be visually inspected the correctness of its progress. Additionally, the implementation of artificial intelligence algorithms for real-time defect detection allows for quick response, which consequently reduces production time and material consumption.

3D printers are rarely equipped with video cameras as standard, and cameras are also rarely used as optional equipment. The camera is most often installed by the user himself. Often, the camera is connected to a previously integrated print server. Depending on its type, the connection is made via a universal USB interface or another specialized interface, e.g. CSI (Camera Serial Interface).

3D printer directed at their worktables. For this purpose, electronic systems of "open-frame" cameras were used, having a 5Mpx matrix and allowing image recording in 1080p resolution at 30 FPS. The systems were built into an articulated handle manufactured by 3D printing (Fig. 3) and then connected to previously retrofitted print servers. This solution enabled remote viewing of the process via the OctoEverywhere cloud platform and real-time supervision of the correctness of printing by the artificial intelligence algorithm - Gadget. A similar solution was used during the construction of the research station. However, in this case, web cameras connected to the print server via a USB interface were used.

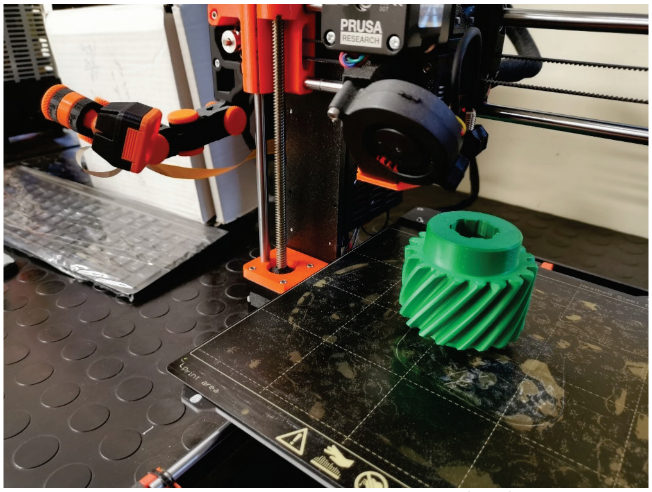


Figure 3 External camera installed on the Prusa i3 MK3 printer

3 STAND FOR TESTING 3D PRINTING AUTOMATION IN THE MEX PROCESS

As part of the research carried out, a concept of 3D printing automation based on the MEX process was developed based on a designed sorting and feeding device intended for 3D printers and applicable to pre- and post-process operations (Fig. 4). The presented concept is related to the use of a completely new approach to manufacturing processes. It involves the integration of various elements necessary to produce products or semi-finished products using additive technologies in a small area. Such a system is equipped with feeding, control, production, storage and control devices, e.g. in the form of 3D scanners and other devices. This solution was created in close cooperation with the socio-economic environment and was submitted for patent protection by PROXIMO AERO Sp. z o. o., application no. P.441182 [WIPO ST 10/C PL441182] of May 14, 2022. The co-authors listed in the patent application are employees of the Rzeszow University of Technology.

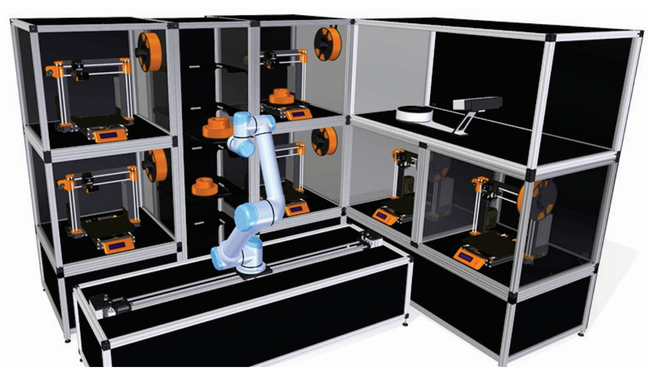


Figure 4 3D-CAD of an automated 3D printing station

The presented system allows for the implementation of works using a collaborative robot as a sorting device and feeding platforms to 3D printers. This solution can be used in a real production environment. The presented solution is based on 3D printers with an open working space, e.g. PRUSA MK3S, thanks to which the robot handle can freely remove the working platform from the printer. To ensure continuity of production, limited by the amount of free resources (filament, work platforms), the solution in question used the Universal-Robots UR5e collaborative robot with an arm reach of 850 mm and a load capacity of 5kg. The robot arm is equipped with a special gripper adapted to work with a specific 3D printer (Fig. 5). The designed solution also enables the station to be expanded with additional segments using a special track on which the robot can be placed.

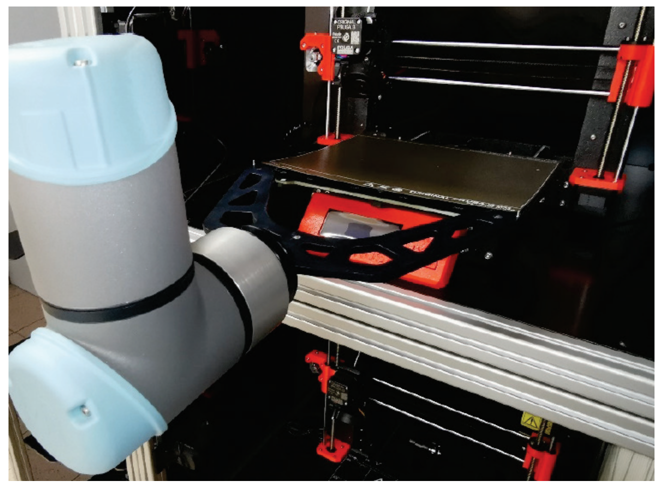


Figure 5 Robot gripper designed for the PRUSA MK3S 3D printer

However, to make this possible, it was necessary to develop a methodology for automating and computerizing the 3D printing process using the layer extrusion process of polymer thermoplastic materials. Within it, two main phases can be distinguished: preparatory and production. The first one aims to summarize the production environment. The phase begins with the preparation of preliminary assumptions regarding the desired production capacities, technological limits and the planned degree of process automation. Subsequent conceptual work concerns determining the type of manufacturing machines planned to be used, quality control systems and the movement of finished products. It is also important to make assumptions for management systems for machines and production materials. The developed concept allows for the selection of machinery. The result of his analysis is a summary of limits regarding possible production materials, maximum dimensions of manufactured objects and their production speed. A view of the material spool holder equipped with a scale and a wireless RFID communication sensor is shown in Fig. 6.

Identification takes place wirelessly when the spool is placed on the holder by reading information from RFID (Radio Frequency Identification) tags in the form of stickers. The read information includes a unique reel ID number allowing it to be uniquely identified in the database. The ID number is programmed once by the operator when the reel is accepted into the production materials warehouse. Reading data from the tag is also possible via most mobile devices. This makes it easier to physically locate a specific reel in the production materials warehouse.

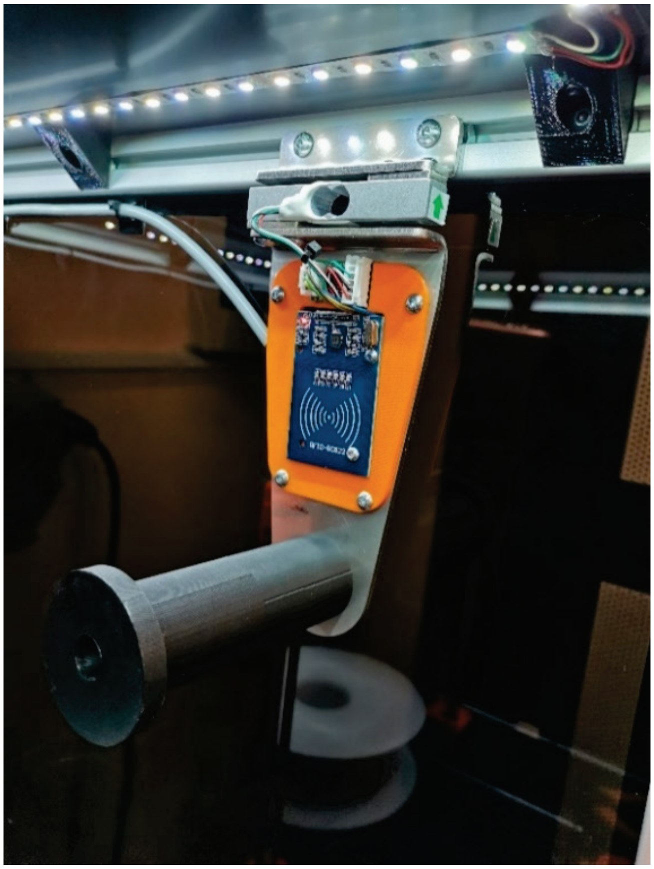


Figure 6 RFID tag reader built into the spool holder

The next step is to carry out network integration of selected devices. This process includes the compilation of network functions and the development and implementation of communication and data transfer protocols, including access authorization methods. One of the key elements of the preparatory phase is the development of an IT supervision system. Correctly designed, it allows for efficient management of machinery, production materials and the manufacturing process. It is important to design a user interface that is user-friendly and effective. Therefore, at this stage it is worth conducting a requirements analysis among potential process operators. The preparatory phase ends with the implementation of automation mechanisms such as quality control systems, optional measurement systems and robotic product transport systems. The manufacturing phase concerns pre- and post-process activities and the actual additive manufacturing process. The phase begins with the operator configuring the manufacturing system using a computer application. The operator selects production files and determines the number of elements to be produced. It is assumed that production files will be prepared in advance using cutting software. After system configuration, resources are automatically allocated to a given process and production files are transferred to appropriate devices via the network. Depending on the level of automation of the production station, the machines are equipped with working platforms by an operator or a robotic system. During the actual production stage, 3D printers and other devices remain under the control of the IT system. Manufactured objects are continuously monitored by artificial intelligence algorithms to detect defects. The operating parameters of 3D printers and the amount of available production material are also monitored. Data is collected in a central database. The operator can visualize them from an application with a graphical interface. According to the adopted approach, two production modes can be distinguished, depending on the degree of process automation. A manufacturing environment with a second degree of automation is characterized by the fact that after production is completed, the physical model together with the working platform is transferred by a robotic arm to the finished products warehouse or to the quality control station. Then the robot equips the 3D printer with an empty platform, and the IT system prints the next object. In a manufacturing environment with the first degree of automation, these activities are made manually by an operator.

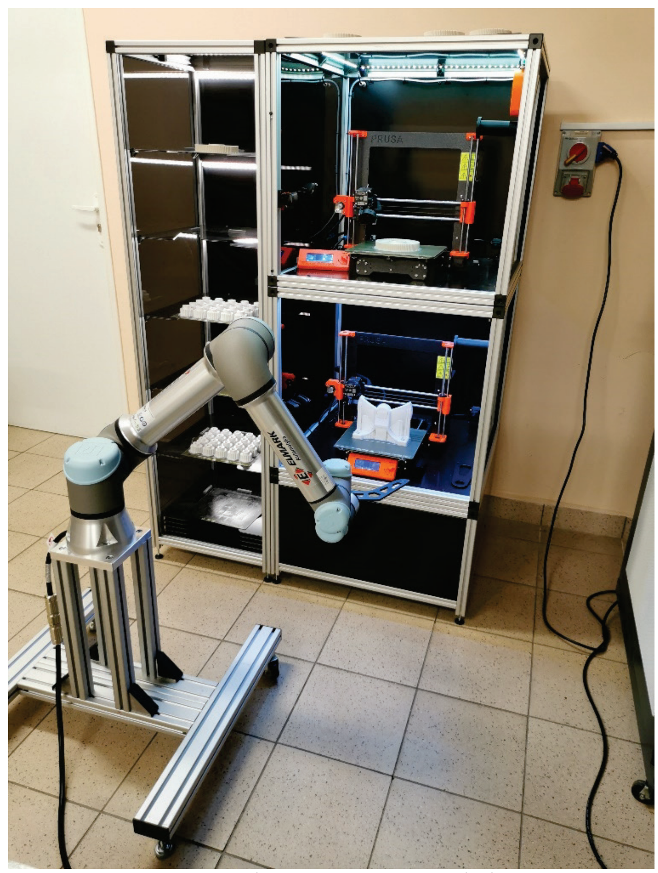


Figure 7 Automated feeding and sorting device for 3D printing

In order to implement the concept of a feeding and sorting station for elements printed in a system of multiple 3D printers, it was necessary to perform work on the analysis of methods of network integration of 3D printers, remote control and monitoring systems of the manufacturing process, mechanisms for controlling the amount of available production material and methods of transferring the manufactured models using collaborative robot. Based on the developed methodology, a research station was built in the form of an automated 3D printing station (Fig. 7) with a monitoring application.

4 CONCLUSIONS

As a result of a series of analyzes and numerous experiments, a methodology was developed for integrating 3D printing systems and processes using network infrastructure, which is the result of achieving the goal of the work. The developed methodology assumes network cooperation of 3D printers, allowing the creation of heterogeneous production environments using layered extrusion technology. On its basis, a research station in the form of an automated 3D printing station with a modular structure was developed and constructed.

The station is operated by the operator remotely via an application from a web browser. The application manages the operation of printing devices and the work of the robot, whose task is to transport the produced models and install clean working platforms. This allows for continuous production until the filament runs out.

As a result of the experimental work carried out, the following conclusions can be drawn:

- The use of network structures enables remote control of the manufacturing process, in particular its starting, stopping and transfer of production files. This allows a printer or set of 3D printers to be armed in advance and ready to remotely start the process on demand when required.
- Transferring production files via a network is less time-consuming compared to conventionally transferring them on a storage medium. This is particularly important in the context of managing large teams of production machines and frequently changing production.
- Early detection of defects occurring during the printing of a physical object allows for reduction of wasted material and 3D printer operating time. This detection can be performed remotely by the operator by periodically inspecting the image from cameras monitoring the process. The operator can observe from one place, which is particularly important in the case of distributed 3D printing systems, in particular its starting, stopping and transfer of production files. This allows a printer or set of 3D printers to be armed in advance and ready to remotely start the process on demand when required.
- Knowing the type and quantity of available production material in a given 3D printer allows for effective management of the machinery. Thanks to this information, the print queuing algorithm implemented in the application that supervises the entire manufacturing environment can automatically redirect the print order to the appropriate 3D printer. This is particularly important in the case of individual or small-scale production to order and in the event of device failure or detection of a print defect.
- Automatic recording of the amount of filament in partially used spools allows for more efficient management of stored materials. Thanks to the knowledge of inventory levels, it is possible to automate certain activities related to logistics. Appropriate algorithms can assess the feasibility of completing a given production order or determine the amount of material needed for the order. It also allows you to quickly estimate the time needed to complete production, which is extremely important information for the ordering party.
- The additional introduction of robotic execution systems allows for the elimination of the operator's participation in activities between subsequent prints. As a result, its presence is only needed to equip the machines with production material.

5 REFERENCES

[1] Barbosa, G. F. & Aroca, R. (2017). An IoT-Based Solution for Control and Monitoring of Additive Manufacturing Processes. *Journal of Powder Metallurgy and Mining, 6*(1), 1000158.

[2] Budzik, G. & Żelechowski, D. (2019). Possibilities of using photopolymer injection molds in small series production in Industry 4.0 structure. *Polimery, 64*(6), 405-410. https://doi.org/10.14314/polimery.2019.6.2

[3] Cui, J., Ren, L., Mai, J., Zheng, P. & Zhang, L. (2022). 3D Printing in the Context of Cloud Manufacturing. *Robotics and Computer-Integrated Manufacturing, 74*(102256). https://doi.org/10.1016/j.rcim.2021.102256

[4] Fuwen, H., Mikolajczyk, T., Pimenov, D. Y. & Gupta, M. K. (2021). Extrusion-based 3D printing of ceramic pastes: Mathematical modelling and in situ shaping retention approach. *Materials, 14*, 1137. https://doi.org/10.3390/MA14051137

[5] Gibson, I., Rosen, D., Stucker, B. & Khorasani, M. (2021). Materials for additive manufacturing. In *Additive Manufacturing Technologies*, Springer: Cham, Switzerland, 379-428. https://doi.org/10.1007/978-3-030-56127-7_14

[6] Wohlers, T. T., Campbell, R. I., Diegel, O., Huff, R., Kowen, J. & Mostow, N. (2021). *Wohlers Report 2021: Additive Manufacturing and 3D Printing State of the Industry*. van Rensburg, J., Ed.; Wohlers Associates, Inc.: Fort Collins, CO, USA.

[7] ISO/ASTM. (2019). ISO/ASTM52921-13: *Standard Terminology for Additive Manufacturing-Coordinate Systems and Test Methodologies*.

