ISSN 1854-6250

APEM *journal*

Advances in Production Engineering & Management

Volume 14 | Number 2 | June 2019

University of Maribor

Published by CPE
apem-journal.org

Advances in Production Engineering & Management

Identification Statement

APEM journal

ISSN 1854-6250 | Abbreviated key title: **Adv produc engineer manag** | Start year: 2006
ISSN 1855-6531 (on-line)

Published quarterly by **Chair of Production Engineering (CPE)**, **University of Maribor**
Smetanova ulica 17, SI – 2000 Maribor, Slovenia, European Union (EU)

Phone: 00386 2 2207522, Fax: 00386 2 2207990
Language of text: English

University of Maribor

APEM homepage: apem-journal.org
University homepage: www.um.si

APEM Editorial

Subscription Rate: 120 EUR for 4 issues (worldwide postage included); 30 EUR for single copies (plus 10 EUR for postage); for details about payment please contact: info@apem-journal.org
Cover and interior design: Miran Brezocnik
Printed: Tiskarna Koštomaj, Celje, Slovenia
Subsidizer: The journal is subsidized by Slovenian Research Agency

Advances in Production Engineering & Management is indexed and abstracted in the **WEB OF SCIENCE** (maintained by **Clarivate Analytics**): **Science Citation Index Expanded**, **Journal Citation Reports** – Science Edition, **Current Contents** – Engineering, Computing and Technology • **Scopus** (maintained by **Elsevier**) • **Inspec** • **EBSCO**: Academic Search Alumni Edition, Academic Search Complete, Academic Search Elite, Academic Search Premier, Engineering Source, Sales & Marketing Source, TOC Premier • **ProQuest**: CSA Engineering Research Database – Cambridge Scientific Abstracts, Materials Business File, Materials Research Database, Mechanical & Transportation Engineering Abstracts, ProQuest SciTech Collection • **TEMA** (**DOMA**) • The journal is listed in **Ulrich's** Periodicals Directory and **Cabell's** Directory

University of Maribor
Chair of Production Engineering (CPE)

Advances in Production Engineering & Management

Volume 14 | Number 2 | June 2019 | pp 139–266

Contents

Journal homepage: apem-journal.org

ISSN 1854-6250 (print)
ISSN 1855-6531 (on-line)

Scope and topics

Advances in Production Engineering & Management (*APEM journal*) is an interdisciplinary refereed international academic journal published quarterly by the *Chair of Production Engineering* at the *University of Maribor*. The main goal of the *APEM journal* is to present original, high quality, theoretical and application-oriented research developments in all areas of production engineering and production management to a broad audience of academics and practitioners. In order to bridge the gap between theory and practice, applications based on advanced theory and case studies are particularly welcome. For theoretical papers, their originality and research contributions are the main factors in the evaluation process. General approaches, formalisms, algorithms or techniques should be illustrated with significant applications that demonstrate their applicability to real-world problems. Although the *APEM journal* main goal is to publish original research papers, review articles and professional papers are occasionally published.

Fields of interest include, but are not limited to:

Additive Manufacturing Processes
Advanced Production Technologies
Artificial Intelligence in Production
Assembly Systems
Automation
Big Data in Production
Computer-Integrated Manufacturing
Cutting and Forming Processes
Decision Support Systems
Deep Learning in Manufacturing
Discrete Systems and Methodology
e-Manufacturing
Evolutionary Computation in Production
Fuzzy Systems
Human Factor Engineering, Ergonomics
Industrial Engineering
Industrial Processes
Industrial Robotics
Intelligent Manufacturing Systems
Joining Processes
Knowledge Management
Logistics in Production
Machine Learning in Production
Machine Tools
Machining Systems
Manufacturing Systems
Materials Science, Multidisciplinary
Mechanical Engineering
Mechatronics
Metrology in Production
Modelling and Simulation
Numerical Techniques
Operations Research
Operations Planning, Scheduling and Control
Optimisation Techniques
Project Management
Quality Management
Risk and Uncertainty
Self-Organizing Systems
Statistical Methods
Supply Chain Management
Virtual Reality in Production

APEM journal

Advances in Production Engineering & Management
Volume 14 | Number 2 | June 2019 | pp 143–152
https://doi.org/10.14743/apem2019.2.317

ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

Effect of process parameters on cutting speed of wire EDM process in machining HSLA steel with cryogenic treated brass wire

Tahir, W.[a,*]**, Jahanzaib, M.**[a]**, Raza, A.**[b]

[a]Industrial Engineering Department, University of Engineering and Technology, Taxila, Pakistan
[b]Industrial Technology Development, National Centre of Physics, Islamabad, Pakistan

A B S T R A C T

Wire electrical discharge machining (wire EDM), a most common non-conventional machine tool, is extensively employed to produce precise, delicate and intricate profiled shaped parts especially from hard to machine materials. The performance of wire EDM is mainly based on the electrical conductivity of both electrode wires and workpiece materials. The aim of research is to increase cutting speed (*CS*) of high strength low alloy (HSLA) hardened steel by determining main contributing input process parameters and effect of cold treatment on electrical conductivity of brass wire at –70 °C. Fractional factorial design is used to determine the relationship of *CS* with input process parameters includes; open voltage, pulse on time, pulse off time, wire tension, flushing pressure of deionized water and brass wires (cold treated – CT, and non-cold treated – NCT). Empirical model for *CS* is developed based on selected input process parameters and their contribution is analyzed through ANOVA technique. It is learned that pulse on time, pulse off time and wire electrode are the main contributing input process parameters that provide assistance to increase *CS* of wire EDM. In wire electrodes, cold treated brass wire is observed as a best alternative to enhance machining performance with an increase of electrical conductivity by 24.5 %.

A R T I C L E I N F O

Keywords:
Wire electrical discharge machining (WEDM);
HSLA steel;
Brass wire;
Cryogenic treatment;
Cutting speed;
Process parameters

**Corresponding author:*
waseem.tahir@uettaxila.edu.pk
(Tahir, W.)

Article history:
Received 30 October 2018
Revised 11 March 2019
Accepted 12 April 2019

1. Introduction

Electric discharge machining (EDM) principle based on erosion phenomena in which high frequency electric sparks removes material of conductive alloys [1]. Inadequacy of conventional machining processes for the machining of intricate and aero profiled shapes of extremely hard materials compels the manufacturer to use non-conventional machine tools. Based on the principle of electric discharge machining, WEDM used wire as cutting tool for machining of intricate profiles with high accuracy [2]. Nowadays, non-conventional machining processes are getting more attention of manufacturer due to less burr formation, residual stresses and low tooling cost [3]. Tools, molds and dies having complex shapes which are mostly used in automobile, aerospace and surgical industries are commonly manufactured by EDM [4].

In electric discharge machining, a series of electric sparks are produced between workpiece and electrode. Current is discharged from the electrode to the workpiece in the presence of dielectric fluid with a very little spark gap in the range of 1/1,000,000 of a second or less [5]. The heat of each electric spark (8,300-11,650 °C), erodes workpiece material in the form of tiny bits that are vaporized and melted from the workpiece surface. Deionized water is mostly used as dielectric fluid, also flushed away debris from tool and workpiece interface. It is noticed that process time of machining is reduced with increase in pulse on time and spark current intensity as compare to pulse off duration [6, 7]. Usually the tools used in WEDM process are tungsten, molybdenum and brass wire. Wire electrodes used in WEDM for cutting having diameter ranges from 0.05-0.3 mm [8].

2. Literature review

In highly competitive and globalized environment, it is necessary for industrial practitioner to work on value added manufacturing techniques to deliver quality product to the customer well in time. It has been seen that in electric discharge machining the properties of electrode significantly affect the machining performance [9]. Different treatments are being developed to improve the mechanical and electrical properties of cutting tools to reduce manufacturing time. It includes also sub-zero treatments of cutting tools other than thermal and coating techniques. A substantial improvements have been seen with sub-zero treatments in physical (wear resistance, hardness, tensile strength etc.), electrical and thermal properties of cutting tools as compared to thermal and coating techniques [10].

Sub-zero or cryogenic treatment is a low temperature processing normally used to alter physical and electrical properties of materials includes wear resistance, electrical conductivity, toughness, dimensional stability etc. Sub-zero treatments with feasible lower most temperature has been classified in three different categories: Cold treatment with a temperature range of 223-193K, shallow cryogenic treatment ranges from 193-113 K and deep cryogenic treatment ranges from 113-77 K [11]. Cutting tools of conventional machining (turning, milling, drilling etc.) have been cryogenically treated to enhance wear resistance, hardness, toughness for better cutting performance.

A limited research has been reported on the cryogenic treatment of wire electrode in WEDM process. A detailed literature study carried out by Kumar *et al.* [10] which clearly depicts the benefits of sub-zero treatment of electrode and workpiece in EDM that are presented in Table 1 and Table 2 respectively. The effects of sub-zero treatment have been analysed on the basis of material removal rate *MRR*, tool wear rate *TWR*, and surface roughness *SR*. Brass wire which mainly consists of zinc (63-65 %) and copper (35-37 %), provide helps to deliver more usable energy to the machined surface area by electric discharges [17]. Addition of zinc contents contribute to increase tensile strength and vapour pressure ratio on account of decrease in electrical conductivity [18].

Sharma *et al.* [8] found that pulse on time T_{on} and pulse off time T_{off} are the main contributing factors for cutting speed in WEDM process for HSLA steel. Maximum *MRR* has been reported while machining at optimum range of T_{on} (3-4μs) and T_{off} (14-16 μs) using Response surface methodology for high speed steel (HRC 62), [25]. Moly wire has been used for the machining of hardened HSLA steel (30CrMnSiA) and found that higher cutting speed *CS* obtained at lower pulse frequency and T_{off} with higher value of power [26]. Selvakumar *et al.* [27] observed that T_{on} and peak current are the most significant input parameters for cutting speed in trim cut while using aluminium AA5083 alloy. Bhuyan *et al.* [28] used Central composite design to investigate the effects of peak current, pulse on time T_{on} and flushing pressure F_p on material removal rate. Multi objective optimization of experimental investigation has been performed using overall evaluation criteria, entropy weight measurement and fuzzy logic techniques.

Singh *et al.* [29] compared zinc coated cryogenically treated brass wire with simple cryogenically treated brass wire as electrode to investigate their effects on cutting performance of AISI D3 die steel. Their results indicate that zinc coated cryogenically treated brass wire yields better *MRR* using Taguchi L9 array for experimental design. Kapoor *et al.* [16] also used deep cryogenically (–184 °C) treated brass wire and simple brass wire as electrode. Taguchi design has been used to investigate the effects of process parameters on *MRR*. Wire type (cryogenic treated wire) has been observed as most effective input parameter followed by pulse width, pulse duration and wire tension.

Table 1 Summary of cryogenic treated electrodes in non-conventional machining [6]

First author [ref.] (Year)	Electrode Tool	Work material	No. of samples	Non-conventional machining	Key findings		
					TWR	*MRR*	*SR*
Sundaram [12] (2009)	Copper	Be-Cu	16	EDD	NC	↑	NR
Kumar [13] (2012)	Copper	Inconel 718	18	Additive Mix EDM	NR	↑	NR
Jafferson [14] (2013)	copper, brass, tungsten	AISI 304	NR	Micro EDM	↓	NR	NR
Sharma [15] (2015)	copper, brass, graphite	AISI D3	3	EDM	58 %↓	↑	↓
Kapoor [16] (2012)	Brass	En-31	9	EDM	NR	↑	NR

↑, increase; ↓, decrease; NC, no change; NR, not reported;

Table 2 Summary of cryogenic treated workpiece in non-conventional machining [6]

First author [ref.] (year)	Work piece material	Tool material	Number of samples	Non-conventional machining	Key findings		
					SR	*MRR*	*TWR*
Gill [19] (2010)	Ti-6246	Electrolyte copper	18	EDD	…	↑8.5 %	↓34.78 %
Yildiz [20] (2011)	Be-Cu	Copper electrode	2	EDM	…	↑30 %	…
Kumar [21] (2014)	Ti, Ti-6Al-4 V and Ti-5Al-2.5Sn	Copper, Copper–Chromium, Copper–Tungsten	1	EDM	…	↑	…
Jatti [22] (2014)	NiTi	Electrolyte copper	5	EDM	…	↑19 %	…
Khanna [23] (2016)	D3		27	WEDM	↓10.6 %	↓5.6 %	…
Goyal [24] (2017)	D2	Copper electrode	3	EDM	↓	↑	↓

↑, increase; ↓, decrease;

The present work studied the effect of cold treated brass wire on the machining performance of HSLA steel using Fractional factorial design 2_{VI}^{6-1}. Experiments are run using different combinations of input process parameters including open voltage, pulse on time, pulse off time, wire type (cold treated – CT, and non-cold treated – NCT), wire tension, flushing pressure of dielectric fluid. This experimentation will certainly useful for industry practitioners to improve productivity by increasing *CS* based on developed empirical models for both CT and NCT brass wires.

3. Materials and methods

This section consists of complete description of CT process for brass wire, testing procedure to measure the significant changes after CT process, Fractional factorial design to execute the WEDM process to analyze the effect of input variables on *CS* of HSLA hardened steel (50-51 HRC).

In the presented research, the following abbreviations are used:

WEDM, Wire EDM	wire electric discharge machining
HSLA	high strength low alloy
CT	cold treated
NCT	non cold treated
ANOVA	analysis of variance
CS	cutting speed
OV	open voltage
T_{on}	pulse on time
T_{off}	pulse off time
W_{t}	wire type
T_{w}	wire tension
F_{p}	flushing pressure
MRR	material removal rate
TWR	tool wear rate
SR	surface roughness

3.1 Cold treatment of brass wire

Temperature chamber (CTT-SC-7520-02FI) is used for the cold treatment of brass wire. The soaking process is carried out at –70 °C for 24 hours at a ramp rate of 2 °C/min. Universal testing machine (Sintech 65G) is used to measure the tensile strength of both CT and NCT brass wires as per ASTM E8M standard while electrical conductivity is measured by portable Kelvin Bridge tester. It is observed that after CT process, electrical conductivity of brass wire is increased by 24.8 % whereas tensile strength is reduced by 5.64 % as showed in Table 3.

Table 3 Brass wire properties

Properties	NCT wire	CT wire	% change	
Tensile Strength (MPa)	727	686	5.64	↓
Electrical Conductivity(S/m)	12.5×10^6	15.6×10^6	24.8	↑

↑, increase; ↓, decrease

3.2 Workpiece material

HSLA steel contains alloying elements as shown in Table 4. It includes Cr, Mn, Si which are responsible for its better strength, forming, impact toughness and corrosion resistant properties. High strength to weight ratio and corrosion resistance of these alloys is the main reason of being widely used in aerospace industry. The carbon content along with other constituent elements makes it a hardened steel alloy.

Spectromax-Ametek® is used to test chemical composition of workpiece material. Specified Index range and actual elemental composition of this steel are enumerated in Table 4. A plate with dimensions of $100 \times 200 \times 15$ mm^3 is hardened by quenching and tempering heat treatment process given in Table 5. After heat treatment, hardness of specimen plate is observed in range of 50-51 HRC measured by hardness tester (INDENTEC:6187.5LK) with diamond indenter of cone angle 120° using minor load of 10 kg and test load of 150 kg.

Table 4 Chemical composition of workpiece material

Comp.	C	Si	Mn	Cr	Cu	Ni	Mo	V	P
weight %	0.29	1.55	0.8	1.1	<0.25	<0.25	0.45	0.09	<.015

Table 5 Heat treatment cycle of HSLA steel

HT Process	Temp. (°C)	Soaking time (min)	Cooling medium
Quenching	920	60	Oil (25 °C)
Tempering	300	160	Air

3.3 Design of experiment

In design of experiment, a full factorial design is considered an appropriate design provide the information of all main effects and all level of interactions (two/three way or of higher orders). However, it seems difficult to run large number of experiments using full factorial design. Fractional factorial design is a reasonable option to evaluate the responses with large number of input parameters [30].

In the present study, Fractional factorial design is selected to evaluate the machining performance based on CS using process parameters OV, T_{on}, T_{off}, W_t, T_w and F_p. The ranges of input process parameters are as follows; OV (75-120) V, T_{on} (1-8) μs, T_{off} (10-48)μs, W_t (CT and NCT), T_w (4-10) g and F_p (3-7) l/min. In design matrix NCT wire is coded as –1 and CT wire as 1 as shown in Table 6. For experimentation, 25 mm length of test pieces is machined with CNC CHMER WEDM. The CT and NCT brass wires of 0.3 mm are used as electrode. Cutting speed is determined by using the expression $CS = L/T$ (mm/min), where L is the length of workpiece in mm and T is the time in min.