[8] Kakade, S., Mulay, A. & Patil, S. (2022). IoT-based real-time online monitoring system for open ware FDM printers. *Materials Today: Proceedings, 67*(2), 363-367. https://doi.org/10.1016/j.matpr.2022.07.210

[9] Magrini, E., Ferraguti, F., Ronga, A. J., Pini, F., Luca, A. & Leali, F. (2020). Human-robot coexistence and interaction in open industrial cells. *Robotics and Computer-Integrated Manufacturing, 61*(101846). https://doi.org/10.1016/j.rcim.2019.101846

[10] Woźniak, J., Budzik, G., Przeszłowski, Ł., Fudali, P., Dziubek, T. & Paszkiewicz, A. (2022). Analysis of the quality of products manufactured with the application of additive manufacturing technologies with the possibility of applying the Industry 4.0 conception. *International Journal for Quality Research, 16*(3), 831-850. https://doi.org/10.24874/IJQR16.03-12

[11] Zhang, L., Luo, X., Ren, L., Mai, J., Pan, F., Zhao, Z. & Li, B. (2020). Cloud based 3D printing service platform for personalized manufacturing. *Science China Information Sciences, 63*, 124201. https://doi.org/10.1007/s11432-018-9942-y

Authors' contacts:

Grzegorz Budzik, Prof. PhD, DSc, Eng.
(Corresponding author)
Academy of Applied Sciences in Nowy Sącz,
ul. Staszica 1, 33–300 Nowy Sącz, Poland
gbudzik@ans-ns.edu.pl
Rzeszów University of Technology
Al. Powstańców Warszawy 12, 35-959 Rzeszów, Poland
gbudzik@prz.edu.pl

Mateusz Przytuła, MSc, Eng.
Rzeszów University of Technology,
Al. Powstańców Warszawy 12, 35-959 Rzeszów, Poland
mprzytula@prz.edu.pl

Andrzej Paszkiewicz, PhD, DSc, Eng. Assistant Prof.
Rzeszów University of Technology,
Al. Powstańców Warszawy 12, 35-959 Rzeszów, Poland
apaszkiewicz@prz.edu.pl

Mariusz Cygnar, Associate Prof. PhD, DSc, Eng.
Academy of Applied Sciences in Nowy Sącz,
ul. Staszica 1, 33–300 Nowy Sącz, Poland
mcygnar@ans-ns.edu.pl

Łukasz Przeszłowski, Assistant Prof. PhD, DSc, Eng.
Rzeszów University of Technology,
Al. Powstańców Warszawy 12, 35-959 Rzeszów, Poland
lprzeszl@prz.edu.pl

Tomasz Dziubek, Associate Prof. PhD, DSc, Eng.
Rzeszów University of Technology,
Al. Powstańców Warszawy 12, 35-959 Rzeszów, Poland
tdziubek@prz.edu.pl

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240502090609

Preliminary communication
Received: 2024-05-02, Accepted: 2024-05-08

Digitalization in Logistics for Competitive Excellence: Case Study of Estonia

Aurelija Burinskienė*, Diana Daškevič

Abstract: Logistics is presently undergoing a transformative phase driven by rapid digitalization. This trend has spurred an in-depth exploration of the industry's theoretical potential and practical challenges. This comprehensive review focuses on practical applications, delving deeply into the increasing interest in digital technologies within logistics. The primary objective of this study is to conduct a bibliometric analysis to unravel the growth and academic development of the concept of logistic digitalization. The investigation examines academic literature from the Web of Science database spanning 2020-2024. Using VOS Viewer, an intellectual structure and bibliographic analysis are performed on selected articles. The cluster that concentrates on articles about the Internet of Things (IoT) technology and its impact on the logistics economy (the fourth cluster) is most closely related to the digital theme. The study combines theoretical framework, including Michael Porter's and the Resource-Based View, to illustrate the connections that improve business competitiveness. The examination of investments in digital technology sheds light on the logistics sector's experiences during the years when these investments impact labor productivity. Descriptive analysis results suggest that companies implementing technology can achieve a workforce productivity increase of up to 20%, underscoring the significance of technology investments for enhancing business competitiveness. These findings shed light on how various logistics sub-sectors respond to investments in software and computers and the rate at which these technologies influence labor productivity metrics. These empirical findings substantially reinforce the theoretical framework, providing practical implications and emphasizing the critical necessity for logistics firms to integrate digital technologies into their operational frameworks seamlessly. In conclusion, as the logistics landscape hovers on the brink of a digital revolution, businesses must adeptly navigate and wholeheartedly embrace digital solutions to ensure competitiveness.

Keywords: digitalization; investments; labor productivity; logistic sector

1 INTRODUCTION

The logistics sector is rapidly transforming through digitalization, becoming crucial in securing competitive advantages in today's dynamic global market. The move to digitalization comes with various advantages, like better access to information, streamlined logistics, real-time data collection, improved inventory management, and increased transparency [1, 2].

In the constantly changing world of logistics, with the rise of Smart Logistics and Smart Supply Chain, technologies like artificial intelligence, blockchain, cloud computing, and the Internet of Things work together to automate processes and planning. This leads to a level of partial autonomy and promotes the sharing of information across organizations [3]. Nevertheless, optimizing the advantages of integrated systems necessitates organizations synchronizing their internal technology implementation with the tempo of external stakeholders. Achieving success in this transformative process requires organizations to allocate resources toward emerging technologies, meticulously oversee the integration of digital tools, and gain a comprehensive understanding of the opportunities and interconnections within logistics processes [4].

This literature review explores the increasing interest in digitalization within logistics, focusing on its profound impact on supply chain management and its integration to enhance competitiveness. Integrating digital technologies into logistics remains pivotal for industry growth [5]. 3PL companies have to provide services to customers. Additionally, retail companies always seek to improve operational efficiency in warehouses due to fierce competition. The outcomes reported by the companies outlined, such as Bltron, Pierre's Ice Cream, and Anacapri Foods, reveal a notable 20% surge in workforce productivity [6]. Warehouse operations, particularly material handling (e.g., storage, picking, etc.), are significantly bolstered by IoT solutions. For instance, "smart" sensors can maintain continuous inventory monitoring, generate replenishment alerts [7], aiming to subsequently reduce operational costs stemming from unmet customer needs or increasing unwanted items. Additionally, autonomous picking and packing robots, as well as collaborative robots working alongside humans, help reduce error rates and increase efficiency [8]. However, these technologies can transform standard processes, make supply chain operations more efficient, and contribute to the success of businesses [9]. Embracing digitalization is vital for logistics firms to ensure they deliver effective and responsive customer services [10, 11].

The study aims to understand how digital transformation is changing the logistics sector and plays a crucial role in gaining a competitive advantage in today's dynamic market [12, 13]. The change propelled by digital technologies brings several benefits, such as improved access to information, streamlined practices, real-time data collection, enhanced inventory management, and greater transparency [14].

The paper aims to bridge theoretical concepts with real-world applications in the logistics sector, exploring the growing significance of digital technologies and their transformative potential. It outlines the empirical study's objective, which focuses on how software and computer investments impact logistics productivity.

This paper aims to address the following research questions:

1) What topics are emphasized in the research paper?
2) What are the specific applications of digital technologies?
3) What is the effect of implementing digital technologies seeking to increase competitive advantage?

The paper is divided into five chapters. It starts with an introduction and a literature review. In the second chapter, bibliometric analysis is conducted, classifying digital technologies used in logistics based on their functional capabilities. This chapter also discusses theories relevant to the application of digital technologies for enhancing logistics competitiveness. In the third chapter, the authors present the research design, methods, materials, and results, which include descriptive, statistical, and regression analysis. Finally, the paper concludes with discussions and conclusions in the fourth and fifth chapters.

2 LITERATURE REVIEW

Analyzing the current and future state of the logistics sector underscores the crucial need to quickly adopt and seamlessly integrate new market demands for a competitive edge. Achieving business efficiency is closely tied to understanding the evolving trends in digital technologies and how they directly impact logistics management operations. [15, 16]. In the current sector, integrating digital technologies becomes a significant challenge. Digitization includes adopting new technologies and converting traditional processes into digital formats [17]. These new digital advancements represent the latest innovations anticipated to significantly reshape corporate strategies and societal environments.

Researchers [18] studied how introducing new digital technologies opens up valuable business opportunities for logistics centers in supply chains. The study's results serve as a basis for proposing alternative strategic options to innovate logistics chains and improve competitiveness.

Studies show that incorporating information technology in a collaborative approach significantly boosts productivity and efficiency in warehouse operations [19]. The impact of information technology on productivity and competitive advantage in logistics is underscored. Productivity, a widely used indicator for performance evaluation [20, 21], remains a key focus. Adopting integrated logistics systems also enables firms to implement lean production methodologies characterized by dependable order cycles and reduced inventory levels [22]. Overall, logistics integration empowers companies and their supply chain partners to operate as a cohesive unit, leading to enhanced performance across the entire chain [23]. Numerous studies have highlighted the manifold logistics advantages of sharing information with supply chain partners, particularly in inventory management [24, 25]. For instance, Vendor-Managed Inventory (VMI) integration with suppliers has been demonstrated to reduce the bullwhip effect [26].

As evident from various sources, information technologies unmistakably reshape the logistics sector and bring strategic value to companies. Their integration into collaborative processes improves productivity and efficiency within warehouse operations [27]. These technologies have become essential tools for gaining a competitive edge in logistics. Supporting productivity, efficiency, and innovation becomes a cornerstone of competitiveness in logistics [28]. Effectively managed information technologies optimize current operations and set the stage for future development and competition.

2.1 The Bibliometric Analysis

This study used a bibliometric approach to understand the academic landscape of digital technologies in logistics competitiveness.

The data collection framework is depicted in Fig. 1.

Web of Science is highlighted as a crucial database, covering a substantial portion of accessible academic literature. The study employs a multi-stage approach to gain a conceptual understanding of the phenomena. Information is gathered from publications on the subject, employing various methods to accomplish this objective. Firstly, a bibliometric analysis is conducted to examine and assess the chosen literature using statistical techniques. The Web of Science is the primary database housing a significant amount of academic literature.

Bibliometric analysis involves exploring and analyzing selected literature using various statistical and mathematical methods. We employed VOS Viewer for bibliometric analysis, covering the period from 2020 to March 2024, comprehensively exploring the concept's evolution.

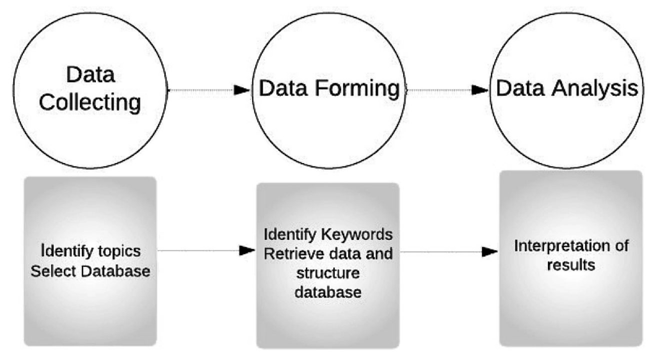


Figure 1 The research process adopted in the study (Author's compilation)

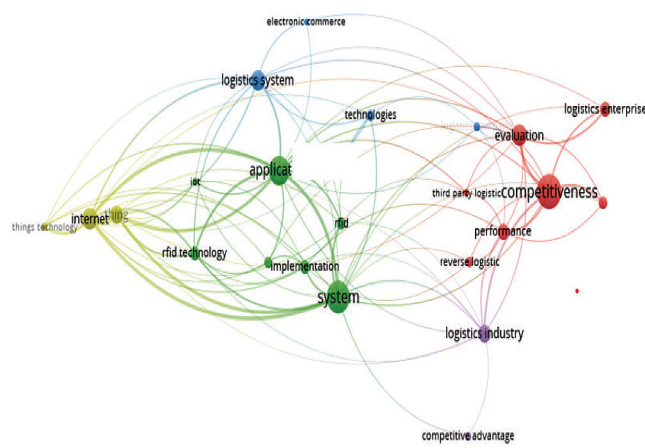


Figure 2 Keyword analysis of authors' keywords (Author's compilation)

VOS Viewer helps create a graphical representation of the dataset, making it easier to interpret extensive data. The generated bibliographic map highlights the articles' most frequently mentioned authors and words. Different colors of circles on bibliometric maps separate clusters to show the keywords 'logistics', 'digitalization', 'technologies', and selected authors in management theory. Circle sizes indicate the importance of the word, with brighter colors signifying greater importance. Lines on the map represent relationships, with brightness indicating the strength of the connection, and different distances define the intensity of the interface.

The constructed bibliometric map consists of five clusters, as illustrated in Fig. 2.

The keywords associated with each cluster, formed through the algorithm, are also presented in the figure.

The first cluster primarily focuses on articles related to competitiveness, evaluation, logistic performance, and logistic enterprise.

The second cluster features articles discussing application, implementation, and logistic information systems.

The third cluster revolves around the common theme of logistics systems, technologies, and logistics companies.

The fourth cluster includes articles discussing Internet of Things (IoT) technology and logistic economy things technology.

The cluster focusing on articles about the Internet of Things (IoT) technology and its impact on the logistics economy (the fourth cluster) appears to be most closely related to the theme of digital technologies in logistics competitiveness.

2.2 Classification of Digital Technologies Used in Logistics

Digitization is how information technologies and digital systems are used to optimize and modernize logistics operations.

It is a field in constant evolution, where new technological solutions are developed to manage the supply chain, warehouses, transportation, information flows, and various aspects of logistics more efficiently. Tab. 1 provides an overview of the integration of digital technologies in various logistics operations.

Digitization in logistics permeates all aspects of operations, from receiving and warehousing goods to optimizing transportation routes and conducting real-time data analytics [10].

These technologies contribute to the digital transformation of logistics operations, making them more efficient and responsive to market demands. Warehousing: This area is related to asset inventory and storage. The technologies include the use of Warehouse Management Systems (WMS), Positioning Systems, RFID (Radio-Frequency Identification) or barcode scanning systems, Augmented Reality (AR), electronic data interchange (EDI), integration with advanced picking systems, and other technologies [29 , 30].

Each business area employs various technologies to enhance operational efficiency, from warehousing and transport management to data analysis and security. This helps optimize the supply chain, reduce time and resource wastage, and increase overall business benefits [18, 32].

Data analysis technologies in logistics and supply chain management provide numerous advantages. These benefits include improved decision-making, cost reduction, enhanced customer satisfaction, and the ability to adapt to changing market conditions, ultimately leading to increased competitiveness in the industry [27, 33].

Table 1 Classification of digital technologies used in logistics

Logistic functional areas	Operations	Technology
Warehousing	Inventory and Warehousing	Warehouse Management System (WMS)
		Positioning System and Augmented Reality (AR)
		Handling Management and Picking Route Optimization
		Storage identification technology
		Cross-Docking Systems
		Internet of Things (IoT) Technology
		RFID (Radio-Frequency Identification) and Barcode Scanning Technologies
		Electronic Data Interchange (EDI)
		Integration with advanced picking systems
Automated Material Handling	Automated Guiding	Automated Storage and Retrieval Systems (AS/RS)
		Automated Guided Vehicles (AGVs)
		Autonomous Mobile Robots (AMRs)
		Guiding Control Software
		IoT Sensors and Data for Decision-Making
Picking and Packing	Automated Picking	Automated picking solutions
		Automated Sortation Systems
		Automated Labeling and Packaging Systems
		Weight Measurement Technology
		Voice Picking Systems
		Real-time data Management Systems (DMS)
Transport Management	Route Planning and Optimization	Global Positioning System (GPS)
		Fleet Management and Route Optimization
		Workload balancing
		Electronic Bill of Lading (eBOL)
		Big Data Analytics
Safety and Security	Safety-oriented transportation	Intelligent transport systems
		AI-driven analytics and warnings (road works, in-vehicle signage, signalized intersections, etc.)
Data Analysis, Interpretation and Exploitation	Demand Forecasting	Big Data Analytics
	Data Collection	Automated Data Collection and Integration
	Labor Augmentation	Robotic Process Automation (RPA)
	Tracking and Monitoring	Real-time IoT Tracking and Monitoring Technology
	Asset Tracking	RFID Asset Tracking, GPS-based asset tracking
	Document Management	Blockchain Technology
	Traceability	Supply Chain Traceability technology
	Real-time decision making	Real-time data and Predictive Modeling, Big data analytics

As we progress in this digital era, logistics professionals and businesses must remain agile and adapt to the evolving landscape. Embracing these digital solutions is no longer an option but a necessity to stay competitive and deliver a seamless logistics experience to customers.

Integrating digital technologies in logistics operations is a game-changer with great promise for the industry. By adopting these innovations, logistics operations can become more efficient and responsive, ultimately contributing to the success of businesses and the entire supply chain.

2.3 Theories that Are Actual for the Application of Logistic Competitiveness

In logistics, the quest for competitiveness is a constant endeavor. Logistics companies must draw upon various theories and models to gain a competitive edge in today's global marketplace. These theories provide an essential framework for understanding and improving logistics operations. In this context, we will explore several competitiveness theories and how they can be applied in logistics, considering the role of digital technologies to enhance efficiency, reduce costs, and provide superior services to customers.

Tab. 2 illustrates the application of various competitiveness theories in the logistics sector and the utilization of digital technologies. Each theory is presented with its application context, highlighting the role of digital technologies in enhancing efficiency, reducing costs, and improving customer service in logistics.

Table 2 Theories that are actual for the application of digital technologies for logistics

Theory	Application in logistics	Role of digital technologies
Michael Porter's Competitive Advantage Theory	Selecting cost leadership or differentiation strategies.	Enhancing differentiation through tracking, data analysis, and customer service, supporting cost leadership.
Michael Porter's Diamond Model	Assessing competitiveness factors in logistics operations.	Improving infrastructure, efficiency, cost reduction, and customer service.
Competitive Strategy Theory	Providing insights into optimal strategies for competitive contexts.	Making data-driven decisions, identifying optimal strategies, and achieving long-term competitive advantages.
Resource-Based View (RBV) Theory	Evaluating unique resources and competencies for a competitive advantage.	Optimizing resources, reducing costs, enhancing efficiency, and improving customer service.
Dynamic Capabilities Theory	Building flexibility, learning from experience, and innovating supply chains.	Enabling real-time adaptation, continuous learning, innovative processes, agile decision-making, and competitive resilience.
Digital Transformation Theory	Adapting to digital trends and excelling in the digital age.	Real-time data collection, enhanced supply chain visibility, efficient inventory management, predictive analytics, improved customer experiences, data-driven decision-making, and sustainable operations.