Table 6 Design matrix with response values

Run No.	Input process parameters						Response variable
	A: Open voltage (OV)	B: Pulse on time (T_{on})	C: Pulse of time (T_{off})	D: Wire type (W_t)	E: Wire tension (T_w)	F: Flushing pressure (F_p)	Cutting speed (CS)
	V	µs	µs	-	G	l/min	mm/min
1	120	8	48	1	4	3	2.4
2	75	8	48	1	10	3	1.8
3	75	1	48	-1	4	7	0.19
4	120	8	48	-1	10	3	1.66
5	120	8	10	1	10	3	3.5
6	75	8	10	-1	10	3	2.14
7	75	1	48	-1	10	3	0.16
8	75	1	10	1	10	3	1.02
9	120	8	10	-1	10	7	3.05
10	75	1	10	1	4	7	0.77
11	75	8	10	-1	4	7	2.14
12	75	8	10	1	4	3	2.89
13	75	1	10	-1	4	3	0.85
14	120	1	48	-1	10	7	0.42
15	120	1	48	-1	4	3	0.41
16	75	8	48	-1	4	3	1.03
17	120	1	10	-1	4	7	0.697
18	120	8	48	1	10	7	2.31
19	120	1	48	-1	10	3	0.566
20	75	1	48	1	10	7	0.4
21	120	8	48	-1	4	7	1.58
22	75	1	48	1	4	3	0.62
23	120	8	10	1	4	7	3.5
24	120	1	10	1	4	3	1.3
25	120	1	48	1	4	7	0.79
26	120	1	10	-1	10	3	0.697
27	75	1	10	-1	10	7	0.394
28	75	8	10	1	10	7	2.83
29	120	8	10	-1	4	3	2.8
30	75	8	48	-1	10	7	1.11
31	75	8	48	1	4	7	1.9
32	120	1	10	1	10	7	1.67

4. Results and discussion

4.1 Statistical modeling and analysis

In statistical analysis, developed regression model depicts the relationship between CS and input process parameters (OV, T_{on}, T_{off}, W_t, T_w and F_p). Mathematical model for CS in term of coded variables is represented in Eq. 1.

$$CS = 1.5 + 0.23 \cdot OV + 0.79 \cdot T_{on} - 0.39 \cdot T_{off} + 0.29 \cdot W_t + 0.1 \cdot T_w - 0.13 \cdot F_p + 0.078 \cdot OV \cdot T_{on} \quad (1)$$

The model for cutting speed is significant as its p-value is less than 0.05 shown in Table 7. Analysis of variance reveals that OV, T_{on}, T_{off} and W_t are significant parameters for cutting speed as they have p-value less than 0.05 whereas T_w and F_p impart little contributions. The value of R^2 is 0.984 which indicates that the developed model for CS is adequate. The predicted R^2 is 0.9646 which isclosed to adjusted R^2 of 0.9771. Adequacy precision ratio is 39.587 indicates an adequate signal as it is more than 4 that is desirable [9].

The Normal plot of residuals in Fig. 1(a) clearly shows errors are normally distributed as residuals which are closer to normal straight line with minor deviations. In Fig. 1(b), plot of residuals versus predicted values confirmed the statistical assumption of independence and constant variance are not varied. It almost reflects the same pattern and structure from left to right.

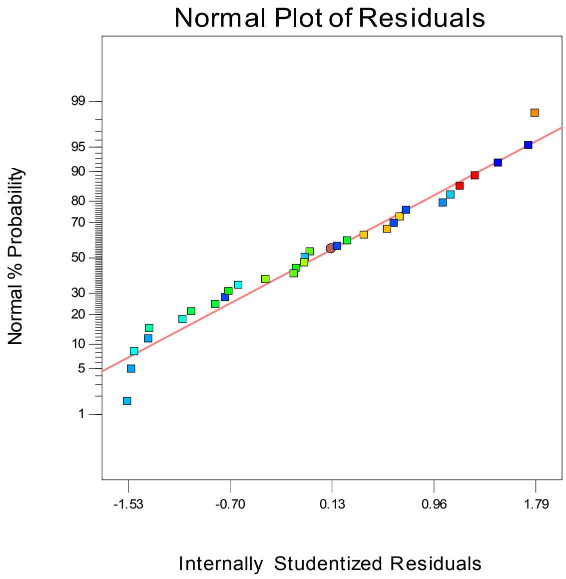
Table 7 ANOVA of model and process variables

Source	Sum of squares	Degree of freedom	Mean square	F value	p-value Prob > F	% Contribution
Model	30.35	9	3.37	143.13	< 0.0001	Significant
A: OV	1.65	1	1.65	69.92	< 0.0001	5.51
B: T_{on}	19.24	1	19.24	816.43	< 0.0001	64.34
C: T_{off}	4.73	1	4.73	200.9	< 0.0001	15.83
D: W_t	2.56	1	2.56	108.68	< 0.0001	8.56
E: T_w	9.45E-04	1	9.45E-04	0.04	0.8432	Not significant
F: F_p	5.02E-03	1	5.02E-03	0.21	0.6491	Not significant
AB	0.19	1	0.19	7.9	0.0105	Not significant
BC	0.92	1	0.92	38.88	< 0.0001	3.06
BD	0.12	1	0.12	4.99	0.0364	0.39
Residual	0.49	21	0.024			
Cor. total	30.85	30				
Std. Dev.		0.15	R^2		0.984	
Mean		1.52	Adj. R^2		0.9771	
C.V. %		10.12	Pred. R^2		0.9646	
PRESS		1.09	Adeq. Precision		39.587	

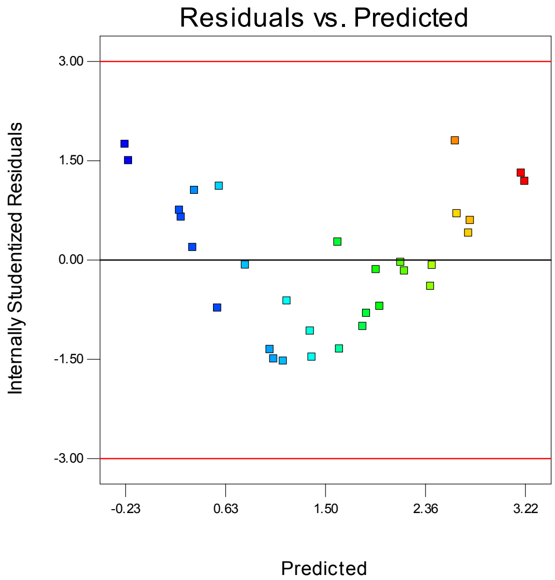
4.2 Effects of process parameters on cutting speed

In WEDM, mainly heat energy removes a very small portion of material by melting and evaporating workpiece material. Discharge process occurred several times in a second during pulse on time which erodes and vaporizes the material. High value of Ton substantially increases machine's *CS* as depicted in Fig. 2(a).

Conversely, high value of T_{off} a decrease in *CS* as shown in Fig. 2(b). Process of resolidification can be reduced by selecting minimum value of T_{off}. For higher production rates, lower value of T_{off} is desired. However, if T_{off} is too short, the eroded debris not properly contributes to reduce the deionization process of dielectric fluid.



(a)



(b)

Fig. 1 (a) Normal plot of residuals, (b) Residual vs. predicted

On the other hand, Fig. 2(c) shows an increase value of electrical conductivity of brass wire by CT process which significantly improves the *CS* with more powerful spark explosions. The effect of *OV* on *CS* has been presented in Fig. 2(d). High voltage produces more energetic pulses leads to increase the *CS*. However, water pressure and wire tension has shown no major contribution (as shown in Table 7) on *CS* as compare to other selected input parameters. Combined effect of significant parameters (*OV*, T_{on} and T_{off}) on *CS* are also considered through contour plots shown in Fig. 2. The Fig. 2(a) shows combined effects of T_{on} and T_{off} on cutting speed. This helps the practitioners to select the desirable value of *CS* by adjusting T_{on} and T_{off}. Contour plot of T_{on} and *OV* is shown in Fig. 2(b).

Contour lines with different *CS* values are shown in Fig. 3. Contour lines provide the option to choose different values of input parameters for the same value of *CS*. For example, in Fig. 3(a), a number of combinations on a similar contour line of T_{on} and T_{off} can be selected to achieve a *CS* of 1.7556 mm/min. Similarly, *CS* of 1.2860 mm/min can be achieved by selecting open voltage and pulse on time in Fig. 3(b).

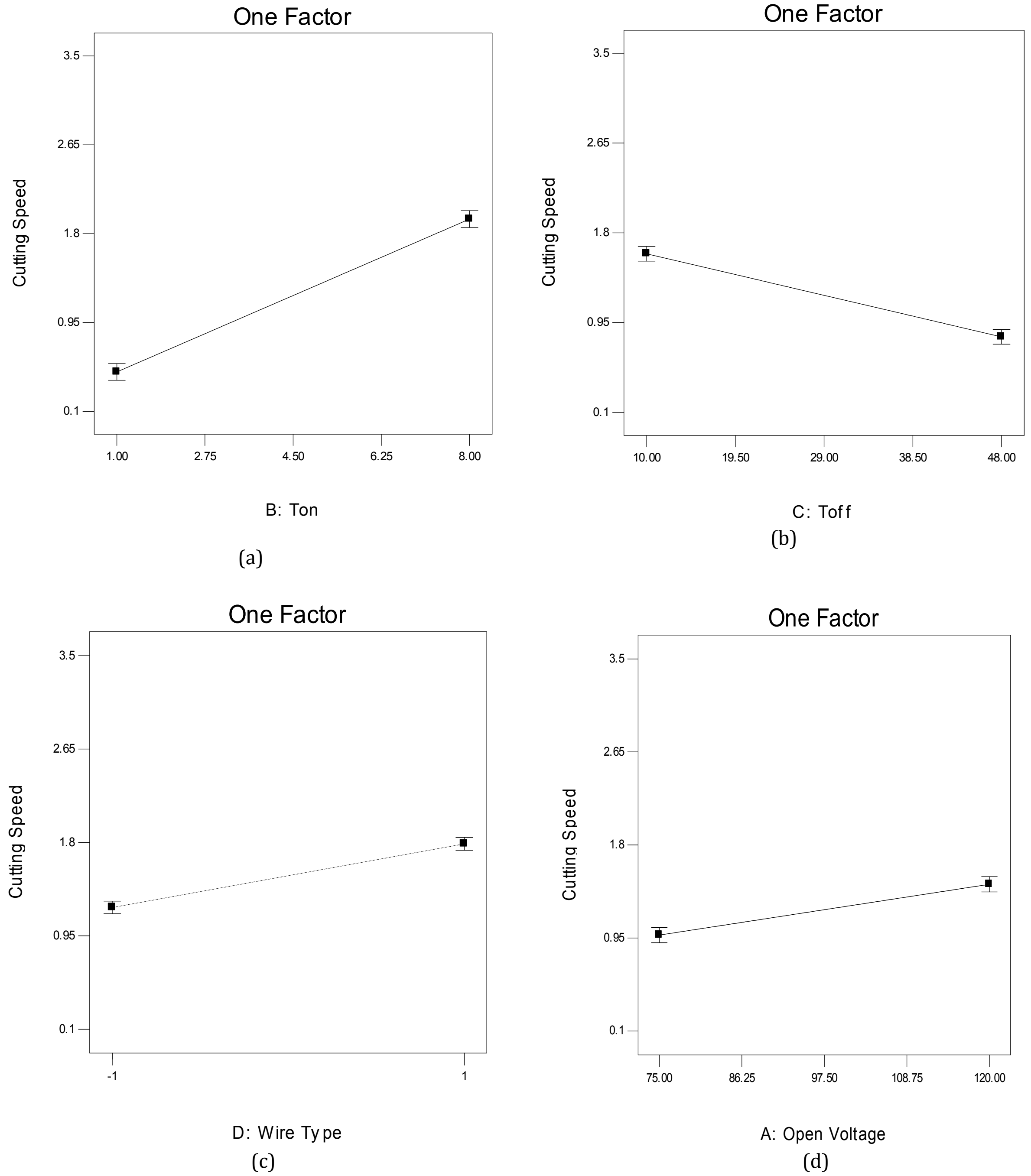
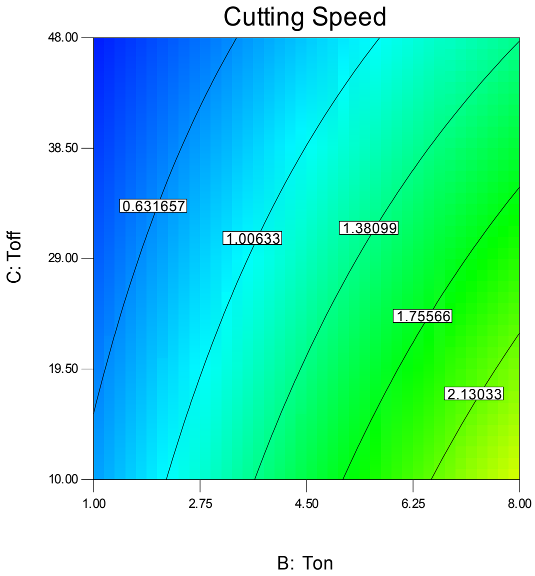
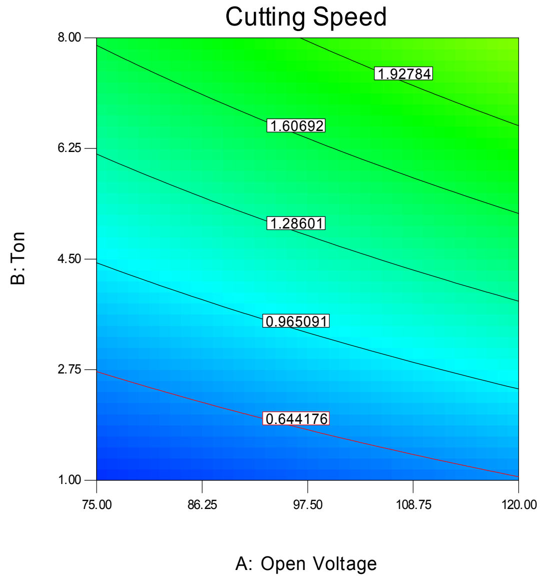


Fig. 2 (a) Effect of T_{on} on *CS* (b) Effect of T_{off} on *CS* (c) Effect of W_t on *CS* (d) Effect of *OV* on *CS*



(a)



(b)

Fig. 3 Contour plots for (a) T_{on} and T_{off} vs. *CS* (b) *OV* and T_{on} vs. *CS*

4.3 Validation tests

Additional experiments have been conducted to validate the statistical model mentioned in Eq. 2 and Eq. 3 for cutting speeds CS_{CT} and CS_{NCT} in case of CT and NCT wires, respectively.

$$CS_{CT} = 0.396 + 0.00585 \cdot OV + 0.223 \cdot T_{on} - 0.00899 \cdot T_{off} + 0.00185 \cdot T_w \quad (2)$$

$$CS_{NCT} = -0.0224 + 0.00585 \cdot OV + 0.187 \cdot T_{on} - 0.00899 \cdot T_{off} + 0.00185 \cdot T_w \quad (3)$$

Treatment combinations with predicted and actual response are presented in Table 8 which clearly shows that percentage error is less than 5 %. These validation runs satisfy the developed model as mentioned in Eq. 2 and Eq. 3 based on fractional factorial design. This model can be used as a reference for production of HSLA steel to determine the *CS* by using these input parameters.

Table 8 Experimentation confirmations

Trial No.	Open voltage	Pulse on time	Pulse off time	Wire type	Wire tension	Flushing pressure	Cutting speed		
							Predicted	Actual	% Error
1	110	7	15	NCT	6	5	2.26	2.157	4.6
2	110	7	15	CT	6	5	2.94	2.854	2.9
3	80	4	25	NCT	8	4	1.017	1.05	3.2
4	80	4	25	CT	8	4	1.5	1.559	3.9

5. Conclusion

In this study, an attempt is made to determine the effect of cold treated brass wire with other main contributing factors *OV*, T_{on} and T_{off} for the machining of HSLA at 51 HRC. From the present research following conclusions can be drawn:

- Improvement in wire conductivity is responsible for reduced machining time with an increase of electrical conductivity by 24.5 % by cold treatment process for 24 hours. However, it reduces the tensile strength by 3.6 %.
- Pulse on time and pulse off time are the main contributing factors for cutting speed with percentage contributions of 64.34 % and 15.83 % respectively.

- Contour plots provide assistance to select optimal process parameters with a simple and efficient way. Maximum cutting speed 2.1 mm/min can be achieved by setting T_{on} and T_{off} values in range of about 6.5-8.0μs and 10-20 μs respectively with the help of contour plot.

In future, both surface roughness and formation of recast layers on HSLA specimens can also be studied and analyzed by using multi objective approach.

Acknowledgement

Special thanks are addressed to Mr. Shahid Aslam, Dr. M. Azam, Dr. M. Amin and Mr. Muzamil Hussain for their suggestions and technical support for lab work at UET, Taxila especially for cryogenic treatment process.