Michael Porter's competitive advantage theory aids in crafting competitive logistics business models. Digital technologies enhance differentiation through tracking, data analysis, and customer service.

Michael Porter's Diamond Model offers a comprehensive framework for assessing logistics competitiveness, focusing on infrastructure, supplier industries, demand conditions, and supporting industries.

Integrating digital technologies enhances infrastructure, efficiency, cost reduction, and customer service [28]. Competitive Strategy Theory aids in strategic decisions with insights into competitive advantages. Digital technologies provide essential data for data-driven decisions. The Resource-Based View (RBV) theory evaluates unique logistics resources. Digital technology integration optimizes resources, reduces costs, and enhances efficiency and customer service [21].

The Dynamic Capabilities Theory empowers logistics companies to adapt and innovate their supply chains with digital technologies for real-time adaptation, learning, and agile decision-making [34]. Digital Transformation Theory guides logistics companies in adopting digital trends [35].

Logistics companies can leverage a range of competitiveness theories to refine their strategies and operations.

The studies on technologies and business competitiveness were formed as two separate streams. However, integration among these two streams is required to expand the concepts of technology application and increase business competitiveness.

In this research, the authors analyzed the different aspects of applying technologies and business competitiveness by delivering general and more specific investigations.

Theories, such as Michael Porter's Competitive Advantage Theory, Resource-Based View (RBV), and others, play a pivotal role in navigating the logistics industry's digitalization landscape. These theories offer frameworks for understanding market dynamics, competitive forces, and strategic positioning in a digital context [28].

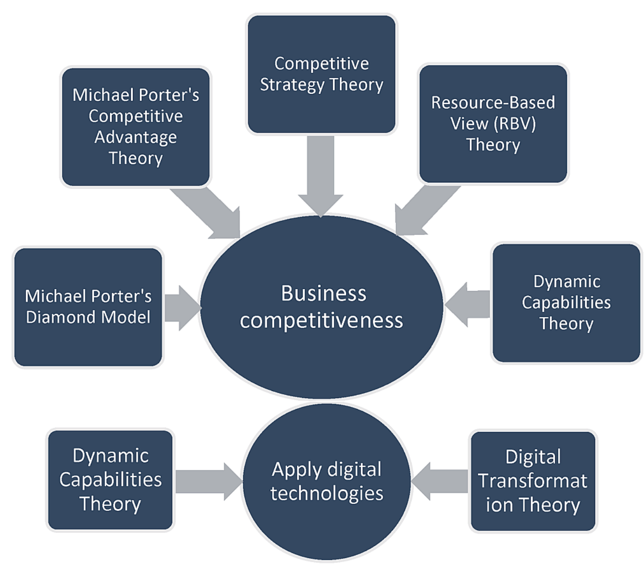


Figure 3 Competitiveness theories and their application

In Fig. 3, theories related to logistics are classified into two categories: those associated with business competitiveness, various competitive strategies, resource evaluation, and identifying unique advantages. The other category comprises strategies focused on applying digital technologies in logistics, such as digital transformation, real-time data collection, supply chain visibility, etc. This emphasizes the significance of these theories in enhancing business competitiveness.

Understanding these theories is crucial in logistics as they aid in strategic planning, decision-making, and optimizing operations. Companies can use these frameworks to assess market dynamics, identify competitive threats and opportunities, allocate resources effectively, and develop strategies to enhance competitiveness. In the digital era, applying these logistics theories helps leverage technology effectively, optimize supply chains, and adapt to changing market conditions to achieve sustained success.

By applying these theories and incorporating these technological advancements, they can achieve improved efficiency, reduced costs, and, most importantly, better meet the ever-evolving needs of their customers in a highly competitive global market.

2.4 The Application of Digital Technologies in Logistics

The evaluation of productivity is commonly assessed by comparing a firm's, an industry's, or a country's performance over time. Indicators linked to the application of digital technologies can generally be categorized into three groups: (1) indicators related to the input application of digital technologies, (2) indicators related to the process of applying digital technologies, and (3) indicators related to the output resulting from the application of digital technologies. These indicators, measured at the firm or sectoral level, affect individual firm productivity. The increasing availability of data and its transformation into economically valuable insights usable for decision-making creates new possibilities for structuring manufacturing within value chains [5].

Digital technologies within the logistics sector can be classified according to their productivity levels, which distinctly shape their impact on operational efficiency. They fall into three productivity categories: low, medium, and high, each exerting a different influence on the logistics industry. Low-productivity digital technologies typically encompass basic systems with restricted functionality. Conversely, high-productivity digital technologies offer sophisticated analysis and optimize logistics activities. For instance, basic data collection and tracking systems, typical of low-productivity technologies, provide essential information on goods' routes but lack advanced analytical capabilities [27].

On the other hand, medium-productivity digital technologies, like certain data analysis platforms, enable in-depth information analysis and decision-making based on analytical data. Customized logistics management systems enhance warehouse operations, route planning, and goods movement, yet may have limitations in functionality. High-productivity digital technologies, such as artificial intelligence and big data analytics, substantially impact logistics sector productivity. They facilitate comprehensive data analysis, market trend forecasting, efficient inventory management, and timely goods delivery. Leveraging artificial intelligence can address intricate logistics challenges, streamline logistics activities, and automate processes, significantly enhancing overall productivity and efficiency [31].

The impact of these technologies on logistics sector productivity varies based on their scope, adaptation, and utilization. Research has demonstrated that utilizing information technologies in logistics can effectively address current and potential future problems, enhance service quality, optimize logistics flows, improve safety standards, reduce resource and product supply costs, enhance information exchange efficiency, and introduce innovative customer support tools [16].

Understanding these distinct productivity levels guides the future direction of digital technologies and empowers businesses to strategically integrate and optimize their technological investments, customizing approaches to augment efficiency and effectiveness within the logistics sector.

3 RESEARCH DESIGN AND RESULTS

3.1 Research Design

The study follows a step-by-step approach to grasp the concepts involved. Initially, the researchers gathered knowledge from data describing advancements in implementing digital application technologies.

The authors use a three-stage methodology. In the first stage, statistical analysis is applied to express and interpret collected data using statistical measures. This method helps understand trends, data distribution, and relationships between variables, providing a foundational basis for subsequent research.

Moving on to the second stage, the authors employ descriptive analysis. This research approach aims to establish facts and describe an object or phenomenon without delving into in-depth explanations or assumptions about causes and effects.

For the third stage, the authors incorporate regression analysis and use heteroscedasticity and autocorrelation-consistent (HAC) standard errors defining the method. Regression analysis explores the relationship between two or more variables, examining the relationship line between a dependent variable and independent variables, enabling the prediction of the dependent variable based on the independent ones.

During the statistical analysis, the authors compare labor productivity across EU logistics sectors and identify productivity gaps between them, with the productivity index taken from Estonia, a leading investor in digital technologies in the Baltic region. In the regression analysis, the authors present regression equations formulated for seven logistics sub-sectors, illustrating diverse effects resulting from the application of digital technologies.

Empirical data from 2011-2020 is used to emphasize the growing emphasis on digitalization within the logistics domain, particularly concerning labor productivity.

In applying the descriptive method, the authors identified nine studies that exemplify the implementation of high-productivity digital technology, specifically voice-picking technology, to enhance logistics management efficiency.

3.2 Review of the EU Logistic Sector

The statistical analysis confirmed that advanced digital technology greatly improves operational efficiency, productivity, and accuracy in logistics across different industries. Companies adopting this technology experienced enhanced workforce performance, shorter training periods, and decreased error rates in logistics operations. This highlights the transformative power of cutting-edge digital technology for effective logistics management in various business settings.

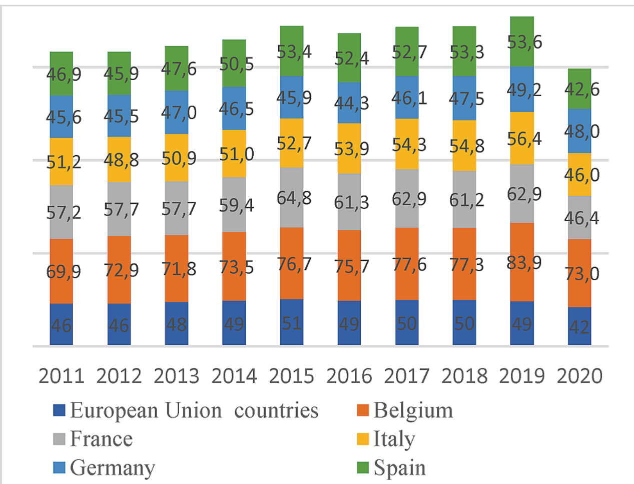


Figure 4 Labor productivity in the logistics sectors of European Union countries [36]

Fig. 4 shows labor productivity in the logistics sector of European Union countries from 2011 to 2020. The percentages indicate the efficiency of generating output relative to the resources used in the logistics domain. For example, Belgium's metrics fluctuate from 69.9% to 83.9%, Germany's from 45.6% to 49.2%, and so on for each country. These percentages reveal the output produced per unit of resources within the logistics sector. The observed fluctuations across countries indicate distinct trends in labor productivity within the logistics industry.

Belgium consistently maintains high logistics labor productivity levels, exceeding 70% for a significant portion of the assessed period. However, looking at the entire European Union, the overall average labor productivity within the logistics sector is approximately 50%. Belgium's exceptional labor productivity suggests they may have implemented various measures to achieve such outstanding indicators. This might involve investments in advanced digital logistics technologies, such as voice picking systems, automated warehouses, or sophisticated inventory management systems, to optimize operational processes and increase efficiency. Large logistics companies and warehouses in various countries, including Germany, France, Spain, and others, have integrated voice-picking technology into their operations. This implementation streamlined order preparation and enhanced warehouse efficiency through voice-controlled digital technology, enabling employees to pick items efficiently. Undoubtedly, this technology has undergone significant changes.

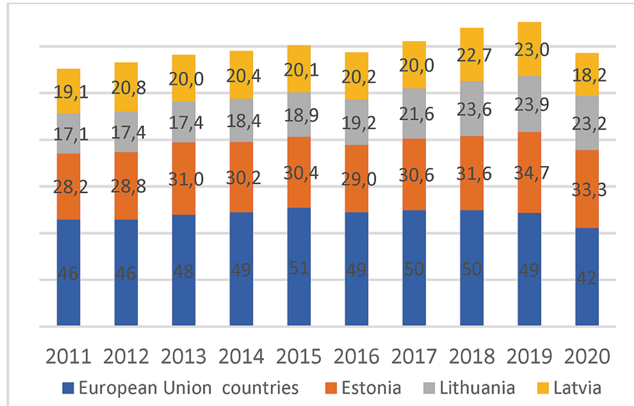


Figure 5 Labor productivity in the logistics sectors of Baltic countries [36]

Fig. 5 delineates labor productivity metrics within the logistics sector across the Baltic States. These metrics represent the percentage scale of productivity levels pertinent to logistics operations. Estonia exhibits a productivity range of 28.2% to 34.7% over the specified duration, Lithuania demonstrates productivity spanning from 17.1% to 23.9%, and Latvia showcases productivity levels oscillating between 18.2% and 23%.

The fluctuations in labor productivity across the Baltic nations uncover distinct trends within the logistics industry. Estonia steadily enhances its performance, maintaining productivity above 30% for most of the analyzed period. Meanwhile, Lithuania and Latvia exhibit varying patterns in their logistics-related labor productivity levels. Among these nations, Estonia consistently maintains higher productivity levels within the logistics sector. They've invested in advanced digital technologies for logistics, including automation, robotics, smart solutions, and data analytics, which effectively streamline processes and boost efficiency in the logistics sector.

3.3 Application of Digital Technologies

Descriptive analysis results show that using advanced digital technology has completely changed how things work in various industries. It has greatly increased productivity, accuracy, and overall efficiency. Companies in different sectors like healthcare and publishing have adopted this technology to make their logistics processes more efficient, and the results have been impressive. This technological shift has improved how employees work and significantly reduced the time it takes for training and the number of errors made.

It highlights how powerful and transformative high-productivity digital technology can be in different business settings.

Tab. 3 overviews how specific advanced digital technologies are used in different logistics operations industries.

Table 3 Results of high-productivity digital technology application for logistics operations [6]

Firms	Application results
Elsevier	Achieving 99.72% accuracy, boosting employee productivity, and quickly realizing a return on investment were notable outcomes. In just 9 months, a significant return on investments was observed. The average training time to familiarize workers with the new system was reduced to 4 hours, an improvement of 2 days compared to RF Scanning. This resulted in a 75% reduction in training time.
Belron	Achieved a 20% increase in productivity, notable enhancements from hands-free work, a higher accuracy rate of 99.95%, improved efficiency resulting in savings of 1.5 million euros, and reduced operating costs.
Simon & Schuster	Achieved a 100% increase in efficiency, quickly realized a return on investment, and eliminated 50 temporary positions. Return on investment was achieved in less than one year. Consolidated the warehouse into one, doubling productivity and efficiency, with 600,000 units shipped daily (150 million units annually).
Pierre's Ice Cream	Improved accuracy by 20%, achieving an accuracy level of 99.96%. This led to higher productivity and accelerated employee training, typically reduced from 7-15 days to 4 hours or less with voice picking, resulting in reduced seasonal hires.
FDL	Improved efficiency in operations, shortened training time, optimized workforce allocation, and analytics solutions aid in order and worker management. These adjustments during peak periods are facilitated, establishing a performance-based intensive pay program.
Anacapri Foods	Boosted productivity without hiring additional staff, minimized errors, expedited employee training, achieved a 20% increase in volume without expanding the workforce, and successfully served over 1800 pizzerias and restaurants.
Frontier Distributing	Attained an outstanding accuracy of 99.9%, boosted productivity, and accelerated employee engagement within the company.
PRSG	Reduced workforce by 25%, improved staff management, minimized travel time, increased on-time orders, and decreased overtime. Enhanced priority management and eliminated manual hand label printing.
Mission Health	Significantly increased accuracy by 30%, improved productivity, enhanced warehouse management, fulfilled more orders, achieved more precise matches on cycle counts, optimized inventory management, and gained the ability to assign priority to work orders.

The outcomes reported by the companies outlined in Tab. 2 reveal a notable 20% surge in workforce productivity. This outcome substantiates its classification as a high-productivity technology. The incorporation of high-productivity digital technology into logistics operations has demonstrably elevated precision, productivity, and operational efficiency within these companies. By integrating this technology, these companies streamlined logistics processes, reduced errors, and optimized workforce management. These findings underscore the versatility and influential impact of high-productivity digital technology in instigating favorable transformations in logistics operations and operational excellence across various sectors.

3.4 The Application of Digital Technologies Estonian Logistic Sector

To thoroughly investigate how digital technologies impact labor productivity in Estonia's logistics sector, this section conducts a detailed analysis from 2005 to 2020. The focus is specifically on investments in computers and software.

The statistical analysis compared labor productivity across various EU logistics sectors, revealing gaps between them. The productivity index was sourced from Estonia, which is known for its significant investments in digital technologies in the Baltic region. The regression analysis formulated regression equations for seven logistics sub-sectors, highlighting the varied effects of digital technology adoption. Empirical data from 2005 to 2020 underscored the increasing focus on digitalization within the logistics domain, particularly concerning labor productivity.

The authors collected data from Estonia Statistics [37] on investments in software and computers by companies in the logistics sector. This covered various sub-sectors, including Postal and courier activities, Warehousing and support activities for transportation, Transportation, and storage, and Transportation by different modes. The intensity of investments varied significantly over the years.

Separate data on labor productivity, calculated by dividing Gross Value Added by the number of persons employed in thousand euros, was gathered for each logistics sub-sector from Estonia Statistics [37].

The analysis centers on labor productivity per person employed, measured by turnover per thousand euros.

The regression analysis reveals changes in labor productivity, notably visible concerning different types of investments (i.e., computer and software) within each logistics sub-sector.

The regression analysis results indicate that understanding the impact of investments in software and computers on labor productivity involves utilizing a regression equation (Eq. (1)).

$$y = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \varepsilon, \quad (1)$$

x_1 – investments in software; x_2 – investments into computers; y – labor productivity; β_0, β_1, β_2 – coefficients of the regression equation; ε – residuals.

The research used the HAC (heteroskedasticity and autocorrelation consistent) test to model and analyze time-series data. This test provides a robust framework for understanding intricate temporal relationships within the dataset and addresses issues of autocorrelation and heteroscedasticity. The method incorporates autoregressive and moving average components of a time series, integrating additional exogenous variables for a more comprehensive and accurate analysis. By using the HAC test, the study aimed to reveal patterns, trends, and dependencies, allowing for a nuanced exploration of the influence of values on the dependent variable. The results underwent validation using test statistics.

Table 4 Formed regression equations (Authors' compilation)

Logistic sub-sector	Regression equation
Investments in software	
Postal and courier activities	$y = 26.3128 + 0.0037 \cdot x_1$
Support activities for transportation	$y = 216.2125 + 0.0077 \cdot x_1$
Warehousing and support activities for transportation	$y = 198.112 + 0.007 \cdot x_1$
Investments in software and computers	
Freight transport by road and removal services	$y = 78.02 + 0.072 \cdot x_1 - 0.0275 \cdot x_2$
Land transport and transport via pipelines	$y = 64.1775 + 0.0336 \cdot x_1 - 0.0064 \cdot x_2$
Transportation and storage	$y = 141.489 + 0.002 \cdot x_1 - 0.009 \cdot x_2$

To understand the impact of investments, regression equations were created for seven sub-sectors in the logistics industry, focusing on two investment categories: software and computers. The coefficients in these equations indicate the change in the dependent variable when the independent variable increases by one unit (see Tab. 4). The authors developed six regression equations.

The regression equation clarifies the relationship between labor productivity and material capital per person employed, quantified in turnover per thousand euros. It's essential to highlight that indicators of investments in computers and software significantly influence logistics sector labor productivity through digital technology. The findings of this analysis are outlined in Tab. 5.

Tab. 5 illustrates that within logistics sub-sectors, investments in software have a discernible impact on labor productivity in Postal and courier activities, Support activities for transportation, and Warehousing and support activities for transportation sub-sectors.

Table 5 Results validation (Authors' compilation)

Variables	FTRRS	LTTP	PCA	SAT	TST	WSAT
β_0	78.0285*** 6.1129	64.1775*** 3.9007	26.3128*** 2.6007	216.2125*** 11.6211	141.489*** 15.5570	198.112*** 10.6550
β_1	0.0720*** 0.0059	0.0336*** 0.0114	0.0037*** 0.0002	0.0077*** 0.0028	0.0020*** 0.0005	0.0070*** 0.0025
β_2	–0.0275*** 0.0110	–0.0064*** 0.0026			–0.0090*** 0.0034	
R-squared	0.3492	0.5999	0.2114	0.1514	0.4926	0.1615
DW		1.5231				
HAC test						
PwlAIC	2		2	2	0	2
BkNW	3		3	3	3	3

Note: *** means that the probability is lower than 0.1. Explanation of abbreviations: FTRRS - Freight transport by road and removal services; LTTP - Land transport and transport via pipelines; PCA - Postal and courier activities; SAT - Support activities for transportation; TST - Transportation and storage; WSAT - Warehousing and support activities for transportation; DW - Durbin-Watson statistics; BkNW - Bartlet kernel, Newey-West fixed bandwidth; PwlAIC - Prewhitening with lags AIC.

We can draw several conclusions based on the findings presented in Tab. 5. Firstly, different sub-sectors show varied responses to investments in software and computers. For instance, all six sub-sectors strongly respond to investments in the corresponding investment year. Each sub-sector demonstrates a distinct correlation between software investments and productivity. However, computer investments positively impact labor productivity in specific sub-sectors, such as Freight transport by road and removal services, Land transport and transport via pipelines, and Transportation and storage.

These findings highlight the diverse influence of software investments on labor productivity across various sub-sectors and types of investments. This emphasizes the importance of tailoring strategies for investments to specific sub-sectors. The formulated regression equations (Tab. 5) reveal that the most significant impacts on labor productivity are associated with investments in software, as all six equations include variables related to such investments. The analysis underscores the varied outcomes of technology investments in different logistics sub-sectors. Additionally, in certain sub-sectors, investments in computers did not show a significant impact on labor productivity. Recognizing these sub-sector-specific characteristics is crucial when planning investments and anticipating their outcomes. Each logistics sub-sector has unique attributes and responds differently to technology investments. Therefore, while technology investments can enhance labor productivity in the logistics sub-sector, it's essential to consider each sub-sector's specific requirements and characteristics when evaluating them. This study provides insights into how digital technologies impact labor productivity in the logistics sector. It highlights a research gap as other authors have not focused on this type of analysis within the logistics sector.

4 DISCUSSION

Recognizing the potential of digitalization in logistics, this paper highlights key areas that require further exploration.

Prajogo et al. (2012) analyzed data from 232 Australian firms and found that logistics integration significantly influences operations performance. Information technology capabilities and sharing are significant in logistics integration [38]. Author Gizetdinov (2024) examined key innovations in transport services, which provide Western companies with a significant advantage in improving logistics efficiency [39]. Al-Ababneh et al. (2023) aimed to define the peculiarities of trends in the evolution of digitalization, innovation in marketing, and logistics among modern companies [40]. Frohlich and Westbrook (2001) modeled supply chain integration in terms of information and material using eight items concerning IT, information sharing, and logistics integration [41]. They found that a wider scope of integration had a positive association with performance improvement. Lee et al. (2000) show that information sharing can lower costs through reductions in inventories and shortages [25]. However, to realize this value, changes in the logistics system are required, such as Vendor-Managed Inventory (VMI) programs, lead time reductions, order quantity reductions, and more frequent deliveries.

Although many researchers have examined the impact of digital technologies on the competitiveness of the logistics sector, there is a clear research gap in exploring future logistics trends. There is a significant opportunity for further scholarly inquiry in this area. Advancements in digital technologies could profoundly influence the competitiveness of the logistics sector. For instance, many studies indicate that information sharing and information technology capabilities are crucial in logistics integration and operational efficiency. However, additional research is needed to further explore how digital technologies may continue to transform and enhance productivity in the logistics sector, thus sustaining business competitiveness.

Bridging the gap between theory and practical applications poses a significant challenge, especially when integrating digital technologies like AI, automation, and data analytics into day-to-day logistics operations. Investigating the obstacles hindering a smooth transition from theory to practice is crucial for successful implementation. The article points out several research gaps:

a) Sectoral Disparities: The analysis reveals diverse responses among different logistics sub-sectors to various technological investments, emphasizing the need to assess these investments while considering the specifics of each sub-sector. This illustrates how identical investments can impact labor productivity differently in various sub-sectors of the logistics industry.

b) Effectiveness of Digital Technology Investments: The effectiveness of digital technology investments can significantly vary depending on the sub-sector. These gaps underscore the importance of considering sub-sector-specific characteristics when planning and implementing technological investments within the logistics sector.

While acknowledging the role of digital technologies in enhancing operational efficiency, understanding the application of specific mechanisms that influence productivity metrics could provide actionable insights for businesses aiming to optimize their logistics operations.

These suggested areas for further investigation aim to address existing research gaps and contribute to a more comprehensive understanding of how digital technologies shape productivity within the logistics sector. Exploring these aspects can provide valuable guidance for businesses navigating the evolving landscape of digital transformations in logistics.

5 CONCLUSION

In summary, our comprehensive analysis underscores the pivotal role of digitalization in transforming the logistics industry and enhancing competitiveness. The focus on applying digital technologies, optimizing efficiency, and understanding emerging trends reveals the transformative potential that lies within the adoption of these technologies. The logistics sector, faced with challenges and opportunities, undergoes a significant reshaping of processes, leading to improved operational efficiency and refined supply chain activities.

Based on papers from the Web of Science, the literature analysis focused on 2020-2024 and delved into various digital technologies, particularly examining their impact on logistics competitiveness. The identified clusters, especially those centered on the Internet of Things (IoT) technology, highlight the interconnectedness of digital technologies and their profound influence on the logistics economy.

Michael Porter's Competitive Advantage Theory, Porter's Diamond Model, Competitive Strategy Theory, Resource-Based View (RBV) theory, Dynamic Capabilities Theory, and Digital Transformation Theory provide valuable frameworks for understanding and improving logistics operations.

Empirical findings illustrate tangible outcomes, indicating a substantial increase in productivity across diverse logistics sectors following technology implementation. The results showcase a 20% boost in productivity, affirming that digital technologies significantly advance company operations, optimize employee activities, and reduce error rates.

However, it's crucial to acknowledge the limitations of our research, particularly the exclusive focus on the Estonian sector due to data constraints. Despite this, Estonia's exceptional labor productivity among Baltic countries is a compelling example of the positive impact of high-productivity digital technologies.

As logistics firms navigate this era of digital adaptation, the presented insights suggest that embracing technology is not merely an option but a necessity for sustained success. The seamless integration of digital solutions positions logistics helps companies meet the evolving demands of tomorrow, fostering a dynamic and efficient industry landscape.

In conclusion, our research illuminates the transformative journey of the logistics sector in the digital age. It reinforces the importance of strategic decision-making based on technology investments, optimizing resource allocation, and cultivating best practices. As logistics companies continue this journey, they address current challenges and position themselves for long-term competitiveness in the ever-changing global marketplace.

In the future, the authors aim to study productivity in other EU countries, focusing on the impact of technology on competitiveness.

6 REFERENCES

[1] Henke, M., Besenfelder, C., Kaczmarek, S. & Fiolka, M. (2020). A vision of digitalization in supply chain management and logistics. In *Proceedings of the Conference on Production Systems and Logistics: CPSL 2020*. https://doi.org/10.15488/9669

[2] Pirogova, O., Plotnikov, V., Makarov, I. & Grafov, A. (2020). The impact of digitalization on the Intellectual capital formation and use (case of service enterprises). *E3S Web of*

Conferences, International Scientific and Practical Conference "Environmental Risks and Safety in Mechanical Engineering" (ERSME-2020), 217, 06009. https://doi.org/10.1051/e3sconf/202021706009

[3] Chen, J. & Zhao, W. (2019). Logistics automation management based on the Internet of Things. *Cluster Computing, 22*, 13627-13634. https://doi.org/10.1007/s10586-018-2041-2

[4] Albrecht, T., Baier, M. S., Gimpel, H., Meierhöfer, S., Röglinger, M., Schlüchtermann, J. & Will, L. (2023). Leveraging Digital Technologies in Logistics 4.0: Insights on Affordances from Intralogistics Processes. *Information Systems Frontiers*, 1-20. https://doi.org/10.1111/isj.12048

[5] Osmolski, W. & Kolinski, A. (2020). Logistics 4.0– Digitalization of the Supply Chains. *Proceedings of the 20th International Scientific Conference Business Logistics in Modern Management*.

[6] Voxware. (2023). Case study. Avalable on https://www.voxware.com/case-studies/

[7] Tang, C. S. & Veelenturf, L. P. (2019). The Strategic Role of Logistics in the Industry 4.0 Era. *Transportation Research Part E: Logistics and Transportation Review, 129*, 1-11. https://doi.org/10.1016/j.tre.2019.06.004

[8] El Zaatari, S., Marei, M., Li, W. & Usman, Z. (2019). Cobot programming for collaborative industrial tasks: An overview. *Robotics and Autonomous Systems, 116*, 162-180. https://doi.org/10.1016/j.robot.2019.03.003

[9] Calatayud, A., Mangan, J. & Christopher, M. (2019). The self-thinking supply chain. *Supply Chain Management, 24*(1), 22-38. https://doi.org/10.1108/SCM-03-2018-0136

[10] Helmke, B. (2022). Digitalization in Logistics. In *Project Management in Logistics and Supply Chain Management: Practical Guide with Examples from Industry, Trade and Services* (pp. 179-201). Wiesbaden: Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978

[11] Kuteyi, D. & Winkler, H. (2022). Logistics Challenges in Sub-Saharan Africa and Opportunities for Digitalization. *Sustainability, 14*, 2399. https://doi.org/10.3390/su14042399

[12] Parfenov, A., Shamina, L., Niu, J. & Yadykin, V. (2021). Transformation of distribution logistics management in the digitalization of the economy. *Journal of Open Innovation: Technology, Market, and Complexity, 7*(1), 58. https://doi.org/10.3390/joitmc7010058

[13] Zaychenko, I., Smirnova, A., Shytova, Y., Mutalieva, B. & Pimenov, N. (2021). Digital logistics transformation: Implementing the Internet of Things (IoT). In *Technological Transformation: A New Role For Human, Machines And Management: TT-2020* (pp. 189-200). Springer International Publishing. https://doi.org/10.1007/978-3-030-64430-7_16

[14] Hoberg, P., Krcmar, H., Oswald, G. & Welz, B. (2017). *Skills for digital transformation*. IDT survey.

[15] Boute, R.N. & Udenio, M. (2023). AI in Logistics and Supply Chain Management. In: Merkert, R. & Hoberg, K. (eds) *Global Logistics and Supply Chain Strategies for the 2020s*. Springer, Cham. https://doi.org/10.1007/978-3-030-95764-3_3

[16] Jahagirdar, A. (2022). *Analysis and development of a digital picking system using IoT and Logistic 4.0*. https://doi.org/10.18419/opus-12634

[17] Brunetti, F., Matt, D. T., Bonfanti, A., Longhi, A. De, Pedrini, G. & Orzes, G. (2020). Digital transformation challenges : strategies emerging from a multi-stakeholder approach. *Digital Transformation Challenges and Strategies, 32*(4), 697-724. https://doi.org/10.1108/tqm-12-2019-0309

[18] Verbivska, L., Zhygalkevych, Z., Fisun, Y., Chobitok, I. & Shvedkyi, V. (2023). Digital technologies as a tool of efficient logistics. *Revista de la Universidad del Zulia*. https://doi.org/10.46925//rdluz.39.28

[19] Christopher, M. (2022). *Logistics and supply chain management*. 6th Edition. Pearson Education Ltd. UK.

[20] Falciola, J., Jansen, M. & Rollo, V. (2020). Defining firm competitiveness: A multidimensional framework. *World Development, 129*, 104857. https://doi.org/10.1016/j.worlddev.2019.104857

[21] Yang, C. S. & Lirn, T. C. (2017). Revisiting the resource-based view on logistics performance in the shipping industry. *International Journal of Physical Distribution & Logistics Management, 47*(9), 884-905. https://doi.org/10.1108/IJPDLM-05-2017-0184

[22] Cagliano, R., Caniato, F. & Spina, G. (2003). E-business strategy: How companies are shaping their supply chain through the internet. *International Journal of Operations & Production Management, 23*(10), 1142-1162. https://doi.org/10.1108/01443570310496607

[23] Tan, K. C., Kannan, V. & Handfield, R. (1998). Supply chain management, supplier performance, and firm performance. *Journal of Purchasing and Materials Management, 34*. https://api.semanticscholar.org/CorpusID:168475226

[24] Cachon, G. P. & Fisher, M. (2000). Supply chain inventory management and the value of shared information. *Management Science, 46*(8), 1032-1048. https://doi.org/10.1287/mnsc.46.8.1032.12029

[25] Lee, H. L., So, K. C. & Tang, C. S. (2000). The value of information sharing in a two-level supply chain. *Management Science, 46*(5), 626-643. https://doi.org/10.1287/mnsc.46.5.626.12047

[26] Disney, S. M. & Towill, D. R. (2003). The effect of vendor managed inventory (VMI) dynamics on the bullwhip effect in supply chains. *International Journal of Production Economics, 85*(2), 199-215. https://doi.org/10.1016/S0925-5273(03)00110-5

[27] Abdul Rahman, N. S. F., Karim, N. H., Md Hanafiah, R., Abdul Hamid, S. & Mohammed, A. (2023). Decision analysis of warehouse productivity performance indicators to enhance logistics operational efficiency. *International Journal of Productivity and Performance Management, 72*(4), 962-985. https://doi.org/10.1108/IJPPM-06-2021-0373

[28] Chung, T. W. (2016). A study on logistics cluster competitiveness among Asia main countries using the Porter's diamond model. *The Asian Journal of Shipping and Logistics, 32*(4), 257-264. https://doi.org/10.1016/j.ajsl.2016.12.010

[29] Borisova, V., Taymashanov, K. & Tasueva, T. (2019). Digital warehousing as a leading logistics potential. In *Sustainable Leadership for Entrepreneurs and Academics: 2018 Prague Institute for Qualification Enhancement (PRIZK) International Conference "Entrepreneurial and Sustainable Academic Leadership" (ESAL2018)* (pp. 279-287). Springer International Publishing. https://doi.org/10.1007/978-3-030-15495-0_29

[30] Agrawal, P. & Narain, R. (2018). Digital supply chain management: An Overview. *IOP Conference Series: Materials Science and Engineering, 455*(1), p. 012074 https://doi.org/10.1088/1757-899X/455/1/012074

[31] Borgi, T., Zoghlami, N., Abed, M. & Naceur, M. S. (2017). Big data for operational efficiency of transport and logistics: a review. *The 6th IEEE International Conference on Advanced Logistics and Transport (ICALT)*, 113-120. https://doi.org/10.1109/ICAdLT.2017.8547029

[32] Ivanova, L., Kurkin, A. & Ivanov, S. (2021). Methods for optimizing routes in digital logistics. *E3S Web of Conferences, Ural Environmental Science Forum "Sustainable Development of Industrial Region" (UESF-2021), 258*, 02015.