References

[1] El-Hofy, H. (2005). *Advanced machining processes: Nontraditional and hybrid machining processes,* McGraw-Hill, New York, USA.
[2] Soni, H., Sannayellappa, N., Rangarasaiah, R.M. (2017). An experimental study of influence of wire electro discharge machining parameters on surface integrity of TiNiCo shape memory alloy, *Journal of Materials Research,* Vol. 32, No. 16, 3100-3108, doi: 10.1557/jmr.2017.137.
[3] Mohd Abbas, N., Solomon, D.G., Bahari, Md.F. (2007). A review on current research trends in electrical discharge machining (EDM), *International Journal of Machine Tools and Manufacture,* Vol. 47, No. 7-8, 1214-1228, doi: 10.1016/j.ijmachtools.2006.08.026.
[4] Chen, Z., Zhang, Y., Zhang, G., Huang, Y., Liu, C. (2017). Theoretical and experimental study of magnetic-assisted finish cutting ferromagnetic material in WEDM, *International Journal of Machine Tools and Manufacture,* Vol. 123, 36-47, doi: 10.1016/j.ijmachtools.2017.07.009.
[5] Singh, B., Misra, J.P. (2016). A critical review of wire electric discharge machining, In: Katalinic, B. (ed.), *DAAAM International Scientific Book 2016,* DAAAM International, Vienna, Austria, doi: 10.2507/daaam.scibook.2016.23.
[6] Jameson, E.C. (2001). *Electrical discharge machining*, Society of Manufacturing Engineers, Michigan, USA.
[7] Sommer, C., Sommer, S. (2005). *Complete EDM Handbook,* Advance Publishing, Texas, USA.
[8] Sharma, N., Khanna, R., Gupta, R.D., Sharma, R. (2013). Modeling and multiresponse optimization on WEDM for HSLA by RSM, *The International Journal of Advanced Manufacturing Technology,* Vol. 67, No. 9-12, 2269-2281, doi: 10.1007/s00170-012-4648-4.
[9] Raza, M.H., Wasim, A., Ali, M.A., Hussain, S., Jahanzaib, M. (2018). Investigating the effects of different electrodes on Al6061-SiC-7.5 wt% during electric discharge machining, *The International Journal of Advanced Manufacturing Technology,* Vol. 99, No. 9-12, 3017-3034, doi: 10.1007/s00170-018-2694-2.
[10] Singla, A.K., Singh, J., Sharma, V.S. (2018). Processing of materials at cryogenic temperature and its implications in manufacturing: A review, *Materials and Manufacturing Processes,* Vol. 33, No. 15, 1603-1640, doi: 10.1080/10426914.2018.1424908.
[11] Mohan Lal, D., Renganarayanan, S., Kalanidhi, A. (2001). Cryogenic treatment to augment wear resistance of tool and die steels, *Cryogenics,* Vol. 41, No. 3, 149-155, doi: 10.1016/S0011-2275(01)00065-0.
[12] Sundaram, M.M., Yildiz, Y., Rajurkar, K.P. (2009). Experimental study of the effect of cryogenic treatment on the performance of electro discharge machining, In: *ASME 2009 International Manufacturing Science and Engineering Conference,* West Lafayette, Indiana, USA, Vol. 2, 215-222, doi: 10.1115/MSEC2009-84247.
[13] Kumar, A., Maheshwari, S., Sharma, C., Beri, N. (2012). Machining efficiency evaluation of cryogenically treated copper electrode in additive mixed EDM, *Materials and Manufacturing Processes,* Vol. 27, No. 10, 1051-1058, doi: 10.1080/10426914.2011.654151.
[14] Jafferson, J.M., Hariharan, P. (2013). Machining performance of cryogenically treated electrodes in microelectric discharge machining: A comparative experimental study, *Materials and Manufacturing Processes,* Vol. 28, No. 4, 397-402, doi: 10.1080/10426914.2013.763955.
[15] Sharma, A., Grover, N.K., Singh, A. (2015). Effect of current and deep cryogenic treated EDM electrodes on surface roughness of AISI D3 steel, *Applied Mechanics and Materials,* Vol. 813-814, 521-525, doi: 10.4028/www.scientific.net/AMM.813-814.521.
[16] Kapoor, J., Singh, S., Khamba, J.S. (2012). Effect of cryogenic treated brass wire electrode on material removal rate in wire electrical discharge machining, *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science,* Vol. 226, No. 11, 2750-2758, doi: 10.1177/0954406212438804.
[17] Schacht, B. (2004). *Composite wire electrodes and alternative dielectric for wire electrical discharge machining, PhD thesis,* Katholieke Universityeit Leuven, Belgium.
[18] Kaneko, H., Onoue, M. (1984). *Electrode material for travelling-wire type electrical discharge machining,* US Patent No. 4424432.
[19] Gill, S.S., Singh, J. (2010). Effect of deep cryogenic treatment on machinability of titanium alloy (Ti-6246) in electric discharge drilling, *Materials and Manufacturing Processes,* Vol. 25, No. 6, 378-385, doi: 10.1080/10426910903179914.

[20] Yildiz, Y., Sundaram, M.M., Rajurkar, K.P., Nalbant, M. (2011). The effects of cold and cryogenic treatments on the machinability of beryllium-copper alloy in electro discharge machining, In: *Proceedings of 44th CIRP Conference on Manufacturing Systems,* Madison, Wisconsin, USA.

[21] Kumar, S., Batish, A., Singh, R., Singh, T.P. (2015). A mathematical model to predict material removal rate during electric discharge machining of cryogenically treated titanium alloys, *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture,* Vol. 229, No. 2, 214-228, doi: 10.1177/0954405414527955.

[22] Jatti, V.S., Singh, T.P. (2014). Effect of deep cryogenic treatment on machinability of NiTi shape memory alloys in electro discharge machining, *Applied Mechanics and Materials,* Vol. 592-594, 197-201, doi: 10.4028/www.scientific.net/AMM.592-594.197.

[23] Khanna, R., Singh, H. (2016). Comparison of optimized settings for cryogenic-treated and normal D-3 steel on WEDM using grey relational theory, *Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications,* Vol. 230, No. 1, 219-232, doi: 10.1177/1464420714565432.

[24] Goyal, R., Sngh, S., Kumar, H. (2018). Performance evaluation of cryogenically assisted electric discharge machining (CEDM) process, *Materials and Manufacturing Processes,* Vol. 33, No. 4, 433-443, doi: 10.1080/10426914.2017.1317789.

[25] Ishfaq, K., Mufti, N.A., Ahmad, J., Sajid, M., Jahanzaib, M. (2018). Analysis of the effect of wire electric dischage machining process parameters for the formation of high speed steel form tool, *Advances in Science and Technology Research Journal,* Vol. 12, No. 1, 89-98, doi: 10.12913/22998624/81001.

[26] Azam, M., Jahanzaib, M., Abbasi, J.A., Wasim, A. (2016). Modeling of cutting speed (CS) for HSLA steel in wire electrical discharge machining (WEDM) using moly wire, *Journal of the Chinese Institute of Engineers,* Vol. 39, No. 7, 802-808, doi: 10.1080/02533839.2016.1191377.

[27] Selvakumar, G., Thiruppathi Kuttalingam, K.G., Ram Prakash, S. (2018). Investigation on machining and surface characteristics of AA5083 for cryogenic applications by adopting trim cut in WEDM, *Journal of the Brazilian Society of Mechanical Sciences and Engineering,* Vol. 40, Article Number 267, doi: 10.1007/s40430-018-1192-7.

[28] Bhuyan, R.K., Routara, B.C., Parida, A.K. (2015). Using entropy weight, OEC and fuzzy logic for optimizing the parameters during EDM of Al-24% SiC P MMC, *Advances in Production Engineering & Management,* Vol. 10, No. 4, 217-227, doi: 10.14743/apem2015.4.204.

[29] Singh, N., Kumar, P., Goyal, K. (2013). Effect of two different cryogenic treated wires in wire electrical discharge machining of AISI D3 die steel, *Journal of Mechanical Engineering,* Vol. 43, No. 2, 54-60, doi: 10.3329/jme.v43i2.17827.

[30] Montgomery, D.C. (2017). *Design and analysis of experiments, 9th edition,* John Wiley & Sons, Hoboken, New York, USA.

APEM journal

Advances in Production Engineering & Management
Volume 14 | Number 2 | June 2019 | pp 153–165
https://doi.org/10.14743/apem2019.2.318

ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

A new architecture model for smart manufacturing: A performance analysis and comparison with the RAMI 4.0 reference model

Resman, M.[a,*], Pipan, M.[a], Šimic, M.[a], Herakovič, N.[a]

[a]Department of Manufacturing Technologies and Systems, Faculty of Mechanical Engineering, University of Ljubljana, Ljubljana, Slovenia

ABSTRACT

In this paper we proposed a new architectural model of the smart factory to allow production experts to make easier and more exact planning of new, smart factories by using all the key technologies of Industry 4.0. The existing complex reference architectural model of Industry 4.0 (RAMI 4.0) offers a good overview of the smart-factory architecture, but it leads to some limitations and a lack of clarity for the users. To overcome these limitations, we have developed a simple model with the entire and very simple architecture of the smart factory, based on the concept of distributed systems with exact information and the data flows between them. The proposed architectural model enables more reliable and simple modelling of the smart factory than the existing RAMI 4.0 model. Our approach improves the existing methodology for planning the smart factory and makes all the necessary steps clearer. At the end of the paper a comparison of the proposed architectural model LASFA (LASIM Smart Factory) with the existing RAMI 4.0 model was made. The developed LASFA model was already successfully implemented in the laboratory environment for building the demo centre of a smart factory.

ARTICLE INFO

Keywords:
Industry 4.0;
Smart manufacturing;
Smart factory;
Architectural model;
Reference architectural model;
RAMI 4.0

**Corresponding author:*
matevz.resman@fs.uni-lj.si
(Resman, M.)

Article history:
Received 20 December 2018
Revised 24 April 2019
Accepted 7 May 2019

1. Introduction

The introduction of Industry 4.0 and with it factories of the future has become an important focus of the world's industry over the past few years. The key technologies of Industry 4.0 can have a major impact on an increase in efficiency and the availability of production assets, raising the efficiency of equipment and production, and increasing the value per employee. At the same time, the goal of introducing smart factories is to reduce costs, lead times, delivery times, etc. The introduction of the technologies of Industry 4.0 into the factories of the future is essential if we want these factories to become flexible and agile. We found that the biggest challenge is how to start planning the factory of the future or how and where to start introducing the new technologies of Industry 4.0 into these factories.

The fourth industrial revolution combines various technologies, such as the digitalisation of production processes and systems (digital twins of production processes and systems), cloud computing in combination with new mathematical algorithms, artificial intelligence, digital agents, the Internet of Things (IoT), big data to create cyber-physical systems (CPS) and smart factories. Industry 4.0 also includes various automated systems that allow the automatic exchange of data [1].

We can assess Industry 4.0's integration into a company using different maturity models that show the maturity level of a company in terms of the integrated technology of Industry 4.0. Some models are web-based, self-assessment tools, others are published in scientific journals [2]. The group of web-based, self-assessment tools covers the maturity models of PwC (PricewaterhouseCoopers) [3], Impuls [4], IHK (Industrie- und Handelskammern) [5] and VDMA (Verband Deutscher Maschinen- und Anlagenbau) [6]. Each of these tools uses different approaches to understand Industry 4.0. We found a variety of models in scientific publications. A popular and frequently mentioned architectural model is *The Reference Architectural Model Industry 4.0* (subsequently referred to as RAMI 4.0) [7]. Other models include *The Industrial Internet Reference Architecture* (IIRA) [8], developed by the Industrial Internet Consortium, and *The Stuttgart IT-Architecture for Manufacturing* (SITAM) [2, 9], developed within several research projects at the Graduate School of Advanced Manufacturing Engineering. Smaller initiatives like *Virtual Fort Knox* and FIWARE also provide data-driven concepts [10, 11].

Many other researchers have looked at architectural models of smart factories and the connections between the systems they contain. Important contributions have been made by Monostori *et al*. [12, 13], Kemeny *et al*. [14, 15], Valckenaers *et al*. [16] Bagheri *et al*. [17], Leitao *et al*. [18], and others [19]. Hussain *et al*. [20] presented a framework for sustainable manufacturing with its associated architecture. Vieira *et al*. [21] presented a literature review of the areas of simulation. In [22], Zheng *et al*. were researching the conceptual framework, the scenarios, and the future perspectives of smart manufacturing systems for Industry 4.0. Zhang *et al*. [23] presented concepts to achieve real-time manufacturing, capturing and integrating three different layers. In Liu *et al*. [24] the authors discussed and compared the concepts of Industry 4.0 and cloud manufacturing.

The first part of this paper presents a detailed description of RAMI 4.0, which is taken as the basis and reference to perform a comparison with the newly proposed architectural model concept. The second section explains the proposed concept of the architectural model (LASIM Smart Factory, referred to as LASFA) in detail, showing all the key elements of a smart factory taken from different vertical layers of RAMI 4.0 and placing them into a two-dimensional platform. The acronym LASIM stands for the original name of the laboratory that proposed the new architectural model. The main difference between the newly proposed model and RAMI 4.0 is the graphical presentation of the elements in a two-dimensional platform. Subsequently, the new model shows the exact locations of the elements as well as the interconnections and the directions of the material/information flows between the elements. The last part includes a comparison of the LASFA model with the RAMI 4.0 architectural model and explains why our model is more useful and easy to understand. The paper concludes with a description of our findings. All the abbreviations used in the paper are explained in the Appendix A.

2. Reference Architectural Model Industry 4.0 (RAMI 4.0)

2.1 Brief overview

The organisations BITKOM, VDMA and ZWEI decided to develop a new architecture model for the needs of Industry 4.0. For this, they took the Smart Grid Architecture Model as a basis [25, 26]. RAMI 4.0 is a three-dimensional model that describes Industry 4.0's space. On the horizontal axis, the layers include different views, such as assets, functional descriptions, data maps, etc. This corresponds with the IT approach of grouping complex projects into subsystems. The other key criteria are the lifecycle (type) and service life (instance) of the products and production systems with the value stream they contain. The vertical axis represents the third type of key aspect, i.e., the allocation of functions and responsibilities within the factories or plants. The combination of a lifecycle and a value stream with a hierarchically structured approach for the definition of Industry 4.0 components is a special feature of RAMI 4.0. The model allows for the logical grouping of functions and the mapping of interfaces and standards [27].

The RAMI 4.0 model is based on the established standards for automation, such as IEC 62890, IEC 62264, IEC 61512/ISA95, as shown in Fig. 1. It combines the key elements and technologies of

Industry 4.0, integrated into a 3D layered model. In this way, a complex system with internal connections can be divided into smaller and, for a better understanding, simpler subsystems [2, 9].

On the right horizontal axis there are the Hierarchy Layers, which are listed in the international standard IEC 62264. This axis represents the various functionalities in companies and factories. In order to present Industry 4.0, the axis has been divided into smaller subsets [2, 9]. The left horizontal axis represents a sustainable cycle of production and product, based on IEC 62890. This axis is further divided into types and instances. The type passes into the instance when the development and the prototype are completed, and the product is in production. The six layers into which the vertical axis is divided serve to describe the splitting of the device, layer by layer [2, 7, 9].

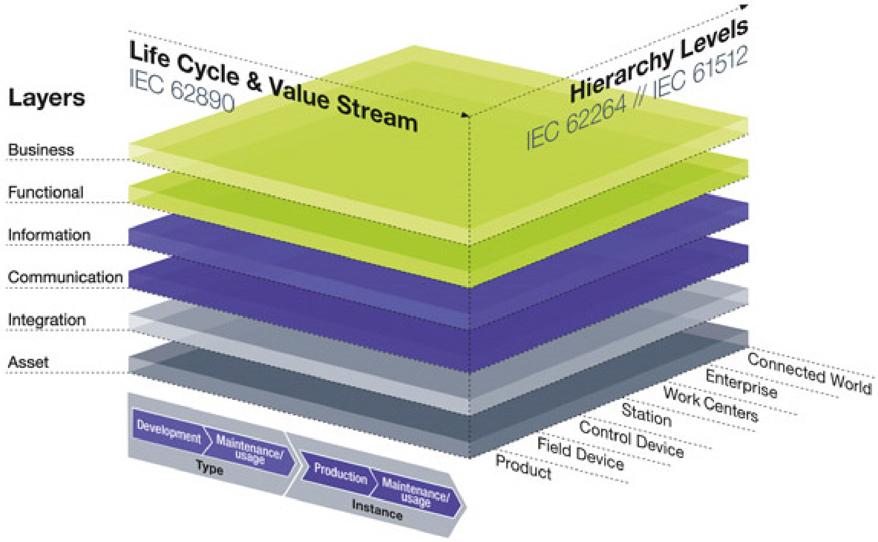


Fig. 1 Reference Architectural Model Industry 4.0 (RAMI 4.0) [2]

2.2 RAMI 4.0 in more detail

The RAMI 4.0 model has six layers on the vertical axis and two on the horizontal axis. Beginning with the vertical axis, the first layer is the 'Asset layer', which shows the physical objects, such as metal parts, documents, archives, diagrams, humans, etc. One layer higher is the 'Integration Layer', where transformations and connections of the physical objects into a digital world takes place. The components of the 'Asset layer' are connected with the digital world by the 'Integration Layer', which deals with the easy processing of information and can be considered as a link between the physical and digital worlds. This layer involves computer control of the process, system drivers, human-machine interface devices, humans, bridge wires, switches, hubs, sensors, RFID (Radio-Frequency Identification), etc. The next layer is the 'Communication layer', which provides standardized communications between the 'Integration layer' and the Information layer. The standardization is achieved with a uniform data format, which is used in the Information layer, which provides the control of the 'Integration layer'. The 'Information layer' holds data in an organized way. The basic purpose of this layer is to provide information about the total number of sales, purchase orders, suppliers, and locations. It holds information about all the products and materials that are manufactured in the industry. It also gives information about the machines and components that are used to build the products. It gives information to customers and saves their feedback. The 'Information layer' is software based, i.e., it might be in the form of applications, data, figures, or files. In this layer the transformation of the received events in data suitable for higher layers takes place. The next layer on the vertical axis is the 'Functional layer', which is responsible for production rules, actions, processing, and system control. It also facilitates users as per product features, like cloud services (restore/backup functionality). Moreover, it involves various other activities, like the coordination of components, system power on/off, testing elements, delivery channels, user inputs, and functions including, but not limited to, alert lights, snapshots, touch screen and fingerprint authentication. The 'Functional layer' includes remote access and horizontal integration. The last layer is the 'Business layer', which is composed of the business strategy, business environment, and business goals. Moreover, it deals with promotions and offers, target locations, advertisements, customer-relationship management, budgets and the pricing model [25, 27].

The horizontal axis on the left-hand side of Fig 1 shows the life cycle and value stream of the industrial production process (Fig. 2). It has two phases: Type and Instance. When the product is under development then it is in the Type phase. When the product is in production it is in the Instance phase. Whenever the same product is under development again it is in the Type phase again. When customers buy products, the products are in the Type phase again. When the products are installed in a system, they are in the Instance phase again. Changing the phase from type to instance can be repeated multiple times [7, 25].

The second horizontal axis represents the Hierarchy Layer, which is shown in Fig. 1 on the right-hand side. The Hierarchy Layer is based on the international standards for enterprise control system integration (IEC 62264 and IEC 61512). In addition to the four layers named 'Enterprise', 'Work Centers', 'Station', and 'Control Device', the last two layers at the bottom are added (but are not included in standards) and are called 'Field Device' and 'Product'. The layer 'Field Devices' makes it possible to control the machines or systems in an intelligent and smart way, e.g., smart sensors. The layer 'Product' takes into account the product homogeneity and the production capacity with their interdependencies. The layer named 'Connected World' is at the top. In this layer the factory can reach external partners through service networks. These layers show the fundamental views for Industry 4.0 organization [28-30]. The RAMI 4.0 model takes into account flexible systems and machines [7].