https://doi.org/10.1051/e3sconf/202125802015
[33] Arunachalam, D., Kumar, N. & Kawalek, J. P. (2018). Understanding big data analytics capabilities in supply chain management: Unravelling the issues, challenges, and implications for practice. Transportation *Research Part E: Logistics and Transportation Review, 114*, 416-436. https://doi.org/10.1016/j.tre.2017.04.001
[34] Bleady, A., Ali, A. H. & Ibrahim, S. B. (2018). Dynamic capabilities theory: pinning down a shifting concept. *Academy of Accounting and Financial Studies Journal, 22*(2), 1-16.
[35] Gudergan, G. & Mugge, P. (2017). The gap between practice and theory of digital transformation. *Proceeding Hawaii International Conference of System Science*, Hawaii, 1-15.
[36] Eurostat (2020). Use of artificial intelligence in enterprises. Available on https://ec.europa.eu/eurostat/databrowser/view/sbs_na_1a_se_r2__custom_9128690/default/table?lang=en.
[37] Estonia Statistics (2023). Available on https://www.stat.ee/en/find-statistics/statistics-theme/economy/financial-statistics-enterprises
[38] Prajogo, D. & Olhager, J. (2012). Supply chain integration and performance: The effects of long-term relationships, information technology and sharing, and logistics integration. *International Journal of Production Economics, 135*(1), 514-522. https://doi.org/10.1016/j.ijpe.2011.09.001
[39] Gizetdinov, R. (2024). Automation and digitalization of transport forwarding services or logistics. *Infrastructure Asset Management*, 1-9. https://doi.org/10.1680/jinam.23.00050
[40] Al-Ababneh, H. A., Dalbouh, M. A. A., Alrhaimi, S. A., Siam, I. M. & Ibragimkhalilova, T. (2023). Digitalization, innovation and marketing in logistics. *Acta Logistica, 10*(4), 615-624. https://doi.org/10.22306/al.v10i4.440
[41] Frohlich, M. T. & Westbrook, R. (2001). Arcs of integration: an international study of supply chain strategies. *Journal of Operations Management, 19*(2), 185-200. https://doi.org/10.1016/S0272-6963(00)00055-3

Authors' contacts:

Aurelija Burinskiene, Associated Professor
Vilnius Gediminas Technical University,
Sauletekio Avenue, 11 10223 Vilnius, Lithuania
E-mail: aurelija.burinskiene@vilniustech.lt

Diana Daškevič, PhD student
Vilnius Gediminas Technical University,
Sauletekio Avenue, 11 10223 Vilnius, Lithuania
E-mail: diana.daskevic@vilniustech.lt

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240423112303

Subject review
Received: 2024-04-23, Accepted: 2024-04-30

Development and Control of Virtual Industrial Process using Factory IO and MATLAB

Goran Munđar*, Miha Kovačič, Uroš Župerl

Abstract: In today's rapidly evolving business landscape, the strategic adoption of virtual manufacturing methods has emerged as a key driver for companies seeking to streamline operations and expedite product launches in a cost-effective manner. This progressive approach involves the creation of a synthetic and interconnected environment, empowered by advanced software tools and systems, including Virtual Reality and Simulation technologies, tailored to optimize industrial processes. Our methodology employs a unique combination of two simulation software tools: Factory I/O for process development and MATLAB for control program implementation. Furthermore, we explore the use of the Modbus TCP/IP communication protocol as the framework for seamless interaction between these software tools during simulation. This research presents practical insights into the transformative potential of virtual manufacturing, showcasing its real-world application in enhancing operational efficiency and agility within industrial settings.

Keywords: Factory I/O; MATLAB; Modbus TCP/IP; simulation technologies; virtual manufacturing

1 INTRODUCTION

In the contemporary era of industrial advancement, the shift towards digitalization and virtual integration marks a pivotal transformation in manufacturing. The advent of Virtual Manufacturing (VM) embodies this transition, heralding a paradigm where operational efficiencies are enhanced and product development cycles are significantly expedited. This paper explores the strategic deployment of VM through a methodology that integrates software tools Factory I/O and MATLAB. While the use of TCP/IP protocol is widely recognized in various communication setups between hardware and software systems, this research delves into its tailored application within the realm of virtual manufacturing simulations, an area less explored and ripe for innovation [1].

VM emerges as a critical enabler for agility and cost-effectiveness in product launches, pivotal for reducing time-to-market in today's dynamic business landscape [2]. By leveraging Virtual Reality (VR) and Simulation technologies, VM allows for the creation of detailed digital twins, enabling predictive analytics, process optimization, and precise decision-making [3]. The integration of Factory I/O with MATLAB, underpinned by the Modbus TCP/IP protocol, though recognized in various technological applications, is leveraged in unique ways specific to virtual manufacturing scenarios. This combination not only exemplifies VM's practical application but also underscores its potential for scalable implementations. To support our approach, we reference similar applications across various industries, demonstrating how TCP/IP methodology improves operational efficiency and agility [4], [5]. In the automotive sector, TCP/IP is extensively utilized to streamline communication across global manufacturing sites, enhancing coordination and efficiency. Parziale et al. [6] explore its broad applications, highlighting its essential role in automotive manufacturing. Da Silva and Shih [7] discuss the use of TCP/IP in automotive automatic manufacturing systems through an ES/CPS architecture that improves the efficiency and reliability of data communications within production lines. Steffen et al. [8] provide insights into an IP-based network architecture for in-car systems, underscoring TCP/IP's potential to support innovative applications. Similarly, Bohuslava et al. [9] examine how TCP/IP facilitates the dynamic control of robotic cells in automotive manufacturing, playing a critical role in adopting Industry 4.0 technologies.

As the manufacturing sector navigates the complexities of digital transformation, our research provides valuable perspectives on leveraging conventional tools in novel configurations, supporting the evolution toward more dynamic and cost-effective manufacturing processes. This comprehensive analysis underscores the transformative potential of integrating established technologies in novel ways, enriching the discourse on virtual manufacturing and its capabilities within the industrial sector.

2 METHODOLOGY

In our study, we employ a comprehensive methodology to explore the integration of VM tools, specifically focusing on Factory I/O and MATLAB. The primary objective is to showcase how these integrations can significantly improve operational efficiency and agility in industrial settings. In the selection of simulation tools, Factory I/O is chosen for its versatility in simulating real-world industrial processes and its compatibility with a diverse range of control systems. It serves as the primary environment for the development and visualization of manufacturing processes [10]. MATLAB is selected for its advanced computational capabilities, particularly in algorithm development, data analysis, and visualization. Within our methodology, MATLAB is utilized to implement control logic and algorithms that effectively manage the simulated manufacturing processes within Factory I/O [11].

As for the communication protocol, Modbus TCP/IP is identified as the preferred choice due to its widespread adoption in industrial applications and its support for both Factory I/O and MATLAB. This selection is based on the necessity for a reliable and efficient means of data exchange between the simulation tools.

Within the scope of our research, we have developed two distinct systems to demonstrate the capabilities of our methodology. The first system, referred to as the test system, was specifically designed for showcasing the Modbus TCP/IP communication protocol. This system serves as a comprehensive illustration of the effective exchange of data between Factory I/O and MATLAB, highlighting the reliability and efficiency of the chosen communication protocol. Additionally, we have created a second system, which we term the industrial system, to further underscore the versatility of our approach. This industrial system emulates a more complex scenario, involving the assembly of two components into a single part. Through this system, we demonstrate that MATLAB can effectively handle intricate control programs, showcasing its proficiency in algorithm development, data analysis, and visualization within the context of more sophisticated manufacturing processes.

3 TEST SYSTEM DEVELOPMENT

In this chapter, we demonstrate the development of our test system, dedicated to demonstrating the effectiveness of the Modbus TCP/IP communication protocol. We detail the design and implementation of the system, offering insights into its structure, functionalities, and the successful outcomes achieved through seamless data exchange between Factory I/O and MATLAB.

3.1 System Development in Factory IO

In the software tool Factory I/O, three components were utilized for the development of the test system:

- Start Button,
- Light Indicator,
- Electric Switchboard.

The arrangement of these components within Factory I/O is depicted in Fig. 1. Specifically, the "Start Button" component serves as the digital input for the test system, the "Light Indicator" represents the digital output, and the "Electric Switchboard" is employed to mount the components to prevent them from floating in the air.

3.2 Definition of Communication Protocol

To establish communication between the virtual system created in Factory I/O and the control program developed in Matlab, we utilized the Modbus TCP/IP communication protocol. In Factory I/O, we defined the communication mode as "Modbus TCP/IP Server". The key steps for establishing communication involved configuring the IP address, port, Slave ID, network adapter, and specifying the type and number of inputs or outputs. The communication configuration in Factory I/O is illustrated in Fig. 2.

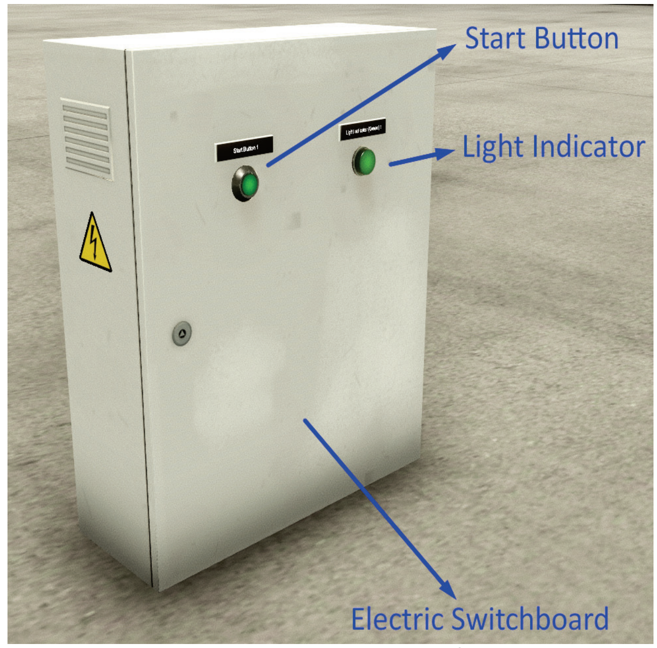


Figure 1 Test system in Factory IO

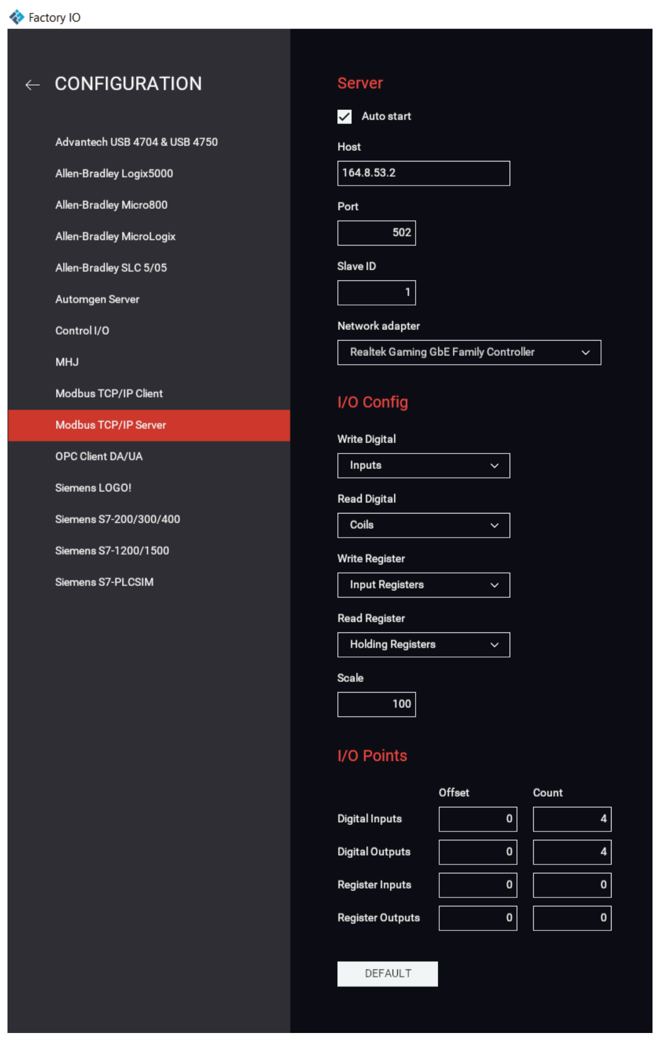


Figure 2 Communication configuration in Factory IO

3.3 Definition of I/O Component Addresses

After the successful configuration of the communication protocol, it is necessary to determine the addresses of the digital input and digital output. The digital input "Start Button" was defined as "Input 0", and the digital output "Light Indicator" was defined as "Coil 0". The described connection is shown in Fig. 3.

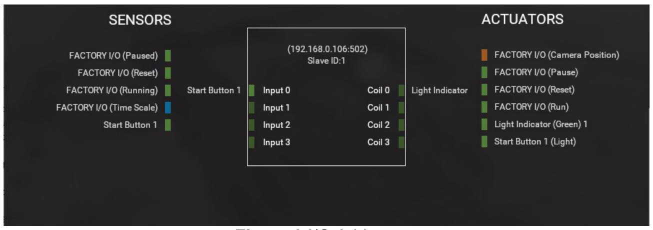


Figure 3 I/O Addresses

3.4 Development of a Control Program in MATLAB and Simulation

To effectively devise a control program within the MATLAB software tool, integration with the "Industrial Communication Toolbox" [12] was imperative. This toolbox facilitates connectivity to Modbus servers and offers diverse functionalities, including the creation of interface objects and functions for data reading, writing, and manipulation.

For the establishment of Modbus communication, the initial step involves creating an interface object using the "modbus" function. Essential data such as "Transport", "Device IP Address", and "Port" must be inputted for proper functionality. Upon successful communication setup, a control program is developed utilizing "read" and "write" functions. The control program follows a sequence: first, the Modbus connection is configured using the Modbus function. Then, in an iterative loop, the status of the Start button in Factory I/O is continually assessed through the read function. Upon detecting a pressed Start button, the write function is employed to activate the "Light Indicator" in Factory I/O. The outlined control program for the test system is provided below:

```
% MODBUS TCP connection with Factory IO
Server
m = modbus("tcpip",'192.168.0.106',502)

% Infinite loop
while true
    %read "button" state
    inputReg0 = read(m,"inputregs",1)

    %turn light indicator on/off
    if (inputReg0 == 1)
        write(m,"coils",1,1)
    else
        write(m,"coils",1,0)
    end
end
```

Post the development of the control program, we conducted a simulation of its operation. Initiation involved running the simulation in Factory I/O followed by launching the control program in MATLAB. The simulation executed successfully, allowing us to toggle the "Light Indicator" on and off by pressing Start button in Factory IO.

4 PRODUCT ASSEMBLY SYSTEM

In this chapter, the creation of a more complex virtual system for product assembly and the development of a control program to manage the developed system will be described.

4.1 Technological Requirements

The primary function of the developed system is to transport two components (bases and covers of the product) along conveyor belts to a workstation. Upon reaching the workstation, the components undergo positioning to the appropriate location. Subsequently, utilizing a two-axis manipulator, the product covers are grasped using a pneumatic gripper and transferred to the product bases, where the assembly takes place. Following a successful assembly, the completed product exits the workstation, making room for new components to undergo the assembly process.

4.2 Development of the Virtual System in Factory I/O

In the creation of the described virtual system in Factory I/O, various actuators, sensors, and structural components were employed. The actuators utilized include:

- 2 Belt Conveyors with a length of 4 m
- Two-Axis Pick & Place manipulator
- 2 Right Positioning pneumatic manipulators
- Digital Display
- 3 Light Indicators.

Ensuring the successful operation of the system required the incorporation of sensors. The following sensors were integrated into our system:

- 3 Diffuse Sensors for detecting elements on the conveyor belts
- 8 limit switches – 2 installed on each pneumatic actuator
- 2 sensors for detecting movement in individual axes of the two-axis manipulator
- 3 control buttons for system operation
- Emergency switch for immediate shutdown.

In addition to actuators and sensors, the constructed system includes structural components such as sensor mounts, an electrical enclosure, an enclosure mount and safety fences. While these components do not directly affect the system's operation, they contribute to a more aesthetically pleasing and illustrative representation of how such a system might appear in a real industrial environment.

Finally, components for element generation and removal were added to the system. Two Emitter components were employed for element generation—one for generating product bases and the other for generating covers of the final products. The system also includes a Remover component, responsible for deleting assembled products when they reach the end of the conveyor belt.

Fig. 4 illustrates the complete virtual industrial system created in the Factory I/O software tool.

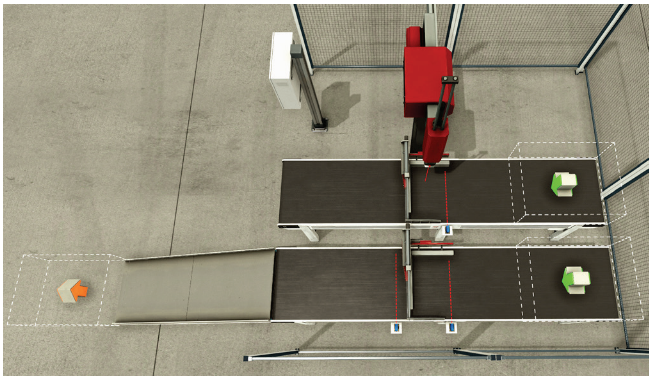


Figure 4 Virtual industrial system

4.3 Definition of I/O Addresses

The addresses of actuators and sensors were defined using the same approach as outlined in the test system described in Chapter 3.3. A total of 16 digital inputs, 12 Coils (digital outputs), and 1 Output Holding Register were employed. The Output Holding Register serves the purpose of controlling the digital display, responsible for monitoring the number of assembled products. Fig. 5 presents the definition of all utilized inputs and outputs in the Factory I/O software tool.