Based on a detailed analysis of the reference model RAMI 4.0 that includes all the elements of the vertical and horizontal axes, there is no exact definition with regards to how the individual elements inside each layer are interconnected with the elements. In our opinion, those interconnections are crucial and have to be defined when planning a new smart factory or upgrading an existing factory to create a smart factory. One of the other important aspects is the integration of digital twins and digital agents into distributed systems, and not as decentralised systems in each vertical layer. This part is still missing from RAMI 4.0.

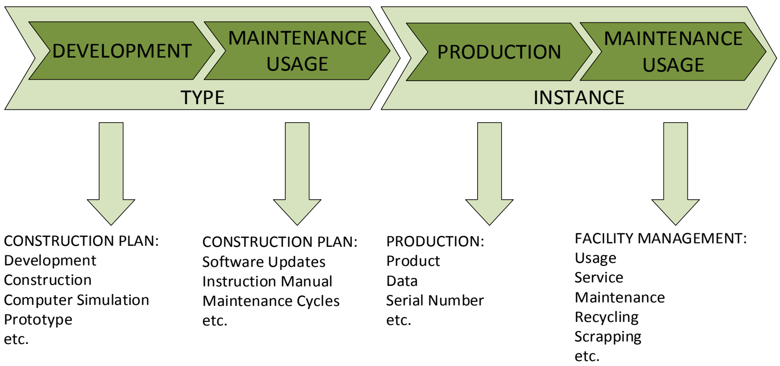


Fig. 2 Product life cycle: From the first idea to the scrapyard [7]

3. Proposed architectural model LASFA – LASIM smart factory

3.1 Global description

The LASFA architectural model (LAsim Smart FActory) is a concept for how to approach the planning and implementation of smart factories. The model was built based on RAMI 4.0, from where we took the hierarchy of the layers. We focused on one of the most important features of smart factories – the communications between systems in the smart factories' distributed systems. Using this model, users will be able to understand the principle operating smart factories.

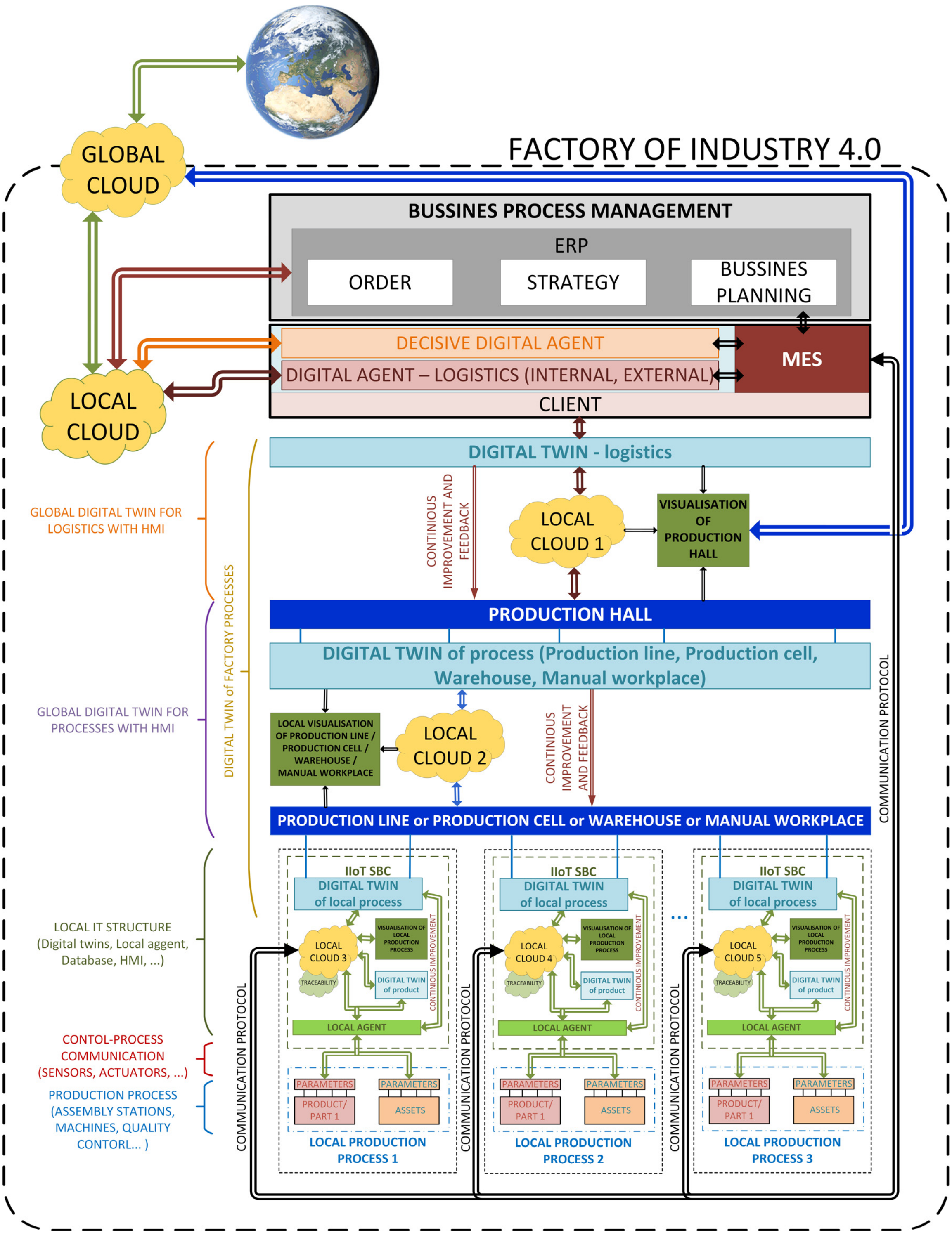


Fig. 3 LASFA model with all systems and processes

The architectural model is based on the results and insights of research projects and information from various European industrial manufacturers. The LASFA model is focused on the main aspects of Industry 4.0, such as horizontal integration, vertical integration, consistent engineering and systems that follow people's needs [7].

The development of the proposed architectural model is based on the reference architectural model RAMI 4.0. Unlike RAMI 4.0, the proposed model is two-dimensional, which makes the concept of a smart factory easier to understand. The RAMI 4.0 reference model is an abstraction for planning smart factories and is therefore not so easy to understand and is not so useful for industry users.

The LASFA architectural model, shown in Fig. 3, presents the individual systems that are combined into a smart factory. It includes several layers, as well as a business process, that result in a product in the production layer. The business layer includes the company's strategy and its leadership in the future as well as the monitoring and the delivery of orders.

Every manufacturer has one or more production lines (more production lines, production cells, warehouses, manual workplaces, etc). The production line consists of several local production systems that are, in this case, treated as distributed systems. Each local production system requires its input and output data. In the following section we will present all the elements in Fig. 3 in more detail.

3.2 Detailed description of the model

In this model we are focused on digital twins in different layers, which means a virtual copy of the real world. The model consists of digital twins for logistics, production lines, and local production processes.

Several production lines / production cells / warehouses / manual workplaces, together form a production hall. As we mentioned before, the proposed model of the smart factory includes its own digital twin for the production hall. Digital twins placed at different levels present a virtual copy of the systems in the real world. One of the important facts is that the digital twins without digital agents do not play an important role in a smart factory, as it only represents a virtual model of a real production system. With the help of various digital agents (each digital twin also has a digital agent – local or global) we get a digital twin that continually sends feedback and new production plans to the real world in real time. The LASFA model includes a decision-making digital agent, despite the fact that all the digital agents are connected. At the current stage of development, the decision-making agent is still human (i.e., a worker). In the future, we can expect that in the cases of production processes, where decisions about improvements to a single production process or production plan will have to be taken in real time, the expert will be replaced by a computer and advanced smart algorithms or artificial intelligence. But the absolute decision-maker should still be the expert. With such an approach we will be able to make the production process more flexible and agile, but the security and other "real-life" decisions will be taken by the expert to ensure the stable functioning of the production processes. Otherwise, if the absolute-decision maker was to be a smart algorithm, this could lead to the uncertainty, instability and insecurity of the production processes.

Fig. 3 shows the links for information exchange between individual systems in a smart factory and different local clouds. The model of a smart factory is built in such a way that each system is an independent unit with its own PlugAndProduce local control unit – distributed systems. This also allows us to add or remove an individual system without major changes to the global system. All the systems inside the model are interconnected. As it is with other models of smart factories, ours also does not have the classic pyramid shape [13]. For the smooth operation of a smart factory, it is also necessary to record and collect the data from sensors and save it in a local cloud.

The LASFA model includes the concept of remote access to the smart factory's data. The concept provides access to some data via the Internet. Users can access specific data within a database (Global cloud) over secure Internet connections. The Global cloud can be inside or outside the Industry 4.0 factory. Users or customers of the product can perform condition monitoring and check the progress of the product being manufactured. The user will be able to change the product's configuration during its manufacturing or production process (changing the colour, components, and accessories, if it is still possible). All the exchange of data is performed in real time, so customers of the product can monitor the production and assembly of the product.

The production line in Fig. 4 shows several different local production processes with different properties and requirements. Each local production process has its own single-board computer (SBC). SBCs are powerful enough to run standard operating systems and mainstream workloads [31]. Information and data exchange with a local cloud is provided through a wireless network or an optical cable for large amounts of data and information. The data and information exchange is bidirectional due to the feedback control performed by local digital agents. Each agent uses the input data and executes the optimization process to calculate the new parameters that are sent back to the real local system. The digital agent with its own database can make decisions independently.

Fig. 6 shows a detailed overview of the production process. All the process parameters are recorded and saved into the local clouds. Due to the bi-directional communications all these parameters are exchanged through the local clouds. This enables all the digital agents to have access to this data and to send new data, as well as to exchange data with other local and global clouds. Each local production process involves a local digital agent that receives information from a local cloud (database). A local agent is a mathematical model or an advanced intelligent algorithm that can also include artificial intelligence (AI). In today's factories, this segment is covered by a human being. The local digital agents use reference parameters and tabulation lists to change the parameter values. In the case of an unusual deviation, the local digital agent detects an error, reports it and can even react. The trend of changing parameters makes it possible to find out when it will be necessary to perform maintenance work (changing a cutting knife, matrix, changing oil and filters, etc.).

An important part of the LASFA model is the clear and concise visualisation. This visualisation enables internal and external users to access information about the condition of their manufactured product using a secure local or global Internet access point. Users, as well as workers on the assembly line, can monitor events with personal computers, smartphones or tablets, as proposed and shown in Fig 5.

Sensors installed in the real systems are used to capture the process parameters and visualize them (for example, the total production in the enterprise, which means the amount of raw material stock, the number of pieces manufactured, the defects of the production system, the identity of a line, the number of workers on a line, the number of shifts on a particular line, a product design plan, etc.) and also in digital form, another digital twin. The data captured from the production line and the data from the digital twin are collected in local clouds. The local clouds are synchronized with the global cloud. The local clouds also include the concept of digital agents that decide which data is important enough to capture and which is not required for subsequent operations.

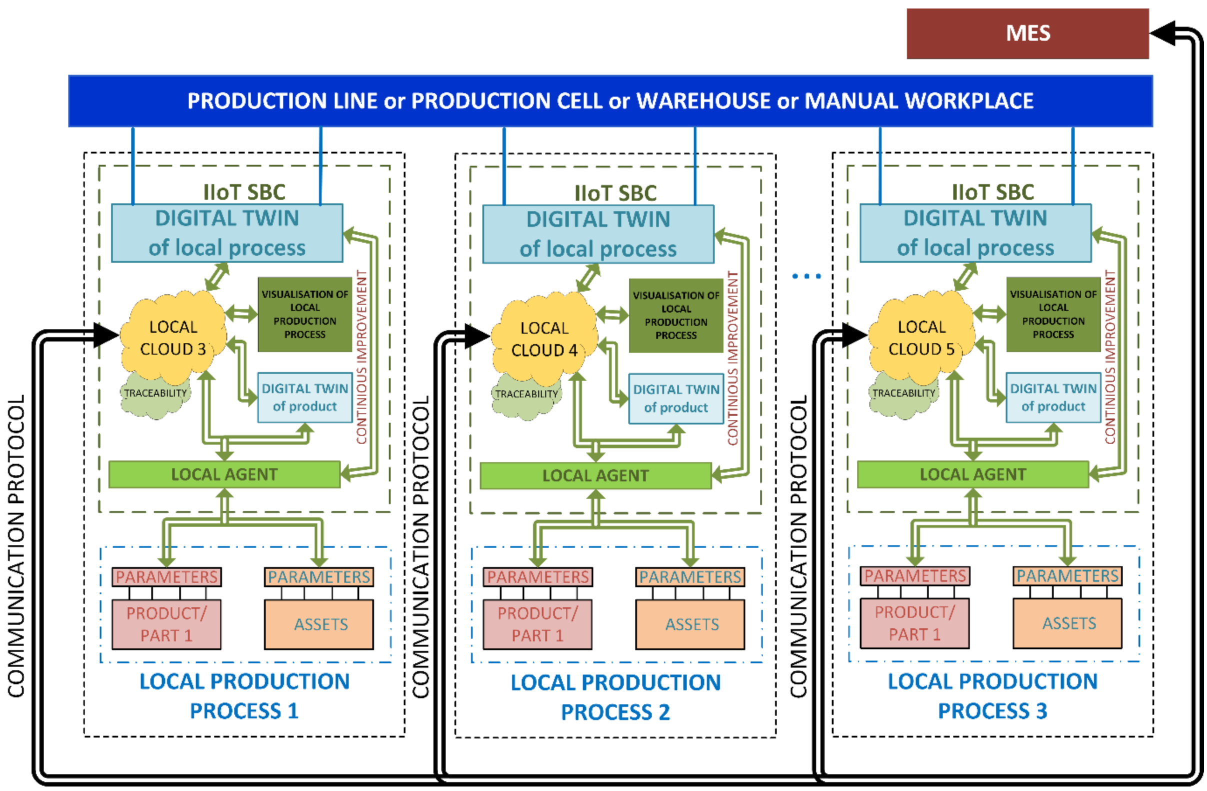


Fig. 4 Detailed view of a production line / production cell / warehouse / manual workplace

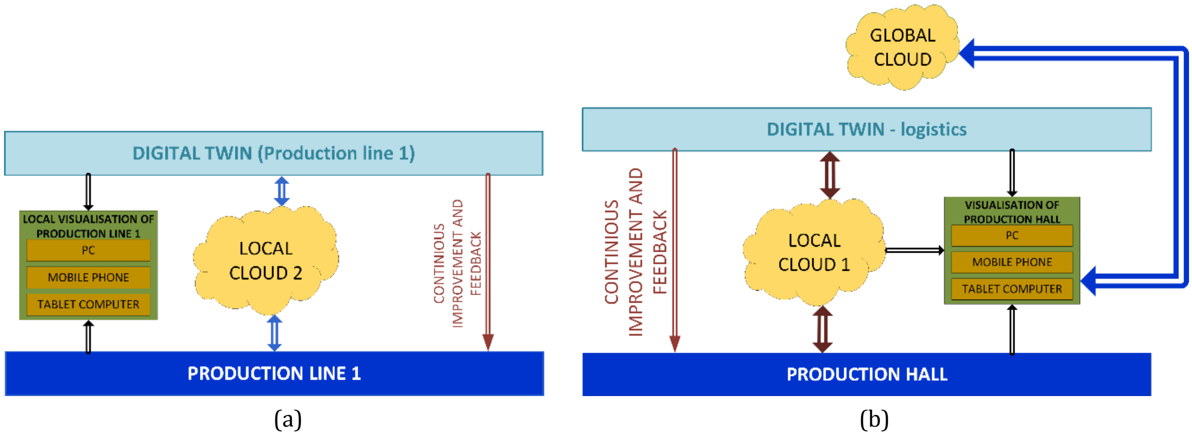


Fig. 5 Visualisation of the production processes (a), the production logistics (b) and the local production process

The main goal of digital twins is to provide continuous control of the production operations, systems and processes. In the event of an unexpected change or shutdown, the digital twin simulates different scenarios and provides the best solution at a given moment in real time. In this way, we have constant improvements to the real system through a feedback loop. Using communication protocols, the local clouds exchange data with the global cloud. Usually, factories include several production lines, manual assembly workplaces, warehouses, production cells and other production systems. Like with the entire production hall where the digital twin for logistics cooperates with local clouds, this layer also includes a digital twin of the production line that works together with its own local cloud. Each system or process has its own local cloud; therefore, we attain a completely distributed system. In the case when a system or a process stops operating, only one part of the production is disabled and not the entire production, as was the case in factories with a centralised database system. The data captured in the production line is stored in the local cloud.

The production line has several local production processes (e.g., a deep-drawing process, product-assembly process, a wire-bending process, a plastic-injection process, etc.). Fig. 6 shows the connections between the systems and the processes in the local production processes. Each local production process has its own digital twin, which constantly improves and optimizes the real system and generates a feedback loop. If we look deeper into the local production process itself, we can recognise many links and locations for data exchange. Each local production process performs a process (in our case a deep-drawing process). The local production process consists of several sub-processes, in our case it consists of measuring systems, a hydraulic process, a control process, and others. The sensors ensure that various data is captured on each subsystem. The data is collected in a local cloud. With this data, it is possible to set up a digital twin of the local production process. The concept is illustrated using various local digital agents and intelligent algorithms. The concept also includes predictive maintenance algorithms, which can be found in the literature [12]. When the parameter values change outside the acceptable range, the algorithm recognizes the error and the system receives the information that the part must be replaced (cutting knife, matrix, etc.). The goal of the digital twin is also to reverse influence. It can change the parameters in the sub-processes. This gives us the best solution at a given moment. The data in the local cloud enables us to use artificial intelligence and machine learning. At the moment, a human is still the main decision-maker, but in the future, the control will be taken over by an agent and a computer in the background. The result of the local production process is a product or a semi-finished product. In the smart factory, the product will also have its own digital twin (see Fig. 4). For this reason, it is necessary to capture all the information that describes the product as well as possible (dimensions, roughness, manufacturing tolerance, geometric tolerances, etc.). As is the case with all the other data, the concept also includes storing this information in a local cloud.

Field devices are also a very important part of smart factories. The data captured in the field devices (sensors) are collected in a local cloud, and this data is shared with other clouds. This

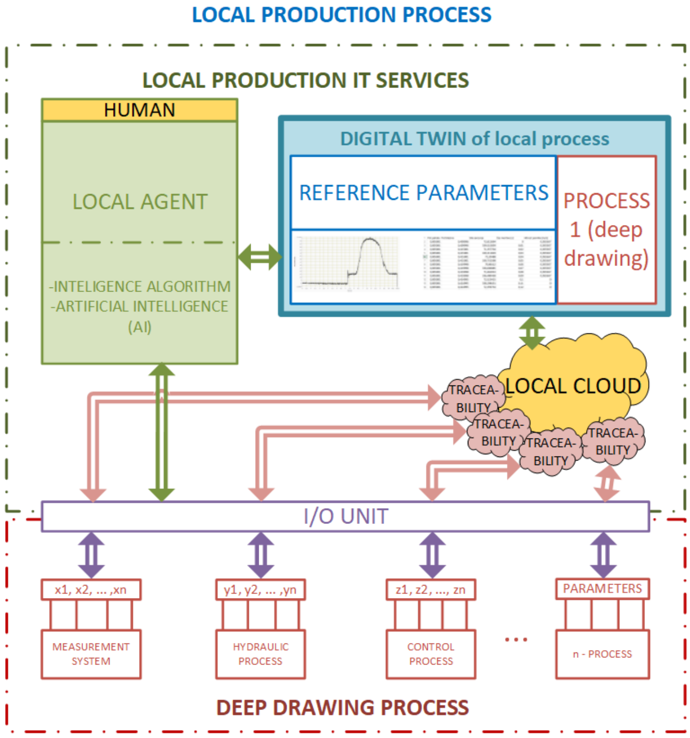


Fig. 6 Detailed view of the local production process

system represents the data captured using RFID, temperature sensors, humidity sensors, presence sensors, different control points, etc. Field devices used for manufacturing process control (valves, breakers, etc.) are controlled via a local agent.

4. Comparison of the LASFA model with RAMI 4.0

When discussing the LASFA and RAMI 4.0 models, and comparing them, we must be careful, because they are different types of models. Nevertheless, we can still make some comparisons between the RAMI and LASFA models in terms of the usefulness and clarity of the data and the information flow between the building blocks of the smart factory for the end user. RAMI 4.0 is based on standards for automation and is very generic. It offers a good overview of all the key technologies of Industry 4.0 and offers various layers and vertical axes as the backbone of the smart factory. As explained at the end of the Section 2, it leads to some limitations regarding an understanding of the exact positioning of different technologies and functions as well as the connectivity between them. This means that the planners of the smart factory do not know exactly where to place and how to interconnect some of the very important technologies, like different kinds of digital twins and digital agents, which is the main challenge for all the planners of smart factories.

On the other hand, the LASFA architectural model is much more specific and offers the end-user a simple visualization of the entire architecture of the smart factory, with the definition of the exact locations and functions of the digital twins and agents, with exact information and the data flow between them. The model shows a very clear distribution and the autonomy of every single building block of the smart factory, from the product to the management. The LASFA model is, therefore, in comparison to the RAMI 4.0 model, ready for direct implementation in the industrial environment and guides the smart-factory planner step by step from the smallest detail of the smart factory to the big picture. It is developed specifically for smart-factory planning and design.

The links between the building blocks of the smart factory in the LASFA model are shown very clearly; they also include the direction of the communication. Connections can also be added and graphically presented in the RAMI 4.0 model by using the Sparx RAMI 4.0 Toolbox [26] software. In this case the connections are made manually by experts, who need a lot of knowledge to plan all the links in all the vertical layers. The RAMI 4.0 model does not come close to the detail and clarity of the links that are important for the design/planning of smart factories.

A very important area covered by the LASFA model is that the functionality of the well-known ERP (Enterprise Resource Planning), MES (Manufacturing execution system) and PLM (Product Lifecycle Management) subsystems are replaced with new digital twins, which have integrated digital agents supported by artificial intelligence. In the model, the location and communication links for exchanging the information and data between the different digital twins in a smart factory are clearly shown. These features are not incorporated into RAMI 4.0 in such a clear manner. RAMI 4.0 does not show the location where ERP, MES and PLM are integrated into the system and modules in smart factories, nor in which layers they are needed, and it does not show which data and information are needed for operation, etc.

Table 1 Feature comparison between the proposed LASFA model and RAMI 4.0

	LASFA model	RAMI 4.0 model
Communication connections between systems and processes inside a smart factory	All the necessary communication links between individual systems are described in detail as well as the direction of communication (the information flow).	Administration shell includes communication protocol in a very general form. The standard does not show the exact connection between different systems inside each layer or vertically between layers.
ERP	Included in the model, the location and interconnections to other subsystems are clearly defined.	May be included, but its location is not defined.
MES	Digital twins with integrated digital agents are used to cover the function of the MES system.	May be included, but its location is not defined.
PLM	Integrated into the architecture model, the location, connection with other systems and main function are defined.	It is not included, but may be added. Hard to define the exact location, interconnections with other system inside the layer as well as between different layers.
Digital agents, artificial intelligence	LASFA represents different local and global digital agents, which are based on mathematical algorithms/models.	No locations for integration of digital agents in the model.
Visualisation of production process and systems	The LASFA model incorporates visualisation in different layers. Visualisation in smart factories is available in the production hall layer, production line layer and in the local production process layer.	Not clearly defined.
Digital twin	The digital twin is the main feature of our smart factory model. The digital twin is necessary for visualisation in different layers, systems operation, and decision making. The data for the digital twin's operation is collected in several local databases.	The model does not show the exact location of the digital twins and its function within the structure.
Visualisation of a decentralised and distributed system	Every local production process has its own local cloud, which is connected with other local clouds over a wireless network. A local cloud and a microcomputer form a decentralised system.	Mainly the visualization of decentralised systems. Distributed systems are briefly visualized.
Defines all the smart factory systems and components with their bi-directional links	Most of the key elements are included.	Not clearly defined
Capture and exchange of data between processes and local clouds	Detail description of the local and global clouds and the direction of the data exchange.	Not clearly defined – sensors, smart data, connections
Included standards for automation	Some standards are included and more can be implemented.	Some standards are included, others are in progress, in development

Another advantage of our model, in comparison to RAMI 4.0, is that it includes different digital agents, which are based on mathematical algorithms/models as well as artificial intelligence, where necessary. In our model we propose two types of digital twins and agents: local and global.

In our opinion, the visualisation of the systems and processes is very important in smart factories. By comparing the LASFA and RAMI 4.0 architectural models, we can see that the LASFA model includes visualisation, which is presented in the various layers of smart factories. It is necessary to visualise the production hall, the production line and the local production processes. The proposed model includes different visualisations, such as a visualisation of the decentralised and distributed system, the visualisation of all the important and necessary modules and systems of smart factories, and the locations for capturing and exchanging data and information between local processes and local/global clouds.

The digital twin is the main feature of our model. The digital twin is necessary for the visualisation of different layers, for the optimisation of systems and processes and for decision making. The data for the digital twin is collected from every local production process as well as the production line and is stored in different local clouds (databases). RAMI 4.0 does not clearly indicate how the systems and processes of smart factories cooperate with each other.

The proposed architectural model of the smart factory enables more reliable, simple and easy modelling of a smart factory than the existing RAMI 4.0 on every scale, from the small and simple to the big and complex production systems. It enables professionals in the production environment to see clearly all the details of the distributed concept of the smart factory; they see very clearly where in the process and in the system they have to position the global digital agent and all the local digital agents as well as where to position the different types of digital twins of the processes, systems and products. The model also gives the end-user very clear information about how to establish all the connections for data and information flow among the digital agents, digital twins and all the other building blocks of the smart factory. Therefore, it is a very suitable and helpful tool for professionals in the production environment. A comparison of the main features of both models is presented in Table 1.

5. Conclusion

In this paper we proposed a new architectural model called LASFA and compared it with the RAMI 4.0 model. The LASFA model combines different layers and shows a clear graphical presentation of all the key elements as well as their interconnections in a two-dimensional platform, which are important for the planning and design of smart factories. Our conceptual model is constructed in a very logical manner and can help industrial companies transform their manufacturing processes and systems into the factories of Industry 4.0.

We explained why the LASFA model is better for planning smart factories than the RAMI 4.0 model, which is an architectural reference for Industry 4.0. We found many advantages of our model, especially the visualisation of digital twins, the integration of different digital agents at different locations and levels, an accurate inventory of the local production process with the location of data capture, links between systems, the concept of integrating ERP, MES, and PLM into smart factories, etc.

On the other hand, RAMI 4.0 integrates different standards for automation, such as IEC 62890, IEC 62264, IEC 61512. The RAMI 4.0 architectural model shows a global view of Industry 4.0 from different vertical layers (Asset, Integration, Communication, Information, Functional, and Business). This gives us a complex and not so transparent three-dimensional view, which greatly enhances the complexity of understanding. In this case each layer on the vertical axis is treated separately and therefore we get a two-dimensional view of each layer.

The research discussed in this paper contributes an innovative architectural model and key design principles of the future factories at our laboratory. The LASFA model enables users to easily plan smart factories with all the necessary systems and to study the communication links between them. The architectural model shows very clearly how the communications between systems take place, where it is necessary to capture the data for the planning of digital twins in different layers of a smart factory, and it shows the visualization and which layer the users can access, etc. The digital twin represents the main feature of a smart factory in the newly proposed model.

In our future research work we plan to constantly improve the architectural model according to the newest scientific and industrial demands. Our main focus will be the area of big data and smart data collection, which are needed as the inputs for the development of different digital twins in different layers.

Acknowledgement

The work was carried out in the framework of the GOSTOP programme (OP20.00361), which is partially financed by the Republic of Slovenia – Ministry of Education, Science and Sport, and the European Union – European Regional Development Fund.

References

[1] Lu, H.-P., Weng, C.-I. (2018). Smart manufacturing technology, market maturity analysis and technology roadmap in the computer and electronic product manufacturing industry, *Technological Forecasting and Social Change*, Vol. 133, 85-94, doi: 10.1016/j.techfore.2018.03.005.

[2] Weber, C., Königsberger, J., Kassner, L., Mitschang, B. (2017). M2DDM – A maturity model for data-driven manufacturing, *Procedia CIRP,* Vol. 63, 173-178, doi: 10.1016/j.procir.2017.03.309.

[3] PwC. Industry 4.0 – Enabling digital operations, from *https://i40-self-assessment.pwc.de/i40/landing/*, accessed November 10, 2018.

[4] Impuls. Industrie 4.0-Readiness: Online-Selbst-Check für Unternehmen, from *https://www.industrie40-readiness.de/*, accessed October 12, 2018.

[5] IHK. Industrie 4.0 Reifegrad – Selbstcheck für Unternehmen, from *https://ihk-industrie40.de/selbstcheck/*, accessed October 15, 2018.

[6] Anderl, R., Picard, A., Wang, Y., Fleischer, J., Dosch, S., Klee, B., Bauer, J. (2015). *Leitfaden Industrie 4.0 – Orientierungshilfe zur Einführung in den Mittelstand*, VDMA, Frankfurt am Main, Germany.

[7] Schweichhart, K. Reference architecture model Industrie 4.0 (RAMI 4.0), from *https://ec.europa.eu/futurium/en/system/files/ged/a2-schweichhart-reference_architectural_model_industrie_4.0_rami_4.0.pdf*, accessed November 27, 2018.

[8] Industrial Internet Consortium. Industrial internet reference architecture, from *https://www.iiconsortium.org/IIRA-1-7-ajs.pdf*, accessed October 17, 2018.

[9] Gröger, C., Kassner, L., Hoos, E., Königsberger, J., Kiefer, C., Silcher, S., Mitschang, B. (2016). The data-driven factory – Leveraging big industrial data for agile, learning and human-centric manufacturing, In: *Proceedings of the 18th International Conference on Enterprise Information Systems,* Rome, Italy, 40-52, doi: 10.5220/0005831500400052.

[10] European Commission. Platforms for connected factories of the future, from *http://ec.europa.eu/information_society/newsroom/image/document/2015-48/workshop_report_platforms_oct15_finaldocx_12361.pdf*, accessed September 22, 2018.

[11] Weyrich, M., Ebert, C. (2016). Reference architectures for the internet of things, *IEEE Software,* Vol. 33, No. 1, 112-116, doi: 10.1109/MS.2016.20.

[12] Monostori, L., Váncza, J., Kumara, S.R.T. (2006). Agent-based systems for manufacturing, *CIRP Annals,* Vol. 55, No. 2, 697-720, doi: 10.1016/j.cirp.2006.10.004.

[13] Monostori, L., Kádár, B., Bauernhansl, T., Kondoh, S., Kumara, S., Reinhart, G., Sauer, O., Schuh, G., Sihn, W., Ueda, K. (2016). Cyber-physical systems in manufacturing, *CIRP Annals,* Vol. 65, No. 2, 621-641, doi: 10.1016/j.cirp.2016.06.005.

[14] Kemény, Z., Beregi, R.J., Erdős, G., Nacsa, J. (2016). The MTA SZTAKI smart factory: Platform for research and project-oriented skill development in higher education, *Procedia CIRP,* Vol. 54, 53-58, doi: 10.1016/j.procir.2016.05.060.

[15] Kemény, Z., Nacsa, J., Erdős, G., Glawar, R., Sihn, W., Monostori, L., Ilie-Zudor, E. (2016). Complementary research and education opportunities – A comparison of learning factory facilities and methodologies at TU Wien and MTA SZTAKI, *Procedia CIRP,* Vol. 54, 47-52, doi: 10.1016/j.procir.2016.05.064.

[16] Valckenaers, P., Van Brussel, H. (2005). Holonic manufacturing execution systems, *CIRP Annals,* Vol. 54, No. 1, 427-432, doi: 10.1016/S0007-8506(07)60137-1.

[17] Bagheri, B., Yang, S., Kao, H.-A., Lee, J. (2015). Cyber-physical systems architecture for self-aware machines in industry 4.0 environment, *IFAC-PapersOnLine,* Vol. 48, No. 3, 1622-1627, doi: 10.1016/j.ifacol.2015.06.318.

[18] Leitão, P., Colombo, A.W., Karnouskos, S. (2016). Industrial automation based on cyber-physical systems technologies: Prototype implementations and challenges, *Computers in Industry,* Vol. 81, 11-25, doi: 10.1016/j.compind.2015.08.004.

[19] Bongaerts, L., Monostori, L., McFarlane, D., Kádár, B. (2000). Hierarchy in distributed shop floor control, *Computers in Industry*, Vol. 43, No. 2, 123-137, doi: 10.1016/S0166-3615(00)00062-2.

[20] Hussain, S., Jahanzaib, M. (2018). Sustainable manufacturing – An overview and a conceptual framework for continuous transformation and competitiveness, *Advances in Production Engineering & Management*, Vol. 13, No. 3, 237-253, doi: 10.14743/apem2018.3.287.
[21] Vieira, A.A.C., Dias, L.M.S., Santos, M.Y., Pereira, G.A.B., Oliveira, J.A. (2018). Setting an industry 4.0 research and development agenda for simulation – A literature review, *International Journal of Simulation Modelling*, Vol. 17, No. 3, 377-390, doi: 10.2507/IJSIMM17(3)429.
[22] Zheng, P., Wang, H., Sang, Z., Zhong, R.Y., Liu, Y., Liu, C., Mubarok, K., Yu, S., Xu, X. (2018). Smart manufacturing systems for industry 4.0: Conceptual framework, scenarios, and future perspectives, *Frontiers of Mechanical Engineering*, Vol. 13, No. 2, 137-150, doi: 10.1007/s11465-018-0499-5.
[23] Zhang, Y., Zhang, G., Wang, J., Sun, S., Si, S., Yang, T. (2015). Real-time information capturing and integration framework of the internet of manufacturing things, *International Journal of Computer Integrated Manufacturing*, Vol. 28, No. 8, 811-822, doi: 10.1080/0951192X.2014.900874.
[24] Liu, Y., Xu, X. (2016). Industry 4.0 and cloud manufacturing: A comparative analysis, *Journal of Manufacturing Science and Engineering*, Vol. 139, No. 3, doi: 10.1115/1.4034667.
[25] Zezulka, F., Marcon, P., Vesely, I., Sajdl, O. (2016). Industry 4.0 – An introduction in the phenomenon, *IFAC-PapersOnLine*, Vol. 49, No. 25, 8-12, doi: 10.1016/j.ifacol.2016.12.002.
[26] Uslar, M., Hanna, S. (2018). Model-driven requirements engineering using RAMI 4.0 based visualization. In: Schaefer, I., Cleophas, L., Felderer, M. (eds.), *Workshops at Modellierung 2018*, Braunschweig, Germany, 21-30.
[27] Phoenix contact. RAMI 4.0 and IIRA reference architecture models – A question of perspective and focus, from *https://www.mynewsdesk.com/material/document/56241/download?resource_type=resource_document*, accessed October 15, 2018.
[28] VDI/VDE-Gesellschaft (2015). VDI/VDE-Gesellschaft Mess- und Automatisierungstechnik, Status report – Reference Architecture Model Industrie 4.0 (RAMI4.0), from *https://www.zvei.org/fileadmin/user_upload/Presse_und_Medien/Publikationen/2016/januar/GMA_Status_Report__Reference_Archtitecture_Model_Industrie_4.0__RAMI_4.0_/GMA-Status-Report-RAMI-40-July-2015.pdf*, accessed September 15, 2018.
[29] Ulrich, E., Bangemann, T., Bauer, C., Bedenbender, H., Diesner, M., Elmas, F., Friedrich, J., Goldschmidt, T., Göbe, F., Grüner, S., Hankel, M., Heidel, R., Hesselmann, K., Hüttemann, G., Kehl, H., Löwen, U., Pfrommer, J., Schleipen, M., Schlich, B., Usländer, T., Westerkamp, C., Winter, A., Wollschlaeger, M. (2016). *Industrie 4.0 – technical assets: basic terminology concepts, life cycles and administration models*, Status report, VDI/VDE-GMA, Düsseldorf, Germany, (Permalink: *http://publica.fraunhofer.de/documents/N-432116.html*).
[30] Pisching, M.A., Pessoa, M.A.O., Junqueira, F., dos Santos Filho, D.J., Miyagi, P.E. (2018). An architecture based on RAMI 4.0 to discover equipment to process operations required by products, *Computers & Industrial Engineering*, Vol. 125, 574-591, doi: 10.1016/j.cie.2017.12.029.
[31] Johnston, S.J., Basford, P.J., Perkins, C.S., Herry, H., Tso, F.P., Pezaros, D., Mullins, R.D., Yoneki, E., Cox, S.J., Singer, J. (2018). Commodity single board computer clusters and their applications, *Future Generation Computer Systems*, Vol. 89, 201-212, doi: 10.1016/j.future.2018.06.048.