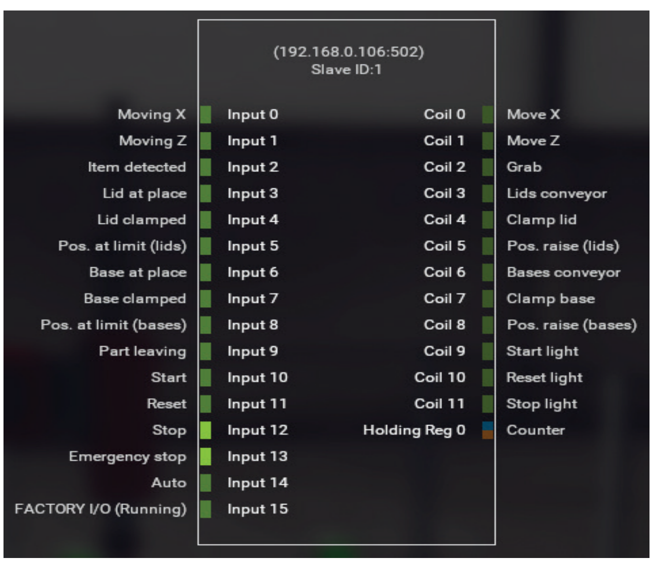


Figure 5 Definition of I/O Addresses

4.4 Development of Control Program and Simulation

The development of the control program in MATLAB followed a similar methodology as outlined in the test system described in Chapter 3.3. To construct the control program, we utilized the "modbus" function for establishing the Modbus TCP connection, the "read" function for retrieving sensor values from Factory I/O, and the "write" function for activating and deactivating actuators.

Once the control program was meticulously crafted, a comprehensive simulation of the system's functionality was conducted. This simulation aimed to validate the effectiveness and efficiency of the control logic in managing the virtual industrial system. The successful execution of the simulation affirmed the reliability of the control program, showcasing its capability to respond appropriately to dynamic inputs and successfully manipulate the virtual components within Factory I/O. System during simulation is shown on Fig. 6.

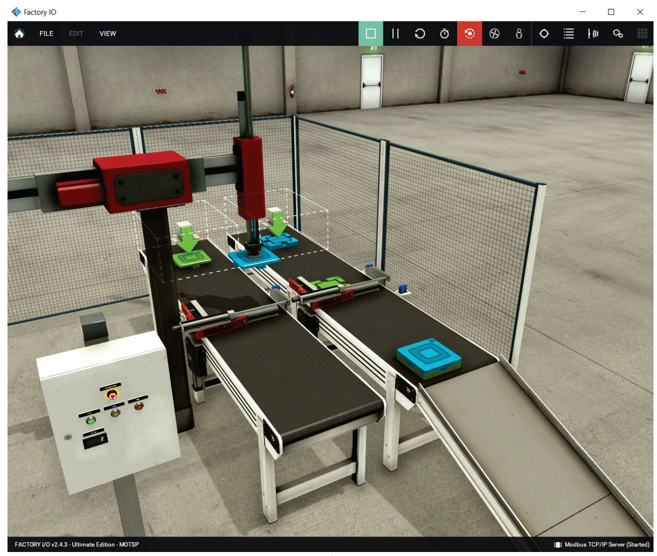


Figure 6 Simulation

5 RESULTS

The simulation of both the test system and the more complex industrial assembly system proved to be highly successful. The Modbus TCP communication protocol demonstrated sufficient speed, ensuring the smooth operation of the simulations without any hindrance.

Following the successful simulations, we conducted further tests to evaluate the performance of communication and control programs at higher simulation speeds. Factory I/O allows simulation speeds to be adjusted from 0.1x to 4x. We systematically tested the functionality at all speeds. The simulation operated seamlessly and could be effectively controlled even at significantly elevated speeds. The results indicated that both the communication protocol and the control program maintained robust functionality, affirming their efficacy even in dynamic and high-speed simulation environments.

Furthermore, the real-time monitoring capabilities of the control program were evaluated during the simulations, showcasing its ability to react promptly to changes in the virtual environment. This responsiveness is a key attribute for real-world applicability, demonstrating the adaptability of the developed system to dynamic manufacturing scenarios.

In summary, the results highlight the successful implementation and robust performance of the proposed virtual manufacturing framework, validating its capability to handle intricate communication and control tasks, even in dynamic and accelerated simulation scenarios.

6 CONCLUSION

This work has presented the process of developing a virtual industrial system using the Factory I/O software tool and creating a control program within MATLAB. The seamless communication between these programs was established through the Modbus TCP communication protocol, selected for its widespread adoption and ease of integration, despite the availability of more secure protocols, which could be considered in future enhancements to improve cybersecurity measures.

MATLAB, as a powerful programming environment, was chosen over traditional PLCs for its superior computational capabilities, enabling more complex simulations and sophisticated data analysis. This decision allows for straightforward control of the constructed systems within Factory I/O, showcasing the potential of high-level software tools in reducing the time and costs associated with physical prototyping. Furthermore, the integration of MATLAB enhances the system's flexibility and scalability, proving essential for addressing complex industrial challenges that PLCs alone might not efficiently manage.

The procedural framework outlined in this work showcases how digital twins of real systems and corresponding control programs can be efficiently developed. This methodology streamlines the creation of control programs, simultaneously reducing costs by facilitating simulation-based testing without the need for physical components. This not only minimizes the potential expenses associated with damage to real system elements but also mitigates the risk of harm to personnel.

The developed system is amenable to further enhancements. MATLAB offers extensive extensions, such as the Simulink® PLC Coder™ [13], enabling the creation of control programs in the Ladder programming language within MATLAB. While automatically generated Ladder diagrams from Simulink might require optimization to remove extraneous elements, they provide a foundational code that can be refined for practical application in PLCs. This could potentially facilitate easier learning of programming and control program development in educational settings, where programmers may lack sufficient experience to test their control programs on real systems.

In essence, the presented approach demonstrates the viability of employing virtual manufacturing frameworks for the design and testing of control programs, providing a valuable alternative to traditional methods. The combination of Factory I/O and MATLAB, coupled with the potential for further expansion, presents a robust solution for the development, testing, and educational application of control systems in diverse industrial contexts.

Acknowledgements

This research was funded by the Slovenian Research and Innovation Agency; grant number L2-3173.

7 REFERENCES

[1] Swales, A. (1999). Open modbus/tcp specification. *Schneider Electric*, *29*(3), 19.
[2] Soori, M., Arezoo, B. & Dastres, R. (2023). Virtual manufacturing in industry 4.0: A review. *Data Science and Management*. https://doi.org/10.1016/j.dsm.2023.10.006
[3] Salah, B., Abidi, M. H., Mian, S. H., Krid, M., Alkhalefah, H. & Abdo, A. (2019). Virtual reality-based engineering education to enhance manufacturing sustainability in industry 4.0. *Sustainability*, *11*(5), 1477. https://doi.org/10.3390/su11051477
[4] Munđar, G. (2022). Development and control of a virtual industrial process for assembling computers. *IOP Conference Series: Materials Science and Engineering, 1268*(1), p. 012017. https://doi.org/10.1088/1757-899X/1268/1/012017
[5] Wang, B. (2020). Research and Development of Virtual Manufacturing. In *International Conference on Applications and Techniques in Cyber Intelligence ATCI 2019: Applications and Techniques in Cyber Intelligence, 7*, 1523-1529. Springer International Publishing. https://doi.org/10.1007/978-3-030-25128-4_188
[6] Parziale, L., Liu, W., Matthews, C., Rosselot, N., Davis, C., Forrester, J. & Britt, D. T. (2006). TCP/IP tutorial and technical overview.
[7] Shih, I. J. L. & da Silva, G. C. (2018). Proposal of ES/CPS Architecture for Automotive Automatic Manufacturing Systems. *Journal of Production and Automation (JPAUT) ISSN 2595-9573*, *1*(1), 19-34.
[8] Steffen, R., Bogenberger, R., Hillebrand, J., Hintermaier, W., Winckler, A. & Rahmani, M. (2010, May). Design and realization of an ip-based in-car network architecture. *The 1st International ICST Symposium on Vehicular Computing Systems*.
[9] Bohuslava, J., Martin, J. & Igor, H. (2017). TCP/IP protocol utilisation in process of dynamic control of robotic cell according industry 4.0 concept. *The IEEE 15th International Symposium on Applied Machine Intelligence and Informatics (SAMI)*, 000217-000222. https://doi.org/10.1109/SAMI.2017.7880306
[10] Factory I/O – Next-Gen PLC Training. https://factoryio.com/
[11] Matlab, S. (2012). Matlab. *The MathWorks, Natick, MA*.
[12] Industrial Communication Toolbox. https://www.mathworks.com/products/industrial-communication.html
[13] Simulink® PLC Coder™ User's Guide. https://www.mathworks.com/help/pdf_doc/plccoder/plc_st_ug.pdf

Authors' contacts:

Goran Munđar, M.Sc. Mech. Eng.
(Corresponding author)
University of Maribor, Faculty of Mechanical Engineering,
Smetanova 17, 2000 Maribor, Slovenia
goran.mundjar@um.si

Miha Kovačič, PhD, Associate Professor
Štore Steel d.o.o.
Železarska 3, 3220 Štore, Slovenia
University of Ljubljana, Faculty of mechanical engineering,
Aškerčeva cesta 6, 1000 Ljubljana, Slovenia
miha.kovacic@store-steel.si

Uroš Župerl, PhD, Associate Professor
University of Maribor, Faculty of Mechanical Engineering,
Smetanova 17, 2000 Maribor, Slovenia
uros.zuperl@um.si

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240423182840

Subject review
Received: 2024-04-23, Accepted: 2024-04-30

Development of Guidelines for the Design of Cantilevers

Filip Šulentić, Daniel Miler, Matija Hoić*

Abstract: Cantilevers are elements typically used in the design of various structures such as factory halls or cantilever cranes. Traditionally, such elements are made of standard profiles or, in the case of large loads, by welding from standard sheets. Often, due to stationary conditions of exploitation, no attention is paid to mass reduction. Increasingly stringent requirements on reducing costs, but also on reducing environmental impact through reducing energy consumption necessary for the production, processing, and transport of materials, facilitate the application of optimization to obtain a construction solution with minimal mass. However, doing so requires advanced high-cost computer tools and increases the required competencies of designers. In this paper, a series of optimizations have been carried out for typical reach and load values of cantilevers aiming to create a database that will facilitate algorithms for selecting cross-sectional dimensions suitable for production using either additive technologies or traditional methods.

Keywords: cantilever; deflection; guidelines; optimization; strength

1 INTRODUCTION

Cantilevers are a common structural element in a range of applications, from common load-bearing structures to energy sources where they convert mechanical vibrations into electrical energy through the piezoelectric effect [1]. The development and application of optimization methods therefore did not bypass even this simple structural element. Newer studies cover a wide range of applied methods, which include the use of neural networks to search for the optimal design [2], topological optimization by subtracting material to achieve the smallest mass [3], and heuristic methods such as simulated annealing [4].

In specific cases, cantilevers have a complex design, that includes active components, so the active control of the deformation of the cantilevers is achieved through the application of dielectric elastomers [5], the application of piezoelectric layers aims to actively manage the vibration modes of the cantilever [6], and the design of composite versions is also being improved [7]. The calculation of large deformations of both simple [8] and multi-stage cantilevers [9] is also investigated.

Advanced methods enable obtaining innovative results; however, for most applications in everyday engineering practice, they may not necessarily be applicable. This is because they require the application of advanced knowledge, and the use of expensive computer programs to solve typical recurring problems. An alternative approach is the formation of design guidelines based on statistical analysis of optimized solutions for a wide range of design parameters for the observed mechanical system.

In this article, an example of a console as a load-bearing element in load-bearing structures is considered. Chapter 2 describes the general engineering problem of designing a load-bearing cantilever with a constant rectangular section. In the third chapter, the application of the parametric optimization module in the Ansys software package is described. Further, optimization results are analysed in Chapter 4, and the guideline formation process is described. The discussion, main conclusions, and possibilities for further development are given in the final chapter.

2 CANTILEVER DESIGN

The cantilever under consideration is a simple structure with a constant rectangular cross-section. The console is fixed at one end, loaded with a concentrated force at the other end and by its mass (Fig. 1a). The section of the console (Fig. 1b) consists of two identical vertical bands of width t and height h and two identical horizontal bands of width b and height t. Vertical bands are spaced at distances equalling $0{,}8b$ as this distance relates to typical classic welded steel sheet design. Moreover, the same design could also be used if manufactured using one of the more recently emerging technologies such as additive manufacturing. The target material is a typical structural steel such as S235.

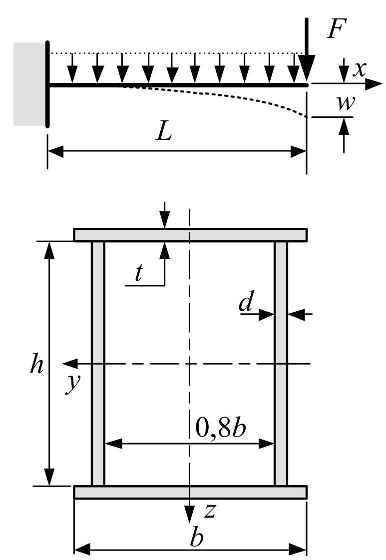


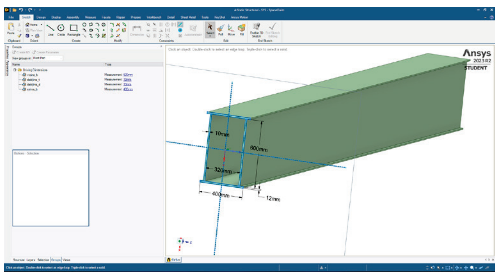
Figure 1 Cantilever beam model (top) and its cross-section (bottom)

The set of typical values of reach L and concentrated loads F are as follows: $L = \{2, 4, 6, 8\}$ m and $F = \{1, 2, 3, 5\}$ kN. The permissible deflection of the console was set to $w = L/600$, while the permissible stress was taken as $\sigma = 120$ MPa.

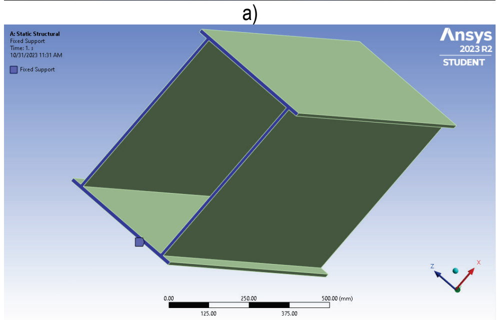
3 OPTIMIZATION PROCEDURE

The optimisation was carried out using a parametric optimization module within *Ansys* based on an Adaptive Single-Objective The selected method utilizes a gradient-based algorithm, thus yielding a refined, global optimum. The method is applicable to single-objective problems with multiple constraints and aims to find the global optimum. It is limited to continuous and manufacturable input parameters.

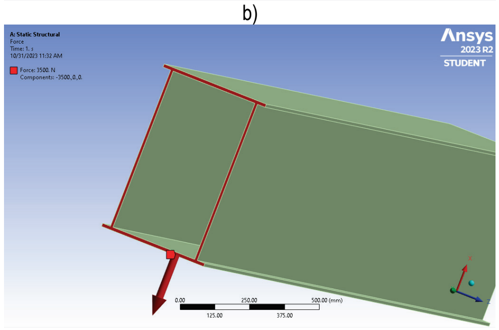
Firstly, the CAD model is designed with nominal values of individual dimension (Fig. 2a), fixed at one end (Fig. 2b), and continous force uniformly distributed across the opposite cross-section (Fig. 2c).



a)



b)



c)

Figure 2 FEM analysis setup within Ansys

Optimization of initial design is conducted through variation of dimension values of the initial design. Upper and lower boundary for each of four dimensions are set at approx. 90 and 110 % of the initial values (Fig. 3).

Optimization problem goal was minimizing the overall mass of the cantilever while adhering to the requirements regarding the maximum allowable deformation and stress (Fig. 4).

The optimization process was started, and the initial FEM analysis is conducted using cubic elements, providing results on stress and deformation (Fig. 5).

Table of Schematic B2: Optimization

	A	B	C	D
1	Input Parameters			
2	Name	Lower Bound	Upper Bound	
3	P1 - visina_h (mm)	500	700	
4	P2 - debljina_t (mm)	10	14	
5	P3 - debljina_d (mm)	8	12	
6	P4 - sirina_b (mm)	320	480	
7	Parameter Relationships			
8	Name	Left Expression	Operator	Right Expression
*	New Parameter Relationship	New Expression	<=	New Expression

Figure 3 Lower and upper boundaries of design variables

Table of Schematic B2: Optimization

	A	B	C	D	E	F	G	H	I
1	Name	Parameter	Objective			Constraint			
2			Type	Target	Tolerance	Type	Lower Bound	Upper Bound	Tolerance
3	P5 <= 13.3 mm	P5 - Total Deformation Maximum	No Objective			Values <= Upper Bound		13.3	0.001
4	P6 <= 120 MPa	P6 - Equivalent Stress Maximum	No Objective			Values <= Upper Bound		120	0.001
5	Minimize P7	P7 - Geometry Mass	Minimize	0		No Constraint			
*		Select a Parameter							

Figure 4 Optimization problem setup

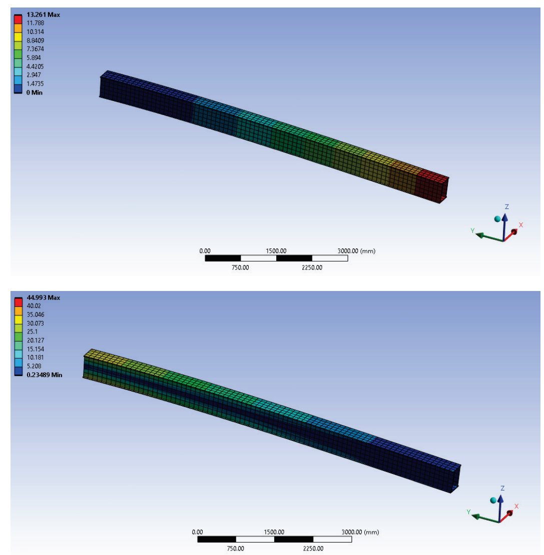


Figure 5 FEM analysis results for L = 8 m, F = 5 kN

If the initial design results in satisfactory stress and deformation, or stress and deformation results that are slightly above target values, the optimization module will generate several additional candidates and repeat the FEM analysis for each candidate. Based on the results of the newly generated solutions, the program generates the subsequent candidates by combining somewhat altered dimensions of the current generation (Fig. 6).