Appendix A

The list of the abbreviations in the paper:

AI	Artificial intelligence
BITKOM	The name of German’s digital association
CPS	Cyber-physical system
ERP	Enterprise resource planning
IHK	Industrie- und Handelskammern
IIoT	Industrial internet of things
IIRA	The industrial internet reference architecture
IoT	Internet of things
I/O unit	Input/output unit
IT	Information technology
LASFA	LASIM smart factory
LASIM	Laboratory for handling, assembly and pneumatics
MES	Manufacturing execution system
PC	Personal computer
PLM	Product lifecycle management
RAMI 4.0	Reference architectural model Industry 4.0
RFID	Radio-frequency identification
SBC	Single board computer
SITAM	The Stuttgart IT-architecture for manufacturing
VDMA	Verband Deutscher Maschinen- und Anlagenbau

APEM journal

Advances in Production Engineering & Management
Volume 14 | Number 2 | June 2019 | pp 166–176
https://doi.org/10.14743/apem2019.2.319

ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

Simulation framework for determining the order and size of the product batches in the flow shop: A case study

Ištoković, D.[a,*]**, Perinić, M.**[a]**, Dobovičekc, S.**[a]**, Bazina, T.**[a]

[a]University of Rijeka, Faculty of Engineering, Rijeka, Croatia

A B S T R A C T

The problems of determining the order and size of the product batches in the flow shop with multiple processors (FSMP) and sequence-dependent setup times are among the most difficult manufacturing planning tasks. In today's environment, where necessity for survival in the market is to deliver the goods in time, it is crucial to optimize production plans. Inspired by real sector manufacturing system, this paper demonstrates the discrete event simulation (DES) supported by the genetic algorithm (GA) optimization tool. The main aim is to develop the simulation framework as a support for the daily planning of manufacturing with emphasis on determining the size and entry order of the product batches within specific requirements. Procedures are developed within the genetic algorithm, which are implemented in Tecnomatix Plant Simulation software package. A genetic algorithm was used to optimize mean flow time (MFT) and total setup time (TST) performance measures. Primary constraint for on-time delivery was imposed on the model. The research results show that solutions are industrially applicable and provide accurate information on the batch size of the defined products, as well as a detailed schedule and timing of entry into the observed system. Display of the solution, in a simple and concise manner, serves as a tool for manufacturing operations planning.

A R T I C L E I N F O

Keywords:
Process planning;
Flow shop;
Sequence-dependent setup time;
Batch scheduling;
Discrete event simulation;
Genetic algorithm (GA)

**Corresponding author:*
distokovic@riteh.hr
(Ištoković, D.)

Article history:
Received 25 February 2019
Revised 18 April 2019
Accepted 7 May 2019

1. Introduction

One of the key problems in operational planning of manufacturing is to determine the order and size of a product batches. In today's environment, the solution to this problem is a necessity for survival in the market. Manufacturing companies must deliver the goods in time to avoid losses and ensure competitiveness where activities are planned to effectively use the available resources [1].

This paper explores the problem of determining the order and size of a product batches with the aim of developing the simulation framework as support for the daily planning of manufacturing. The study deals with the problem of the real sector.

FSMP planning includes, among other things, the order of jobs in unidirectional flow system where at any workstation more than one machine of the same type can be located. The machine can process at most one job at any given time. All jobs are subject to priorities that limit them to the same processing order throughout all processing stages [2]. Each customer order must be processed in all or some of the workstations. Special attention is paid to the presence of the sequence-dependent setup times, the size of the product batches and the availability of production resources [3]. Setup time is the time required for the staff to prepare and provide a job for uninterrupted work when changing the operation. It does not create any additional value, and by increasing the number of setups, available machine processing time is consequently reduced. Reducing the available processing time increases the possibility of delays with delivery, which every company wants to avoid.

This paper presents a discrete model of the manufacturing system, made using the simulation package Tecnomatix Plant Simulation. Using the genetic algorithm implemented within the simulation package itself, the problem of determining the order and size of the product batches is solved. This simulation shows the exact values of the required parameters that are applicable to any system of that type.

The rest of the article is organized as follows. A review of relevant literature is presented in the section 2. Section 3 introduces the formulation of the observed problem, as well as an overview of the development of the simulation model. Section 4 shows the details of the simulation experiments. Section 5 provides an analysis of experimental results. Finally, section 6 gives final notes.

2. Literature review

Discrete event simulations (DES) are powerful and effective tools for solving many real-world problems. They are also one of the most commonly used techniques for analysing and understanding the dynamics of manufacturing systems. The proof of this is a large number of published scientific papers on this particular subject. Negahban and Smith provided a comprehensive review of discrete event simulation publications with a focus on applications in manufacturing, including manufacturing operations planning and scheduling problems [4]. Another comprehensive survey on scheduling problems, which provides an extensive review of about 500 papers that have appeared since the mid of 2006 to the end of 2014, was presented by Allahverdi [5]. He introduced very significant classification and notation of scheduling problems, based on shop ambient, process features, setup conditions and performance measures. I. Ribas *et al.* provided overarching review of recently published articles about the problems of scheduling hybrid flow shop (HFS). The works are divided into two categories based on the characteristics of the HFS and production constraints and in view of the proposed approach to problem solving [6]. The most important surveys of using DES for solving flow-shop specific manufacturing operations planning and scheduling problems are reviewed in this section.

Gourgand *et al.* [7] researched scheduling problems in two and *m* machine stochastic flow shop with infinite buffers. They implemented hybrid approach, consisting of recursive heuristics or metaheuristics and performance evaluation algorithm. Makespan, used for measuring performance of generated feasible job schedules, was computed using Markov chain, or estimated using DES. Wang *et al.* [8] used genetic algorithms (GA) for stochastic flow-shop scheduling problem. Their objective was to avoid premature convergence of the GA. Yang *et al.* [9] solved a multi-attribute combinatorial dispatching (MACD) decision problem in a flow shop with multiple processors (FSMP) environment. The same problem solved Azadeh *et al.* [10]. I. A. Chaudhry and M. Usman used a genetic algorithm and independent spreadsheets to simultaneously solve scheduling problems and process planning in a job shop environment [11]. R. Meolic and Z. Brezočnik proposed a new approach to solving job shop scheduling problems with an emphasis on identifying feasible solutions. The new approach allows all schedules of relatively large systems to be found using the data structure, the zero-suppressed binary decision diagrams [12].

Hendizadeh *et al.* [13] considered a flow shop scheduling problem of a manufacturing cell that contains families of jobs. Setup times are sequence-dependent of the families. To minimize makespan and total flow time, the authors proposed multi-objective genetic algorithm (MOGA). The same problem has been considered by Lin and Ying [14] using a two-level multi-start simulated annealing (TLMSA). Lee [15] dealt with a problem of estimating order lead time in hybrid flow shops, where orders arrive dynamically. Alfieri [3] proposed the solution for multiple objective flow shop scheduling problem on model with work calendars on resources, multi-machine stages, re-entrant flows, external operations, sequence-dependent setup times, and transfer product batches between stages. Dugardin *et al.* [16] presented L-NSGA multi-objective

GA which uses the Lorenz dominance relationship for a re-entrant hybrid flow shop scheduling problem. They had two objectives: maximum utilization rate of the bottleneck and minimum completion time. Ladhari *et al.* [17] researched two-machine permutation flow shop problem with the sequence-independent setup times. Their objectives were minimizing the sum of completion times. Ying *et al.* [18] examined the no-wait flow shop manufacturing cell scheduling problem (FMCSP) with the sequence-dependent family setup times with makespan criterion as objective. Galzina *et al.* [19] deal with the flow shop scheduling problem using hybrid fuzzy logic and intelligent swarm method. The compiled model was compared with stochastic algorithms for assessing applicability to general problems. Chen and Hao [20] solved the problem of distributing the flow shop by applying a non-dominated sorting genetic algorithm (NSGA). They use it for multi-objective optimization of non-compact flow shops in view of process linking. Yan *et al.* used a Tabu search algorithm and particle swarm optimization in a two-stage semi-continuous flow shop to optimize the production and distribution decisions at the same time [21].

Liu *et al.* [22] compiled an overview of different optimization approaches for solving manufacturing planning and scheduling problems. They listed numerous global and local optimization methods, along with application examples and associated constraints. Frequently used techniques, like response surface methodology, gradient-based methods and evolutionary algorithms, as well as emerging ones, like stochastic approximation, particle swarm optimization and ant colony optimization were encompassed. Supsomboon and Vajasuvimon proposed simulation model using Tecnomatix Plant Simulation for making machine parts in the job shop. The simulation model shows that job expansion, plant allocation, group technology, and capacity expansion ultimately contribute to lower operating costs and increase employee utilization [23].

3. Materials and methods

The objectives of each manufacturing are to achieve the required product quality with the least cost of manufacturing, and delivery on time [24]. Delivering on time is the primary condition which must be satisfied. This generally means that the total required quantity of products q_j must be produced within the observed period, i.e. the completion time of the last workpiece (makespan) C_{max} must be less than the maximum available time of the manufacturing equipment $C_{max,goal}$ for the observed period, Eq. 1.

$$C_{max} < C_{max,goal} \tag{1}$$

Batch size of the *j*-product Lot_j is defined as a natural number in a given interval, Eq. 2. *Dg* represents the minimum batch size, and *gg* the maximum batch size of the *j*-product that cannot be greater than the total quantity of product q_j.

$$Lot_j \in [dg, gg], \quad \forall\, Lot_j \in \mathbb{N} \tag{2}$$

Taking any number from the defined range as the batch size enables a lot more potential optimization solutions than [25] the use of different batch size with the assumption that the total quantity of the product must be multiplied by the batch size (the total quantity of the product must be divided by the batch size). This means that each batch size is the same, which greatly reduces the ability to find a better solution.

If it is assumed that each batch size of the same product is equal to the total quantity of the product, it is likely that more products will be produced than needed. This ultimately does not change the mean flow time, but it extends the total processing duration. Also, this creates a stock that creates additional cost that is undesirable for the company.

Simple example: It is necessary to produce 1,234 workpieces in 2 weeks. A batch size is 400 workpieces. It is evident that if three batches of 400 workpieces are produced, there are still 34 workpieces left to produce. If four batches of 400 workpieces are produced, the stock will be 366 workpieces, which will increase the makespan, and therefore the possibility of not delivering on time. For this reason, it is ensured that the last batch of the same product is equal to the Eq. 3.

$$Lot_{j,last} = q_j - Lot\_Num_j \cdot Lot_j \tag{3}$$

where $Lot_{j,\text{last}}$ is the size of the last batch of the same product, $Lot_{j,\text{last}} \in [1, Lot_j]$; q_j is the total amount of the same product; Lot_Num_j is the amount of produced batches of the same product; Lot_j is the batch size of the same product.

In a flow shop production system the products travel in batches through the system. The batch size directly affects the flow time in a way that increasing the batch size linearly increases the flow time, worth and vice versa. The flow time F_j is defined as the time the j-product batch performs in the system, i.e. the difference between the j-product's output time and the j-product's input time. The flow time of each j-product batch is different because of uneven waiting times on the processing. For simpler further optimization their mean value is calculated by Eq. 4. Thus, the mean flow time MFT_j is actually the average time of all j-product flow times. By introducing the total setup time, the mean flow time does not change, but the makespan does. The bigger the total setup time, the bigger the makespan.

$$MFT_j = \frac{\sum_1^{Lot\_Num_j} F_j}{Lot\_Num_j} \tag{4}$$

In order to determine the entry sequence of product batches, a second variable is introduced -the probability of entering the j-product batches into the system $Perc_j$, defined as a natural number at a given interval, Eq. 5.

$$Perc_j \in [dg, gg], \quad \forall\, Perc_j \in \mathbb{N} \tag{5}$$

The real probability of entering the j-product batches $RealPerc_j$ is defined as ratio of the probability of entering the j-product batches and sum of the probability of entering product batches, according to Eq. 6.

$$RealPerc_j = \frac{Perc_j}{Perc_1 + Perc_2 + \cdots + Perc_j + \cdots + Perc_n} \tag{6}$$

The entry sequence of a product batches has no effect on the mean flow time, but when entering different product batch, the need for setup time appears. When one batch of a particular product is completed on the same production equipment, then a new batch of a particular product comes in. If the new product batch is the same as the previous product batch, then setup time is not required, i.e. the setup time is equal to zero. If the new product batch is different from the previous product batch, then the setup of the workplace is required before the start of processing. The setup time is randomly selected in a uniform distribution, according to [26]. At the arrival of the first batch of any product, the setup of the workplace is also carried out.

The question is why the setup time cannot be clearly displayed (if not automated, i.e. if it performs at least partially by a man)? There are many reasons, including working staff that performs the setup job is different (work in multiple shifts), fatigue and motivation of the staff members, etc.

It is not possible to analytically determine how much is the total setup time of the j-product for the observed period. It is determined by simulation. The total setup time of the j-product TST_j represents the sum of all setup times of the j-product ST_j, according to Eq. 7. The ST_Num_j represents the number of impressions of setup time ST during the production of the j-product.

$$TST_j = \sum_1^{ST\_Num_j} ST_j \tag{7}$$

It is concluded that the minimum total setup time will ideally be when the production is in a unit as large batches. Also, from the standpoint of the minimum mean flow time, it is preferred that the production takes place in the lowest unit of product batches, which leads to contradictions. For this reason, in order to simultaneously minimize both, the mean flow time and total setup time, it is necessary to conduct optimization.

Optimization will be performed using a genetic algorithm that is embedded in the used software package Tecnomatix Plant Simulation. The structure of genetic algorithm is shown in Fig. 1. Whereby *pop* means population, *gen* means generation and *gen_num* means maximum number of generations.

Based on the defined input variables (batch size Lot_j, probability of entering the batch $Perc_j$) and the simulation model of the production system, the fitness method will be minimized. The fitness method is defined as the sum of individual members where each member has a certain importance.

```
gen = 0
Generating an Initial Population pop(0)
if gen < gen_num
    gen = gen + 1
    Fitness proportionate selection pop'(gen) from pop(gen - 1)
    Crossover in two points pop'(gen) and saving in pop''(gen)
    Mutation pop''(gen)
    Probabilistic selection pop(gen) from pop''(gen) and pop(gen - 1)
minimize fitness
```

Fig. 1 Structure of genetic algorithm

Thus, each individual member is multiplied by the weight factor, with the higher weight factor being more important for the overall result. According to the above, each member for the mean flow time of *j*-product MFT_j would be multiplied with the weight factor a_j, and each member for the total setup time of the *j*-product TST_j would be multiplied with the weight factor b_j, shown in Eq. 8. Considering that the total sum of all weight factors must be equal to 1, Eq. 9.

$$fitness = min\ (a_j \cdot MFT_j + b_j \cdot TST_j) \tag{8}$$

$$a_1 + a_2 + \cdots + a_j + \cdots + a_n + b_1 + b_2 + \cdots + b_j + \cdots + b_n = 1 \tag{9}$$

4. Presentation of the problem

4.1 General

Studied production system is designed according to the production plant companies from the real sector which are producing families of technologically similar products. Technologically similar products are those that have a high degree of similarity to the order of processing and duration of the operations. It is assumed that the production of three products (*D*, *E*, *F*) is foreseen for delivery every two weeks, or more precisely every second Friday after the second shift at 10:00 p.m. The two-week quantity determined for the three products as well as the order and duration of the operations are assumed and shown in Table 1. The operation times are set in hours.

Table 1 Example data

j	D	E	F	M_{iC}
q_j	4616	3232	2616	
i				
1	10-0.028	10-0.035	10-0.036	3
2	20-0.031	20-0.035	20-0.036	3
3	30-0.012	30-0.01	30-0.01	1
4	40-0.02	40-0.019	40-0.017	2
5	50-0.08	50-0.012	50-0.14	1

The working week lasts for five days and takes place in two shifts. Operating hours per shift are eight. From this, according to the Eq. 6, the maximum availability of production equipment can be calculated C_{max} is 160 hours. The machines cannot operate continuously, without interruption, so the utilization time is 0.85, according to [27]. The required number of *i*-th production equipment M_{iC} has been obtained by [28], assuming that the reliability of production equipment equals 1. Production takes place at five workstations, where all three products pass unidirectional through the system and are processed at each workstation. Each workstation consists of the M_{iC} number of the same production equipment. All of the above defines the observed production system as a flow shop with multiple processors (FSMP). Using the Tecnomatix Plant Simulation software package, a discrete FSMP model was developed.

The setup time is defined as relatively large due to the fact that in a serial production, except change of tools, jig, etc., it is extremely important to check the first workpiece of the batch. This is sometimes a request by clients in some industries (e.g., automotive industry). Therefore, the setup time ST_j for workstations 1 and 2 is defined by a uniform distribution between 10 and 18 minutes, while for workstations 3, 4 and 5 is defined by a uniform distribution between 12 and 20 minutes. Bearing in mind that these are technologically similar products, the setup times for any combination of the previous and next batch of j-products are approximately equal ($D \rightarrow E$, $E \rightarrow F$, etc.).

The boundary conditions for batch size Lot_j and probability of entering the j-product $Perc_j$ were given by Eq. 11 and Eq. 12.

$$Lot_j \in [10{,}1000] \quad \forall\, Lot_j \in \mathbb{N} \tag{11}$$

$$Perc_j \in [100{,}1000] \quad \forall\, Perc_j \in \mathbb{N} \tag{12}$$

4.2 Genetic algorithms parameters

The values of genetic operators used were as follows: *probability of crossover* and *probability of mutation* were 0.8 and 0.15, respectively. When triggering optimization, the following limitations for the genetic algorithm were determined: number of population *pop_num* = 50, number of generations *gen_num* = 250, and number of observation *obs_num* = 10.

4.3 Coding of organisms

The genetic algorithm at the beginning of the optimization randomly generates an initial population of 50 individuals. Each individual (chromosome) consists of six genes that represent a specific property: first gene is batch size of the product D, second gene is batch size of the product E, third gene is batch size of the product F, fourth gene is probability of entering into system batch of the product D, fifth gene is probability of entering into system batch of the product E, sixth gene is probability of entering into system batch of the product F.

4.4 Definition of fitness function

The observed optimization task is the assignment task. Therefore, to solve this problem, a given gene assigns a random value:

- according to Eq. 11, for gene of batch size,
- according to Eq. 12, for the probability of entering a certain batch of the product.

When initial individuals with corresponding values (genes) are defined, the fitness method for all individuals within the population is calculated using Eq. 13:

$$\begin{aligned} fitness = C_{max,fit} + a_D \cdot MFT_D + a_E \cdot MFT_E + a_F \cdot MFT_F + b_D \cdot TST_D \\ + b_E \cdot TST_E + b_F \cdot TST_F \end{aligned} \tag{13}$$

Earlier defined objectives are that all the mean flow times are as small as possible, as well as all total setup times. Thus, the task of the genetic algorithm is to find the least value (minimum) of the fitness method. Furthermore, as all the optimization parameters are equally important, the assumption is that all weight factors are equal, i.e.:

$$a_D = a_E = a_F = b_D = b_E = b_F = \frac{1}{6}$$

Delivering on time is the primary condition which must be satisfied, as such, it should be a part of the objective function. However, it is not necessary for the makespan to be as small as possible, only to be satisfied. Therefore, the makespan will not be part of the objective function. In order for the genetic algorithm to "move away" from poor (unsatisfactory) solutions and "approach" better solutions, a penalty $C_{max,penalty}$ should be introduced according to Fig. 2. Value of $C_{max,fit}$ is added to the objective function. If $C_{max,fit} = 0$, the delivery condition is satisfied and will not have any effect on the objective function, but if $C_{max,fit} > 0$ then this means that the condition is not satisfied and that the value of the goal function will increase, which will ultimately result in moving genetic algorithm from bad solutions. In this way, it is achieved that products are made on time and that delivery is not delayed.

```
var Cmax,penalty : real := [ (Cmax - Cmax,goal) / Cmax,goal ]
var Cmax,fit : time
if   Cmax,penalty  >  0
     Cmax,fit := ( 1 + Cmax,penalty ) · ( Cmax - Cmax,goal )
else
     Cmax,fit := 0
end
```

Fig. 2 Penalty condition

The objective function was calculated in the example of two selected individuals.

	Lot_D	Lot_E	Lot_F	$Perc_D$	$Perc_E$	$Perc_F$
Individual 1	255	703	444	236	199	704
Individual 2	275	325	335	770	400	350

By simulation, the mean flow time and the total setup time are obtained. The time format used in the following text is days:hours:minutes:seconds (d:h:m:s).

	MFT_D	MFT_E	MFT_F	TST_D	TST_E	TST_F	C_{max}
Individual 1	2:11:38:44	2:01:57:41	1:15:18:29	2:04:22:00	1:19:08:15	3:23:06:42	23:02:46:56
Individual 2	0:23:49:18	1:00:46:18	1:01:28:57	1:19:53:49	1:02:33:06	0:22:43:36	5:22:33:50

Using penalty condition (Fig. 2) the penalties for each individual are obtained:

Individual 1
$C_{max,penalty}$ = [(23:02:46:56 - 6:16:00:00) / 6:16:00:00] = 2.467 > 0
$C_{max,fit}$ = (1 + 2.467) · (23:02:46:56 - 6:16:00:00) = 40:14:04:53

Individual 2
$C_{max,penalty}$ = [(5:22:33:50 - 6:16:00:00) / 6:16:00:00] = -0.122 ≤ 0
$C_{max,fit}$ = 0

Furthermore, in Eq. 13 the objective function is calculated for the two individuals mentioned.

Individual 1

$$fitness = 40{:}14{:}04{:}53 + \frac{1}{6}\cdot 2{:}11{:}38{:}44 + \frac{1}{6}\cdot 2{:}01{:}57{:}41 + \frac{1}{6}\cdot 1{:}15{:}18{:}29 + \frac{1}{6}\cdot 2{:}04{:}22{:}00 + \frac{1}{6}\cdot 1{:}19{:}08{:}15 + \frac{1}{6}\cdot 3{:}23{:}06{:}42 = \mathbf{42{:}22{:}40{:}11}$$

Individual 2

$$fitness = 0 + \frac{1}{6}\cdot 23{:}49{:}18 + \frac{1}{6}\cdot 1{:}00{:}46{:}18 + \frac{1}{6}\cdot 1{:}01{:}28{:}57 + \frac{1}{6}\cdot 1{:}19{:}53{:}49 + \frac{1}{6}\cdot 1{:}02{:}33{:}06 + \frac{1}{6}\cdot 22{:}43{:}36 = \mathbf{1{:}03{:}52{:}31}$$

The mentioned individuals (parents) can be selected for cloning by roulette wheel selection, whereby Individual 2 being much more likely to be selected than Individual 1. Every individual can be selected more than once, but it is also possible not to be selected even once. Genetic operators (2-point crossover, mutation) are applied on cloned individuals (offsprings) [29]. Then the probabilistic method selects individuals for the next generation between parents and offsprings. The process is repeated until it reaches the 250 generation. The genetic algorithm then shows the best solutions.

5. Results and discussion

The previously defined optimization by genetic algorithm was performed on an 8-core processor of 2.66 GHz, with duration of 2:04:40:50 (d:h:m:s). The best generated solution for a defined optimization task is 1:00:55:36 (d:h:m:s), which is the minimum value of the objective function. The optimization parameters obtained are given in Table 2.

From an evolutionary diagram, Figure 3, are visible solutions obtained during a defined number of generations. Also, it can be observed that the genetic algorithm has relatively quickly found a fairly good solution, but an increased number of generations were given an even better solution.

Through optimization, the values for batch size of *j*-product are obtained. By initiating the simulation of the production system model for previously obtained optimization results, the sequence of entering of the *j*-products is determined. Fig. 4, besides the sequence of the *j*-product entries, also presents other values such as: batch size of *j*-product, finished quantity of *j*-product, completion time (makespan), mean flow time of *j*-product batch, total number of setup times, total setup time of *j*-product and others.

Since the resulting makespan is smaller than the delivery deadline, the start of production can be shifted from Monday 6:00 a.m. to Tuesday 10:49:30 a.m. Also, by means of the Gantt chart, the correct timing of the *j*-product batch is visible. Due to a large number of batches and workdays, the Gantt chart is large and unobtrusive on a small display. For this reason, the Gantt chart for one working day is shown in Fig. 5. In addition to the start date of the product batches on particular production equipment, the end date of the processing of the last workpiece from the batch is shown, as well as the total processing time (the time the product batch spent on particular production equipment).

Table 2 Batch size of products D, E, F for the observed case

Lot_D	Lot_E	Lot_F	$Perc_D$	$Perc_E$	$Perc_F$
274	324	330	776	425	351

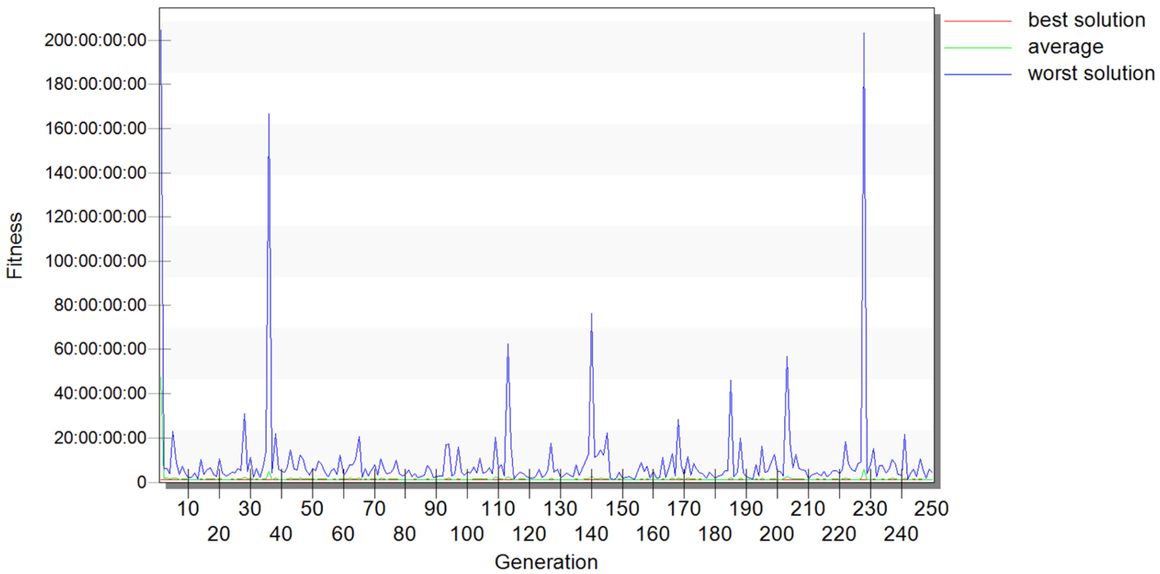


Fig. 3 Evolution diagram shows a quick finding of a enough good solution

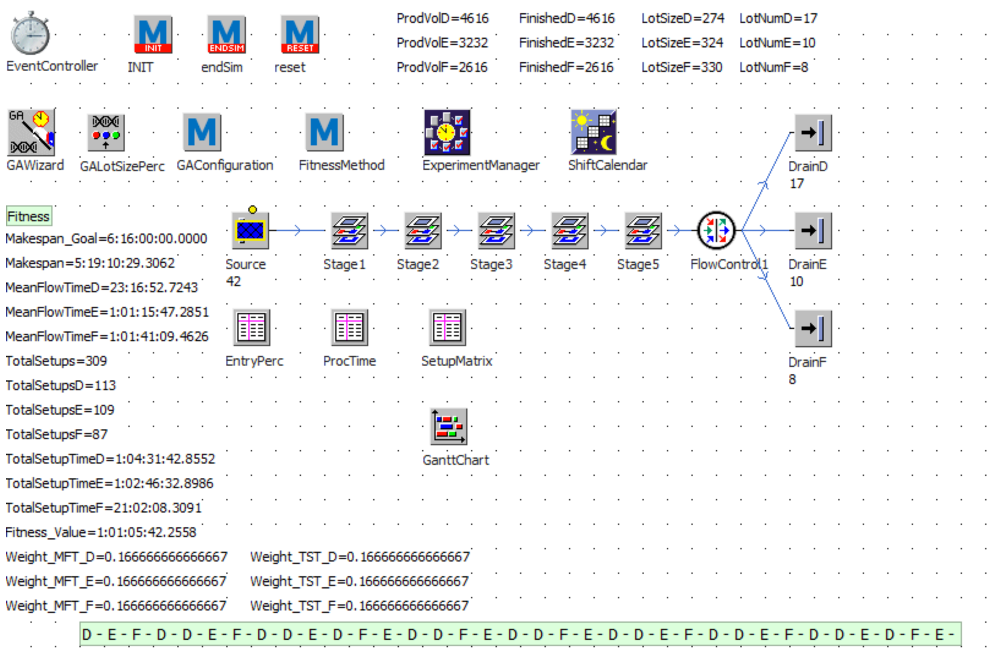


Fig. 4 FSMP model with displayed results

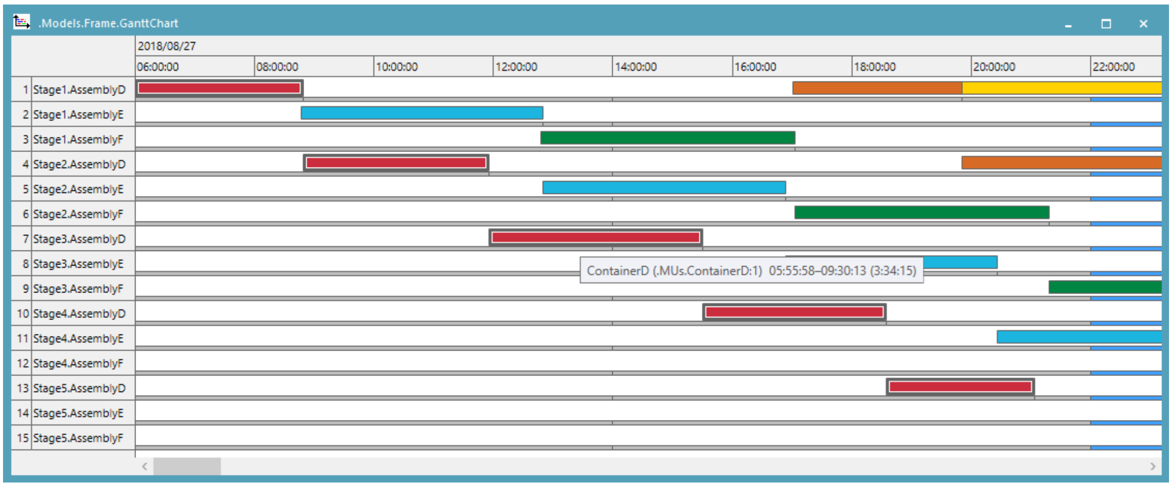


Fig. 5 The Gantt chart of products D, E, F for one working day

6. Conclusion

This paper deals with the problem of determining the batch size and the sequence of entering the product batches into the system, focusing on sequence-dependent setup times. A GA simulation approach was presented as a combination of stochastic modelling of discrete event simulation capabilities and intelligent GA search algorithm. A discrete model for flow shop with multiple processors (FSMP) was developed using the Tecnomatix Plant Simulation software package, specialized in manufacturing engineering as a tool to support the manufacturing planning of technologically similar products.

Based on the developed model and developed procedures within the genetic algorithm, optimal values for the mean flow time and total setup time are obtained, along with the primary condition of delivering the product on time. The GA simulation approach has shown that for the defined performance measures, the mean flowtime and the total setup time, the discrete model provides good solutions. The solution is applicable and shows the exact batch entering into the process and gives a detailed order and timing of entering a particular product batch into the default system. The main contribution of this paper is the simplicity and concision of the display solution that serves as a tool for manufacturing operations planning.

Acknowledgement

This work was financially supported by University of Rijeka, Croatia, contract number uniri-tehnic-18-223-6165 and by University of Rijeka, Croatia, contract number uniri-tehnic-18-100-5893. The authors would like to thank the company iTCR d.o.o. for their extensive help.

References

[1] Pinedo, M.L. (2008). *Scheduling: Theory, algorithms and systems*, 3rd edition, Springer, New York, USA, doi: 10.1007/978-0-387-78935-4.

[2] 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*, Vol. 78, No. 1, 40-56, doi: 10.1016/j.matcom.2007.06.002.

[3] Alfieri, A. (2009). Workload simulation and optimisation in multi-criteria hybrid flowshop scheduling: A case study, *International Journal of Production Research*, Vol. 47, No. 18, 5129-5145, doi: 10.1080/00207540802010823.

[4] Negahban, A., Smith, J.S. (2014). Simulation for manufacturing system design and operation: Literature review and analysis, *Journal of Manufacturing Systems*, Vol. 33, No. 2, 241-261, doi: 10.1016/j.jmsy.2013.12.007.

[5] Allahverdi, A. (2015). The third comprehensive survey on scheduling problems with setup times/costs, *European Journal of Operational Research*, Vol. 246, No. 2, 345-378, doi: 10.1016/j.ejor.2015.04.004.

[6] 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*, Vol. 37, No. 8, 1439-1454, doi: 10.1016/j.cor.2009.11.001.

[7] Gourgand, M., Grangeon, N., Norre, S. (2003). A contribution to the stochastic flow shop scheduling problem, *European Journal of Operational Research*, Vol. 151, No. 2, 415-433, doi: 10.1016/S0377-2217(02)00835-4.

[8] Wang, L., Zhang, L., Zheng, D.-Z. (2005). A class of hypothesis-test-based genetic algorithms for flow shop scheduling with stochastic processing time, *The International Journal of Advanced Manufacturing Technology*, Vol. 25, No. 11-12, 1157-1163, doi: 10.1007/s00170-003-1961-y.