Table of Schematic B2: Optimization

	A	B	C	D	E	F	G	H
1	Name	P5 - b (mm)	P6 - h (mm)	P7 - t (mm)	P8 - d (mm)	P9 - Total Deformation Maximum (mm)	P10 - Equivalent Stress Maximum (MPa)	P11 - Geometry Mass (kg)
4	3 DP 14	335.62	516.09	10.078	12.364			
5	4	373.86	475.04	10.188	11.932			
6	5 DP 16	361.11	545.42	9.4209	11.357			
7	6	395.1	510.23	9.5305	11.645			
8	7	327.12	469.18	9.0923	12.22			
9	8	365.36	451.58	9.3114	11.501			
10	9	348.36	463.31	9.8591	13.083			
11	10	356.86	504.36	8.5446	11.214			
12	11	386.6	498.5	9.64	13.226			
13	12 DP 23	369.61	539.55	9.9687	12.508			
14	13 DP 24	344.12	521.96	9.2019	13.37			
15	14	339.87	492.64	9.7496	11.07			
16	15 DP 26	331.37	527.82	8.8732	12.076			
17	16	390.85	480.91	8.6541	11.789			
18	17 DP 28	382.35	533.69	8.7637	12.651			

Figure 6 Generation of further potential optimal solutions

By variation of dimensions for successful solutions, the procedure results in improved solutions which tend to approach an optimum. Once the improvement in mass reduction is no longer significant, the procedure stops resulting in a solution that is very near to the optimal one (Fig. 7).

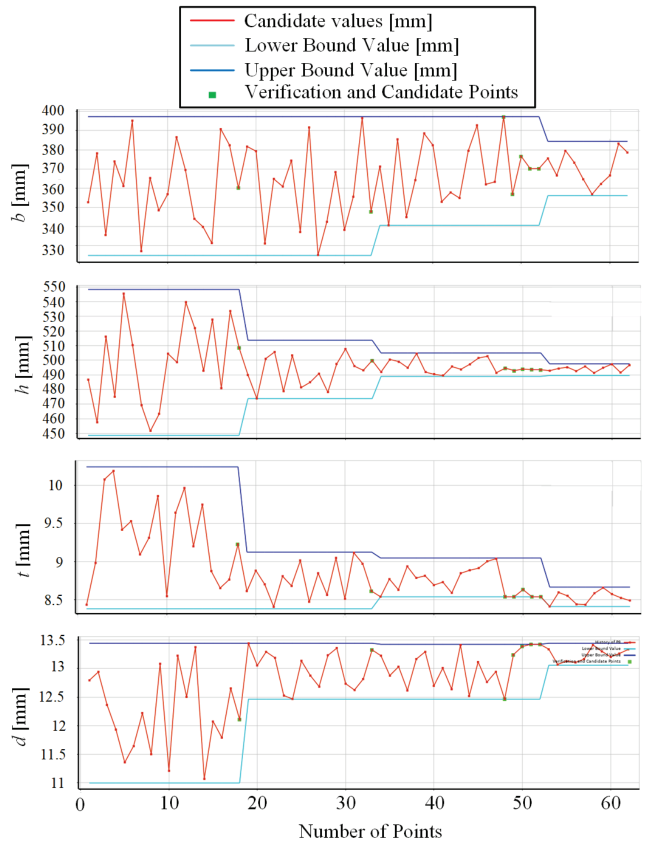


Figure 7 Generation of further potential optimal solutions

It should be noted that the choice of upper and lower boundaries might prevent the algorithm in finding either feasible (if the upper dimension values are too small to achieve acceptable stress and/or deformation) or optimal solution (if the lower boundaries results in stress and/or deformation which are well above set limits). One solution would be to widen the boundaries. However, wider boundaries will induce a wide variation in potential solutions, prolonging the optimization process. This, on the other hand, may significantly increase optimization times for individual set of parameters to several hours, possibly even hindering the ability to find optimum solution. The choice of initial solution and boundary width thus have a significant impact on the overall required time.

Additionally, the size of the finite elements was set automatically, i.e. the mesh density may not be sufficient to result in a realistic stiffness and consequently stress values. More then 60 FEM analysis was conducted per each optimization which resulted in several hours of computing per each of 18 investigated combinations using a moderate power personal computer. Mesh with higher density would thus necessitate usage of workstation grade computer which was not available in this case.

Herein, the procedure was started by finding optimal solution for the greatest force and reach. Several attempts were needed until a good set of initial parameters and boundary widths was found. Once the optimal solution was found for this extreme case, optimization was conducted for neighbouring sets of loads and reach with initial sets of dimensions being set as approx. 10 % lower than the optimum solution for the extreme case. The same approach was then used for each subsequent set of smaller loads and reaches.

As the algorithm progresses, it might find that the limits are narrowed down, and it will do so dynamically to speed up the process (blue and light blue lines in Fig. 7). For each combination of load, reach, initial dimensions, and boundaries, properly set limits will result in the optimum solution being within boundaries. Further, it may be observed that the algorithm will widen the limits from their contracted values if needed; however, not beyond limits initially set by the user.

Once the improvement in the obtained solutions falls under the set limit, the program will stop and offer the three best solutions, i.e. three solutions with the smallest overall masses. Herein, the solution with the smallest mass of three is selected for further analysis.

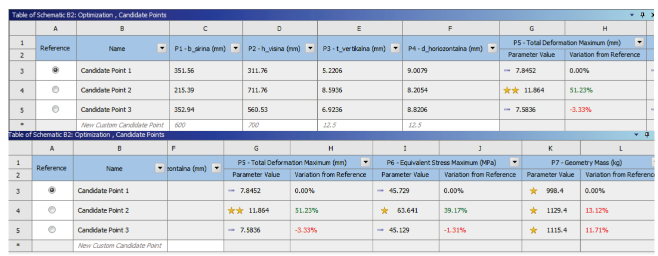


Table of Schematic B2: Optimization , Candidate Points

	A	B	C	D	E	F	G	H
1	Reference	Name	P1 - b_sirina (mm)	P2 - h_visina (mm)	P3 - t_vertikalna (mm)	P4 - d_horiozontalna (mm)	P5 - Total Deformation Maximum (mm)	
2							Parameter Value	Variation from Reference
3	◉	Candidate Point 1	351.56	311.76	5.2206	9.0079	7.8452	0.00%
4	○	Candidate Point 2	215.39	711.76	8.5936	8.2054	★★ 11.864	51.23%
5	○	Candidate Point 3	352.94	560.53	6.9236	8.8206	7.5836	-3.33%
*		New Custom Candidate Point	600	700	12.5	12.5		

Table of Schematic B2: Optimization , Candidate Points

	A	B	F	G	H	I	J	K	L
1	Reference	Name	zontalna (mm)	P5 - Total Deformation Maximum (mm)		P6 - Equivalent Stress Maximum (MPa)		P7 - Geometry Mass (kg)	
2				Parameter Value	Variation from Reference	Parameter Value	Variation from Reference	Parameter Value	Variation from Referenc
3	◉	Candidate Point 1		7.8452	0.00%	45.729	0.00%	★ 998.4	0.00%
4	○	Candidate Point 2		★★ 11.864	51.23%	★ 63.641	39.17%	★ 1129.4	13.12%
5	○	Candidate Point 3		7.5836	-3.33%	45.129	-1.31%	★ 1115.4	11.71%
*		New Custom Candidate Point							

Figure 8 Example of three offered solutions after the termination of optimization procedure

4 ANALYSES OF RESULTS

A total of 18 combinations were investigated for the target range of loads and reaches with basic 16 combinations made from four values for force and lever length and two additional outlying sets which include smaller force and longer reach and vice versa. The results are presented in Tab. 1 and Fig. 9.

Table 1 Overview of obtained solutions

	F	*L*	*b*	*h*	*t*	*d*
1	5	8000	333,6	450,1	8,4	10,9
2	3	8000	308,2	433,7	8,2	10,1
3	5	6000	310,9	406,2	8,8	10,6
4	3	6000	259,7	413,4	7,9	9,2
5	2	8000	248,7	364,2	7,6	8,9
6	5	4000	269,9	337,8	7,4	8,6
7	3	4000	207,1	264,0	6,9	7,6
8	2	4000	177,2	224,4	7,0	8,4
9	5	2000	183,0	288,0	6,2	6,5
10	2	6000	182,4	284,7	6,8	7,7
11	1	6000	151,9	302,1	6,3	6,7
12	1	8000	199,4	304,1	7,2	8,3
13	1	4000	125,3	277,3	5,2	5,1
14	3	2000	93,7	183,5	5,0	6,1
15	2	2000	79,8	150,3	3,2	4,5
16	1	2000	62,3	153,5	3,1	4,4
17	5	9000	375,5	492,8	8,4	13,3
18	0,5	1000	74,2	155,2	4,1	4,6

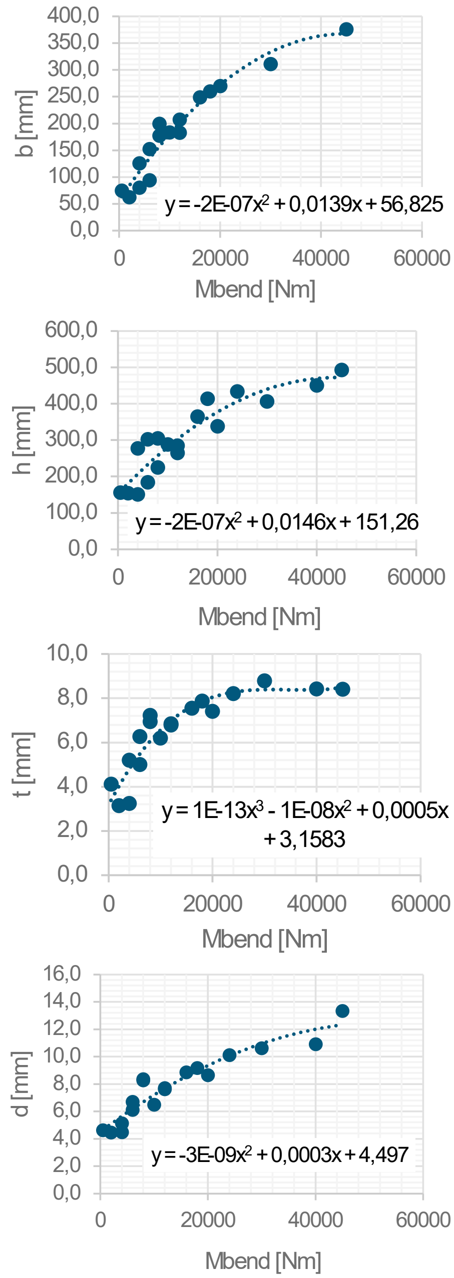


Figure 9 Example of three offered solutions after the termination of the optimization procedure

Obtained optimal dimensions for a given combinations of loads and reaches are plotted in Fig. 9 where *x*-axis shows the bending torque calculated as $M_{\text{bend}} = F \cdot L$, and the *y*-axis shows corresponding dimension values. It may be observed that the results can be described well using a 2nd or a 3rd order polynomial. The $y = f(x)$ equations shown in the figures thus allow the user to calculate the near optimum dimensions of the cantilever cross section for the known required reach and concentrated load at that reach.

Some variations in Fig. 9 can be observed with several points of optimum solutions being quite far from the regression curve. Hence, further analysis of stress and deformation were carried out as the obtained values from FEM analysis were used as limits during optimization. Analytical equations were used to obtain stress and deformation values for the obtained optimal solutions [10]:

$$w = \frac{FL^3}{3\text{EI}} + \frac{qL^4}{8\text{EI}}, \tag{1}$$

$$\sigma = \frac{FL}{W} + \frac{qL^2}{2W}, \tag{2}$$

where the cross-section characteristics are as follows:

$$I_y = 2\left[\frac{dh^3}{12} + \frac{bt^3}{12} + \left(\frac{h+2t}{2}\right)^2 tb\right], \tag{3}$$

$$W_y = \frac{I_y}{2(h+2t)}. \tag{4}$$

Furthermore, continuous load due to the cantilever weight is:

$$q = g \cdot A \cdot \rho, \tag{5}$$

where g = 9.81 N/s², A is the cantilever cross-section area, and ρ = 7850 kg/m³ is the density of steel.

Comparison of results is given in Tab. 2 and Fig. 10. Obtained results have shown that both deformations and stresses were lower for numerical solutions compared to results obtained via analytical equations. It is reasoned that this is a result of the high stiffness of the FEM model, indicating that the mesh density should be increased.

Table 2 Comparison of stress and deformation obtained for optimum solutions using FEM (σ_{op}, w_{op}) and analytical equations (σ_{an}, w_{an})

	σ_{op}	w_{op}	σ_{an}	w_{an}
1	44,99	13,26	23,33	8,92
2	30,92	12,70	20,01	11,67
3	26,87	9,56	17,17	5,54
4	46,02	10,00	15,08	4,96
5	28,33	12,48	22,75	15,90
6	55,68	5,87	15,07	2,42
7	65,65	5,39	16,92	3,50
8	67,07	6,42	16,41	4,06
9	73,23	3,29	12,79	0,56
10	73,05	9,83	22,88	10,70
11	67,26	9,86	17,85	8,36
12	63,37	13,18	23,56	20,34
13	46,49	3,34	13,03	2,75
14	68,02	3,17	26,27	1,79
15	65,00	3,28	37,70	3,15
16	35,55	1,51	23,59	1,99
17	60,91	14,89	24,70	16,03
18	33,62	1,49	4,28	0,09

Correlation factors between numerical and analytical results are 0,88 and 0,23 for deformation and stress, respectively. It may be reasoned that maximum stress is a single value, i.e. a value of a single finite element, which may, particularly in the case of inner angles be unrealistic, i.e. the value may simply refer to the stress concentration rather than realistic value. On the contrary, maximum deformation is a sum of deformations of many individual elements where a single unrealistic value will have an insignificant effect on the overall value. Again, increasing the mesh density may partially resolve this problem, albeit values that are more realistic may also necessitate redefinition of boundary conditions.

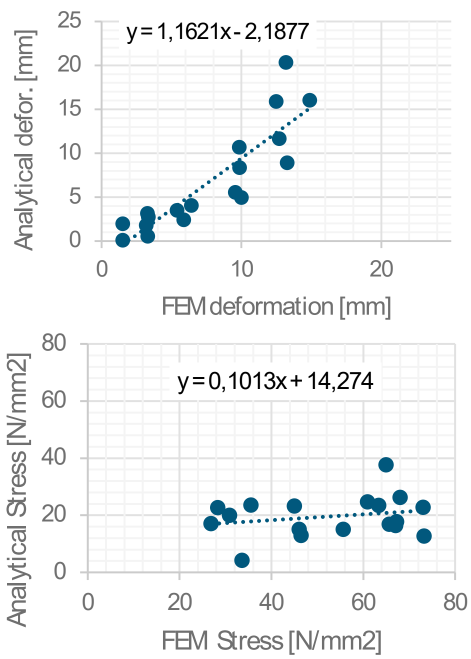


Figure 10 Comparison of deformation and stress using analytical and FEM methods

5 CONCLUSIONS

The paper presents a proposal on the approach of development of guidelines for designing a cantilever for a typical set of parameters occurring in steel structures. The application of Ansys optimization parametric tools facilitated a fast and consistent tool to generate optimum solutions for a wide range of design parameters.

It should be noted that the applied geometry is typically related to the welded steel sheet design, however, similar geometry may be obtained using different technologies including extrusion and additive manufacturing, hence developed guidelines may enable simple straightforward means of designing near optimized cantilever design elements resulting in reduced requirements for material, energy, transportation, and assembly requirements.

Analysis of obtained results, and particularly comparison of numerical and analytical results of calculated stresses and deformation show significant differences. Results point to the conclusion that the numerical analysis should be further developed and detailed to ensure that it provides reliable values. Hence, the guidelines obtained herein are not suitable for general application, and a more detailed procedure for the development of design guidelines should be developed. However, obtained results show that a functional relationship between design requirements consisting of load and reach of the cantilever and values of cross-section dimensions can be defined, i.e. that the proposed approach has merit and that investing additional resources, particularly computational power, should results in obtaining functional guidelines for designing near optimum cantilever design.

6 REFERENCES

[1] Zamanian, M. & Firouzi, B. (2024). Analysis and Optimal Design of Vibration-Based Paddle Type Piezoelectric Energy Harvester Under Electrostatic Actuation. *J. Vib. Eng. Technol.* https://doi.org/10.1007/s42417-024-01280-9

[2] Lew A. J. & Buehler, M. J. (2021). Encoding and exploring latent design space of optimal material structures via a VAE-LSTM model. *Forces in Mechanics, 5*, 100054. https://doi.org/10.1016/j.finmec.2021.100054

[3] Qiu, C., Du, S. & Yang, J. (2021). A deep learning approach for efficient topology optimization based on the element removal strategy. *Materials & Design, 212*, 110179. https://doi.org/10.1016/j.matdes.2021.110179

[4] Rostami Najafabadi, H., Martins, T. C., Tsuzuki, M. S. G. & Barari, A. (2004). Structural Design with Self-Weight and Inertial Loading Using Simulated Annealing for Non-Gradient Topology Optimization. *Machines, 12*(1), 25. https://doi.org/10.3390/machines12010025

[5] Liu, C., Mao, B., Huang, G., Wu, Q., Xie S. & Xu, M. (2018). Optimization of shape control of a cantilever beam using dielectric elastomer actuators. *AIP Advances, 8*(5), 055015. https://doi.org/10.1063/1.5026160

[6] Huang, Z., Huang, F., Wang, X., Chu, F. (2023). Active Vibration Control of Composite Cantilever Beams. *Materials, 16*(1), 95. https://doi.org/10.3390/ma16010095

[7] Singh S., Pflug L., Mergheim J. & Stingl M. (2023). Robust design optimization for enhancing delamination resistance of composites. *Int J Numer Methods Eng, 124*(6), 1381-1404. https://doi.org/10.1002/nme.7168

[8] Cui Y., Hong Y., Khan N. A. & Sulaiman M. (2021). Application of Soft Computing Paradigm to Large Deformation Analysis of Cantilever Beam under Point Load. *Complexity, 2021*, Article ID 2182693. https://doi.org/10.1155/2021/2182693

[9] Sargent, B. S., Ynchausti, C. R., Nelson, T. G. & Howell, L. L. (2022). The Mixed-Body Model: A Method for Predicting Large Deflections in Stepped Cantilever Beams. *ASME. J. Mechanisms Robotics, 14*(4), 041001. https://doi.org/10.1115/1.4053376

[10] Den Hartog, J. P. (2012). *Strength of Materials*. Courier Corporation.

Authors' contacts:

Filip Šulentić
University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture,
Ivana Lučića 5, 10000 Zagreb, Croatia
fs220977@stud.fsb.hr

Daniel Miler, PhD
University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture,
Ivana Lučića 5, 10000 Zagreb, Croatia
daniel.miler@fsb.unizg.hr

Matija Hoić, PhD
(Corresponding author)
University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture,
Ivana Lučića 5, 10000 Zagreb, Croatia
matija.hoic@fsb.unizg.hr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20...........

Categorization of paper
Received: y-m-d, Accepted: y1-m1-d1

14pt

14pt

Article Title Only in English (Style: Arial Narrow, Bold, 14pt)

14pt

Ivan Horvat, Thomas Johnson, Marko Marić (Style: Arial Narrow, Normal, 10pt)

14pt

Abstract: Article abstract contains maximum of 150 words and is written in the language of the article. The abstract should reflect the content of the article as precisely as possible. TECHNICAL JOURNAL is a trade journal that publishes scientific and professional papers from the domain(s) of mechanical engineering, electrical engineering, civil engineering, multimedia, logistics, etc., and their boundary areas. This document must be used as the template for writing articles so that all the articles have the same layout. (Style: Arial Narraw, 8pt)

8pt

Keywords: keywords in alphabetical order (5-6 key words). Keywords are generally taken from the article title and/or from the abstract. (Style: Arial Narraw, 8pt)

10pt

10pt

1 INTRODUCTION (Article Design)
(Style: Arial Narrow, Bold, 10pt)

10pt

(Tab 6 mm) The article is written in Latin script and Greek symbols can be used for labelling. The length of the article is limited to eight pages of international paper size of Letter (in accordance with the template with all the tables and figures included). When formatting the text the syllabification option is not to be used.

10pt

1.1 Subtitle 1 (Writing Instructions)
(Style: Arial Narrow, 10pt, Bold, Align Left)

10pt

The document format is Letter with margins in accordance with the template. A two column layout is used with the column spacing of 10 mm. The running text is written in Times New Roman with single line spacing, font size 10 pt, alignment justified.

Article title must clearly reflect the issues covered by the article (it should not contain more than 15 words).

Body of the text is divided into chapters and the chapters are divided into subchapters, if needed. Chapters are numbered with Arabic numerals (followed by a period). Subchapters, as a part of a chapter, are marked with two Arabic numerals i.e. 1.1, 1.2, 1.3, etc. Subchapters can be divided into even smaller units that are marked with three Arabic numerals i.e. 1.1.1, 1.1.2, etc. Further divisions are not to be made.

Titles of chapters are written in capital letters (uppercase) and are aligned in the centre. The titles of subchapters (and smaller units) are written in small letters (lowercase) and are aligned left. If the text in the title of the subchapter is longer than one line, no hanging indents.

10pt

Typographical symbols (bullets), which are being used for marking an item in a list or for enumeration, are placed at a beginning of a line. There is a spacing of 10pt following the last item:

- Item 1
- Item 2
- Item 3

10pt

The same rule is valid when items are numbered in a list:

1) Item 1
2) Item 2
3) Item 3

10pt

1.2 Formatting of Pictures, Tables and Equations
(Style: Arial Narrow, 10pt, Bold, Align Left)

10pt

Figures (drawings, diagrams, photographs) that are part of the content are embedded into the article and aligned in the centre. In order for the figure to always be in the same position in relation to the text, the following settings should be defined when importing it: text wrapping / in line with text.

Pictures must be formatted for graphic reproduction with minimal resolution of 300 dpi. Pictures downloaded from the internet in ratio 1:1 are not suitable for print reproduction because of unsatisfying quality.

10pt

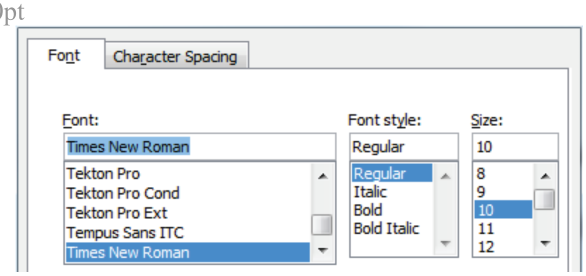


Figure 1 Text under the figure [1]
(Style: Arial Narraw, 8pt, Align Centre)

10pt

The journal is printed in black ink and the figures have to be prepared accordingly so that bright tones are printed in a satisfactory manner and are readable. Figures are to be in colour for the purpose of digital format publishing. Figures in the article are numbered with Arabic numerals (followed by a period).

Text and other data in tables are formatted - Times New Roman, 8pt, Normal, Align Center.

When describing figures and tables, physical units and their factors are written in italics with Latin or Greek letters, while the measuring values and numbers are written upright.

10pt

Table 1 Table title aligned centre
(Style: Arial Narraw, 8pt, Align Centre)

	1	2	3	4	5	6
ABC	ab	ab	ab	ab	ab	ab
DEF	cd	cd	cd	cd	cd	cd
GHI	ef	ef	ef	ef	ef	ef

10 pt

Equations in the text are numbered with Arabic numerals inside the round brackets on the right side of the text. Inside the text they are referred to with equation number inside the round brackets i.e. ".... from Eq. (5) follows ...." (Create equations with MathType Equation Editor - some examples are given below).

10pt

$$F_{\text{avg}}(t,t_0)=\frac{1}{t}\int_{t_0}^{t_0+t}F\left[q(\tau),p(\tau)\right]\mathrm{d}\tau, \tag{1}$$

$$\cos\alpha+\cos\beta=2\cos\frac{\alpha+\beta}{2}\cdot\cos\frac{\alpha-\beta}{2}, \tag{2}$$

$$(\boldsymbol{AB})^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}. \tag{3}$$

10pt

Variables that are used in equations and also in the text or tables of the article are formatted as *italics* in the same font size as the text.

10pt

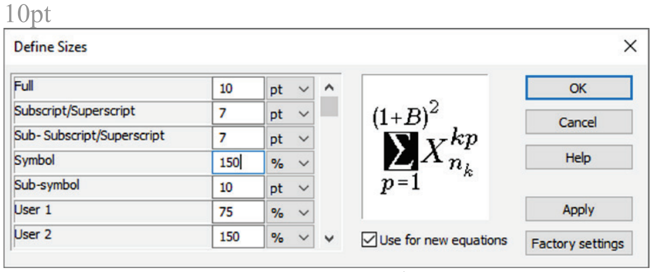


Figure 2 The texts under figures
(Style: Arial Narraw, 8pt, Align Centre)

10pt

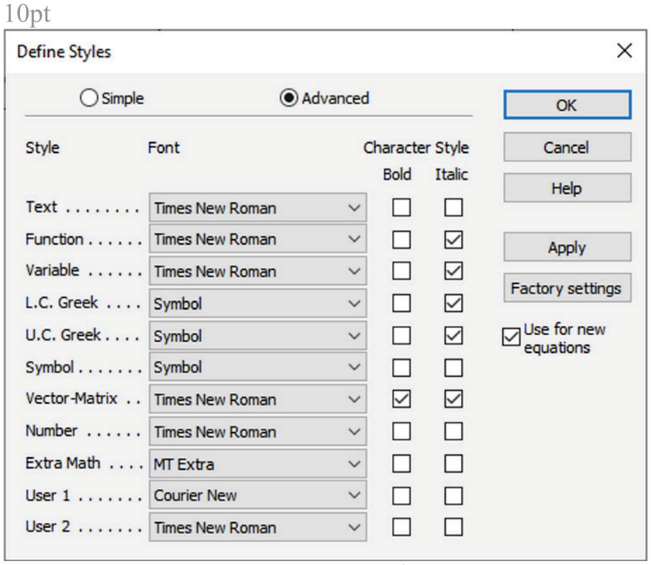


Figure 3 The texts under figures
(Style: Arial Narraw, 8pt, Align Centre)

Figures and tables that are a part of the article have to be mentioned inside the text and thus connected to the content i.e. " … as shown in Fig. 1…" or "data from Tab. 1…" and similar.

10pt

2 PRELIMINARY ANNOTATION

10pt

Article that is offered for publication cannot be published beforehand, be it in the same or similar form, and it cannot be offered at the same time to a different journal. Author or authors are solely responsible for the content of the article and the authenticity of information and statements written in the article.

Articles that are accepted for publishing are classified into four categories: original scientific papers, preliminary communications, subject reviews and professional papers.

Original scientific papers are articles that according to the reviewer and the editorial board contain original theoretical or practical results of research. These articles need to be written in such a way that based on the information given, the experiment can be repeated and the results described can be achieved together with the author's observations, theoretical statements or measurements.

Preliminary communication contains one or more pieces of new scientific information, but without details that allow recollection as in original scientific papers. Preliminary communication can give results of an experimental research, results of a shorter research or research in progress that is deemed useful for publishing.

Subject review contains a complete depiction of conditions and tendencies of a specific domain of theory, technology or application. Articles in this category have an overview character with a critical review and evaluation. Cited literature must be complete enough to allow a good insight and comprehension of the depicted domain.

Professional paper can contain a description of an original solution to a device, assembly or instrument, depiction of important practical solutions, and similar. The article need not be related to the original research, but it should contains a contribution to an application of known scientific results and their adaptation to practical needs, so it presents a contribution to spreading knowledge, etc.

Outside the mentioned categorization, the Editorial board of the journal will publish articles of interesting content in a special column. These articles provide descriptions of practical implementation and solutions from the area of production, experiences from device application, and similar.

10pt

3 WRITING AN ARTICLE

10pt

Article is written in the English language and the terminology and the measurement system should be adjusted to legal regulations, standards and the International System of Units (SI) (Quantities and Units: ISO 80 000 - from Part 1 to Part 14). The article should be written in third person.

Introduction contains the depiction of the problem and an account of important results that come from the articles that are listed in the cited literature.

Main section of the article can be divided into several parts or chapters. Mathematical statements that obstruct the reading of the article should be avoided. Mathematical statements that cannot be avoided can be written as one or more addendums, when needed. It is recommended to use an example when an experiment procedure, the use of the work in a concrete situation or an algorithm of the suggested method must be illustrated. In general, an analysis should be experimentally confirmed.

Conclusion is a part of the article where the results are being given and efficiency of the procedure used is emphasized. Possible procedure and domain constraints where the obtained results can be applied should be emphasized.

AI and AI-assisted tools do not qualify for authorship under TG/TJ's authorship policy. Authors who use AI or AI-assisted tools during the manuscript writing process are asked to disclose their use in a separate section of the manuscript. The publishing agreement process works as usual, with the authors keeping the copyright to their own work.

10pt

4 RECAPITULATION ANNOTATION

10pt

In order for the articles to be formatted in the same manner as in this template, this document is recommended for use when writing the article. Finished articles written in MS Word for Windows and formatted according to this template must be submitted using our The Paper Submission Tool (PST) (https://tehnickiglasnik.unin.hr/authors.php) or eventually sent to the Editorial board of the Technical Journal to the following e-mail address: tehnickiglasnik@unin.hr

The editorial board reserves the right to minor redaction corrections of the article within the framework of prepress procedures. Articles that in any way do not follow these authors' instructions will be returned to the author by the editorial board. Should any questions arise, the editorial board contacts only the first author and accepts only the reflections given by the first author.

10pt

5 REFERENCES (According to APA)

10pt

The literature is cited in the order it is used in the article. No more than 35 references are recommended. Individual references from the listed literature inside the text are addressed with the corresponding number inside square brackets i.e. "… in [7] is shown …". If the literature references are web links, the hyperlink is to be removed as shown with the reference number 8. Also, the hyperlinks from the e-mail addresses of the authors are to be removed. In the literature list, each unit is marked with a number and listed according to the following examples (omit the subtitles over the references – they are here only to show possible types of references):

9pt

[1] See http://www.bibme.org/citation-guide/apa/
[2] See http://sites.umuc.edu/library/libhow/apa_examples.cfm
[3] (Style: Times New Roman, 9pt, according to APA)
[4] Amidzic, O., Riehle, H. J. & Elbert, T. (2006). Toward a psychophysiology of expertise: Focal magnetic gamma bursts as a signature of memory chunks and the aptitude of chess players. *Journal of Psychophysiology, 20*(4), 253-258. https://doi.org/10.1027/0269-8803.20.4.253
[5] Reitzes, D. C. & Mutran, E. J. (2004). The transition to retirement: Stages and factors that influence retirement adjustment. *International Journal of Aging and Human Development, 59*(1), 63-84. Retrieved from http://www.baywood.com/journals/PreviewJournals.asp?Id=0091-4150
[6] Jans, N. (1993). *The last light breaking: Life among Alaska's Inupiat Eskimos*. Anchorage, AK: Alaska Northwest Books.
[7] Miller, J. & Smith, T. (Eds.). (1996). *Cape Cod stories: Tales from Cape Cod, Nantucket, and Martha's Vineyard*. San Francisco, CA: Chronicle Books.
[8] Chaffe-Stengel, P. & Stengel, D. (2012). *Working with sample data: Exploration and inference*. https://doi.org/10.4128/9781606492147
[9] Freitas, N. (2015, January 6). People around the world are voluntarily submitting to China's Great Firewall. Why? Retrieved from http://www.slate.com/blogs/future_tense/2015/01/06/tencent_s_wechat_worldwide_internet_users_are_voluntarily_submitting_to.html
(Style: Times New Roman, 9pt, according to APA)

10pt

10pt

Authors' contacts:

8pt

Full Name, title
Institution, company
Address
Tel./Fax, e-mail

8pt

Full Name, title
Institution, company
Address
Tel./Fax, e-mail

Note: Gray text should be removed in the final version of the article because it is for guidance only.


WIT CONFERENCES
Call for Papers
HPSM/OPTI
2025
12th International Conference on High Performance and Optimum Design of Structures and Materials
10 - 12 June 2025 | Edinburgh, UK
Organised by
Wessex Institute, UK
University of A Coruña, Spain
Sponsored by
WIT Transactions on the Built Environment
www.witconferences.com/hpsm2025

ERES
2025
WIT CONFERENCES
Call for Papers
KEEP RIGHT
14th International Conference on Earthquake Resistant Engineering Structures
11 – 13 June 2025 | Edinburgh, UK
Organised by
Wessex Institute, UK
Sponsored by
WIT Transactions on the Built Environment
www.witconferences.com/eres2025

TEHNIČKI GLASNIK / TECHNICAL JOURNAL – GODIŠTE / VOLUME 18 – BROJ / ISSUE 3

RUJAN 2024 / SEPTEMBER 2024 – STRANICA / PAGES 337-506

SVEUČILIŠTE SJEVER / UNIVERSITY NORTH – CROATIA – EUROPE

ISSN 1846-6168 (PRINT) / ISSN 1848-5588 (ONLINE)

TEHNICKIGLASNIK@UNIN.HR – HTTP://TEHNICKIGLASNIK.UNIN.HR