[9] Yang, T., Kuo, Y., Cho, C. (2007). A genetic algorithms simulation approach for the multi-attribute combinatorial dispatching decision problem, *European Journal of Operational Research*, Vol. 176, No. 3, 1859-1873, doi: 10.1016/j.ejor.2005.10.048.

[10] Azadeh, A., Moghaddam, M., Geranmayeh, P., Naghavi, A. (2010). A flexible artificial neural network – Fuzzy simulation algorithm for scheduling a flow shop with multiple processors, *The International Journal of Advanced Manufacturing Technology*, Vol. 50, No. 5-8, 699-715, doi: 10.1007/s00170-010-2533-6.

[11] Chaudhry, I.A., Usman, M. (2017). Integrated process planning and scheduling using genetic algorithms, *Tehnički Vjesnik – Technical Gazette*, Vol. 24, No. 5, 1401-1409, doi: 10.17559/TV-20151121212910.

[12] Meolic, R., Brezočnik, Z. (2018). Flexible job shop scheduling using zero-suppressed binary decision diagrams, *Advances in Production Engineering & Management*, Vol. 13, No. 4, 373-388, doi: 10.14743/apem2018.4.297.

[13] Hendizadeh, S.H., ElMekkawy, T.Y., Wang, G.G. (2007). Bi-criteria scheduling of a flowshop manufacturing cell with sequence dependent setup times, *European Journal of Industrial Engineering*, Vol. 1, No. 4, 391-413, doi: 10.1504/EJIE.2007.015388.

[14] Lin, S.-W., Ying, K.-C. (2012). Scheduling a bi-criteria flowshop manufacturing cell with sequence-dependent family setup times, *European Journal of Industrial Engineering*, Vol. 6, No. 4, 474-496, doi: 10.1504/EJIE.2012.047666.

[15] Lee, G.-C. (2009). Estimating order lead times in hybrid flowshops with different scheduling rules, *Computers & Industrial Engineering*, Vol. 56, No. 4, 1668-1674, doi: 10.1016/j.cie.2008.10.016.

[16] Dugardin, F., Yalaoui, F., Amodeo, L. (2010). New multi-objective method to solve reentrant hybrid flow shop scheduling problem, *European Journal of Operational Research*, Vol. 203, No. 1, 22-31, doi: 10.1016/j.ejor.2009.06.031.

[17] Ladhari, T., Msakni, M.K., Allahverdi, A. (2012). Minimizing the total completion time in a two-machine flowshop with sequence-independent setup times, *Journal of the Operational Research Society*, Vol. 63, No. 4, 445-459, doi: 10.1057/jors.2011.37.

[18] Ying, K.-C., Lee, Z.-J., Lu, C.-C., Lin, S.-W. (2012). Metaheuristics for scheduling a no-wait flowshop manufacturing cell with sequence-dependent family setups, *The International Journal of Advanced Manufacturing Technology*, Vol. 58, No. 5-8, 671-682, doi: 10.1007/s00170-011-3419-y.

[19] Galzina, V., Lujić, R., Šarić, T. (2012). Adaptive fuzzy particle swarm optimization for flow-shop scheduling problem, *Tehnički Vjesnik – Technical Gazette*, Vol. 19, No. 1, 151-157.

[20] Chen, W., Hao, Y.F. (2018). Genetic algorithm-based design and simulation of manufacturing flow shop scheduling, *International Journal of Simulation Modelling*, Vol. 17, No. 4, 702-711, doi: 10.2507/IJSIMM17(4)CO17.

[21] Yan, R., Li, M.M., Wei, W.C. (2018). Integrated production scheduling and distribution planning with a two-stage semi-continuous flow shop environment, *International Journal of Simulation Modelling*, Vol. 17, No. 3, 553-561, doi: 10.2507/IJSIMM17(3)CO15.

[22] Liu, R., Xie, X., Yu, K., Hu, Q. (2018). A survey on simulation optimization for the manufacturing system operation, *International Journal of Modelling and Simulation*, Vol. 38, No. 2, 116-127, doi: 10.1080/02286203.2017.1401418.

[23] Supsomboon, S., Vajasuvimon, A. (2016). Simulation model for job shop production process improvement in machine parts manufacturing, *International Journal of Simulation Modelling*, Vol. 15, No. 4, 611-622, doi: 10.2507/IJSIMM15(4)3.352.
[24] Schroeder, R.G. (1993). Operations management: Decision making in the operations function, McGraw-Hill Education, New York, USA.
[25] Brajković, T., Perinić, M., Ikonić, M. (2018). Production planning and optimization of work launch orders using genetic algorithm, *Tehnički Vjesnik – Technical Gazette,* Vol. 25 No. 5, 1278-1285, doi: 10.17559/TV-20161207195125.
[26] Sharma, P., Jain, A. (2015). Performance analysis of dispatching rules in a stochastic dynamic job shop manufacturing system with sequence-dependent setup times: Simulation approach, *CIRP Journal of Manufacturing Science and Technology*, Vol. 10, 110-119, doi: 10.1016/j.cirpj.2015.03.003.
[27] Mikac, T. (1994). *Optimization of the manufacturing system concept,* PhD thesis (in Croatian), Faculty of Engineering, Rijeka, Croatia.
[28] Koren, Y., Shpitalni, M. (2010). Design of reconfigurable manufacturing systems, *Journal of Manufacturing Systems*, Vol. 29, No. 4, 130-141, doi: 10.1016/j.jmsy.2011.01.001.
[29] Gen, M., Cheng, R. (1997). *Genetic algorithms and engineering design*, John Wiley & Sons, New York, USA, doi: 10.1002/9780470172254.

APEM journal

Advances in Production Engineering & Management
Volume 14 | Number 2 | June 2019 | pp 177–188
https://doi.org/10.14743/apem2019.2.320

ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

An improved flower pollination algorithm for optimization of intelligent logistics distribution center

Hu, W.[a,*]

[a]School of Economics and Management, Shanghai University of Electric Power, Shanghai, P.R. China

A B S T R A C T

It is easy to fall into local optimal solution in solving the optimal location of intelligent logistics distribution center by traditional method and the result of optimization is not ideal. For this, the study puts forward an optimization method of intelligent logistics distribution center based on improved flower pollination algorithm. This method uses the logic self-mapping function to carry out chaotic disturbance to the pollen grains, so that the pollen grain set lacking the mutation mechanism has strong self-adaptability, and the convergence of the optimal solution in the later stage of the algorithm is effectively prevented. The boundary buffer factor is used to buffer the cross-boundary pollen grains adaptively so as to prevent the algorithm from the local optimization, and the convergence speed and the optimization accuracy of the algorithm can be improved obviously in processing the optimal location of intelligent logistics distribution center. The convergence of the algorithm is analyzed theoretically by using the real number coding method, and the biological model and theoretical basis of the algorithm are given. The experimental results show that the proposed method has better performance than the traditional one, and the algorithm outperforms a genetic algorithm and particle swarm algorithm. It provides a feasible solution for the intelligent logistics distribution center location strategy. It affords a good reference for improving and optimizing the internal logistics of the manufacturing system and the operational efficiency of the entire intelligent logistics system.

A R T I C L E I N F O

Keywords:
Intelligent logistics;
Distribution center;
Optimization;
Intelligent optimization algorithm;
Flower pollination algorithm;
Intelligent location

**Corresponding author:*
hwshiep@shiep.edu.cn
(Hu, W.)

Article history:
Received 18 October 2018
Revised 21 April 2019
Accepted 15 May 2019

1. Introduction

With the increase of vehicle ownership in China, the problem of traffic jam becomes more and more serious. It's a bottleneck to choose the optimal route of distribution center in a real time manner according to the road condition in the field of intelligent logistics [1-2]. The traditional solution is to optimize the distribution center route under certain constraints, and realize the optimization of the distribution center route under the condition of only considering the single index such as the least cost or the shortest path [3-4]. With the increase of road condition and distribution task complexity, it is difficult to satisfy the demand of intelligent logistics distribution index with only single constraint condition. The optimization of distribution center's optimal route gradually evolves into a multi-constraint optimization problem that satisfies several expectations of distribution personnel in the road network [5-6]. Logistics distribution center is a bridge between supply and demand, its location optimization strategy is the core content of logistics network system analysis, and often determines the distribution system and mode of intelligent logistics system. Therefore, it's theoretically and practically significant to select a reasonable logistics distribution center, achieve the best balance between supply and demand, and thus improve the operation efficiency of the entire intelligent logistics system [7-8].

In recent years, researchers have adopted a lot of methods to study location strategies of intelligent logistics distribution center, and achieved a lot of influential results. Research scholars at home and abroad have deeply studied the location optimization of intelligent logistics center, and classified it into continuous location model and discrete location model [9-10]. The continuous method mainly solves the choice and decision-making of logistics node in plane area. Many scholars have studied the uncertainty of parameters in the model. Literature [11] studied the stochastic programming model of supply chain network under uncertain conditions. Document [12] delt with the location of commercial facilities in uncertain environments. Literature [13] established an interval node decision-making model of logistics network based on barycenter method, and constructed an interval decision-making model of multi-commodity and multi-node continuous logistics facility location. Literature [14] proposes a logistics distribution center location model based on bi-level programming. The distribution points and user cost are considered respectively in the upper and lower levels of the model, and a simple heuristic algorithm is used to optimize the model. Literature [15] used AFS and TOPSIS related theories to explore the location model of logistics distribution center. Literature [16] studied the location issue of distribution center based on rough set and interactive multi-objective fuzzy decision theory. These methods above focus on qualitative or quantitative research on the optimization of location of intelligent logistics distribution centers, and have achieved certain results, which provide a reference for further improving and optimizing the operational efficiency of the entire intelligent logistics system.

From existing researches, it's found that there are two problems in the route optimization of intelligent logistics distribution center: (1) Few researches focus on the specific application requirements of intelligent optimization algorithm in the route optimization of intelligent logistics distribution centers but the influence of the algorithm on the optimization precision and convergence speed is ignored; (2) The setting of the route optimization constraints is too simple to meet the optimization precision requirements under the current complex road conditions. Aiming at the problems in previous researches, this study proposes an optimization method of intelligent logistics distribution centers based on improved flower pollination algorithm.

Flower pollination algorithm (FPA) is a new swarm intelligence optimization algorithm. It is a stochastic global optimization algorithm which mimics the mechanism of self-pollination and cross-pollination of flowering plants in the biological world [17]. The algorithm has the characteristics of easy implementation, strong universality and fast convergence speed, and has wide application. The algorithm is simple to implement and has less parameters, the given conversion probability can set the conversion threshold of global search and local search, and the algorithm adopts Lévy flight mechanism, which has excellent global optimization effect, so the algorithm can be used to solve many complicated optimization problems. On the basis of analyzing the optimization nature of standard flower pollination algorithm, this study presents an improved flower pollination algorithm, analyzes the convergence of the algorithm, optimizes the location model of logistics distribution centers with this improved algorithm, and compares the obtained results with the results of genetic algorithm and standard particle swarm algorithm. The experimental results show that under the premise of satisfying supply and demand, the method proposed in this study obtains the best location scheme for intelligent logistics distribution centers, and its performance is superior to the other two algorithms. The scheme in this study has a good guiding significance for the concrete practice.

2. Intelligent logistics distribution center location model

The problem of the vehicle travel path belongs to the optimal scheduling problem, and the cost is reduced by solving the optimal driving route. The related theories and solving algorithms are of great significance for improving the efficiency of logistics transportation, so it has always been the focus of relevant experts. In recent years, research on the problem of vehicle travel paths has produced many results, such as vehicle routing problems with multiple stations, vehicle routing problems with time window constraints, and vehicle routing problems for loading and unloading cargo. At the same time, vehicle travel routes have applications in all aspects of life, such as product distribution, cargo transportation, and mitigation of traffic jams.

It is difficult to change the location of an intelligent logistics distribution center once it's determined. Therefore, in the process of constructing the distribution center location model, the factors such as fixed cost, management cost and maximum inventory capacity should be taken into account. In the logistics network system, the demand of the demand points should be less than or equal to the size capacity of the distribution center. Under the condition of satisfying the distance upper limit, it is necessary to find out the distribution center from the known demand points and distribute the goods to each demand point. Based on the above problems, the intelligent logistics distribution center location model can be expressed as:

$$\mathrm{Min}T \sum_{j=1}^{M} h_j C_j + \sum_{i=1}^{N}\sum_{j=1}^{M} g_j W_{ij} + \sum_{i=1}^{N}\sum_{j=1}^{M} W_{ij}\, d_{ij} z_{ij} \tag{1}$$

$$\text{s.t} \sum_{j=1}^{M} W_{ij} \le B_i, \quad (i = 1,2,\cdots,N) \tag{2}$$

$$\sum_{j=1}^{M} Z_{ij} = 1, \quad (i = 1,2,\cdots,N) \tag{3}$$

$$Z_{ij} \le h_j, \quad (i = 1,2,\cdots,N) \tag{4}$$

$$\sum_{j=1}^{M} h_j = p \tag{5}$$

$$d_{ij} \le l, \quad (i \in M, j \in N) \tag{6}$$

where, N represents a set of ordinal numbers for all demand points; M is a set of demand points selected as the distribution center; C_j is the cost of building a distribution center; $h_j \in \{0,1\}$, when it is 1, the point j is selected as the distribution center; g_j indicates the unit management cost of material circulation in the distribution centers; W_{ij} represents the demand at the demand point i; d_{ij} represents the distance between demand point i and its nearest distribution center j; $Z_{ij} \in \{0,1\}$ represents the service allocation relationship between the demand point and the distribution center, when it's 1, it indicates that the demand of the demand point i is supplied by the distribution center j, otherwise $Z_{ij} = 0$. l indicates the upper limit of the distance between the demand point and the distribution center. Eq. 2 indicates that the demand of users should be less than or equal to the size capacity of the distribution center; Eq. 3 represents ensuring that each demand point is served by a distribution center closest to it; Eq. 4 indicates that there is no customer at a location without a distribution center; Eq. 5 shows that p demand point(s) is (are) selected as distribution center(s); Eq. 6 shows that the distribution center supplies the nearby demand points only within a limited range.

In the material and product distribution part of the manufacturing system, this method can also be used to select the distribution center, and the target function can be changed or added according to its own needs. For example, when distributing the materials needed for manufacturing, if the arrival time of materials has a certain limit, you can add time constraints, and then use the improved flower pollination algorithm to solve. It can be seen that the method can be used not only in the manufacturing system but also in any system that needs to be distributed.

3. Improved flower pollination algorithm

3.1 Standard flower pollination algorithm

Flower pollination algorithm is a kind of random search algorithm constructed by simulating the process of flower pollination in the nature, which embodies the preference mechanism of the nature [18]. The bionic principle is that flowers breed their offspring by pollination. The pollination process can be carried out by insects or natural wind and water. Butterflies are attracted by the color and smell of flowers, and fly to X_{best} after collecting nectar on the pollen X_i, so as to realize pollen transfer between flowers, which is called cross-pollination or global pollination. Under the action of wind, the pollen transfer between adjacent flowers X_i and X_i is realized. This kind of pollination is called self-pollination or local pollination. Through global pollination and local pollination, flowers flexibly achieve the process of pollen transfer. The ideal conditions of the algorithm are assumed as follows:

- In the process of cross-pollination, pollinators carry out pollination through Lévy flight.
- Self-pollination is the pollination process of adjacent flowers under the action of natural force. It is a kind of local pollination. The pollination mode is determined randomly by transition probability $p \in [0,1]$.
- In general, each flowering plant can bloom a lot of flowers, producing millions of pollen gametes, and in order to simplify the problem, we make the hypothesis that each plant only blooms one flower and each flower only produces one pollen gamete, which means that a flower or pollen gamete corresponds to a solution to the optimization problem.

For the formal description of the flower pollination algorithm, the following mathematical model is established:

In global pollination, pollinator carry pollen grains for large-scale and long-distance search, so as to ensure the optimal individual pollination and propagation. In the process of global pollination, the formula for updating the position of pollen is:

$$X_i^{t+1} = X_i^t + L(X_i^t - g^*) \tag{7}$$

where, X_i^t is the spatial location corresponding to the i-th pollen in the t-th iteration; g^* is the spatial position of the optimal solution in the current iteration (i.e., the t-th iteration). The parameter $L > 0$ is the step size and follows the Levy distribution, as shown in Eq. 8.

$$L \sim \frac{\lambda \Gamma(\lambda) \sin(\pi\lambda/2)}{\pi} \frac{1}{s^{1+\lambda}}, (s >> s_0 > 0) \tag{8}$$

where, $\Gamma(\lambda)$ is a standard gamma function, parameter $1 < \lambda \leq 3$ is $\lambda = 1.5$ in this study.

The formula of pollen grain position renewal for local pollination of adjacent flowers is as follows:

$$X_i^{t+1} = X_i^t + \varepsilon(X_j^t - X_k^t) \tag{9}$$

where, X_j^t and X_k^t are the spatial positions of two pollen grains randomly selected, and ε is a random operator uniformly distributed on the [0,1] interval.

The transition probability P is the probability determining whether pollen grains are pollinated globally or locally, and is a constant, with $p \in [0,1]$. The higher the value of P is, the higher the probability of carrying out the local pollination activity is, and the lower the chance of carrying out the global pollination is.

Due to the fact that adjacent flowers are easier to pollinate, the probability of local pollination will be higher than that of global pollination. Through the study of parameters, it is found that the optimization problem $p = 0.8$ is the best choice of parameters. The process of optimizing the algorithm: randomly disperse the population of pollen grains in solution space, and by comparing the size of random production number *rand* and transition probability P, and determine whether each pollen grain will carry on global pollination or local pollination; after several movements and updates of positions, all the pollen grains will converge to the place of optimum fitness; so, the successful pollination and insemination of pollen grains can be realized, and the optimization can be achieved. The steps of the algorithm are as follows:

Step 1: Initialize basic parameters of the algorithm. Set the number of pollen grains as n, transition probability as P and the maximum value of the number of iterations as $MaxT$;

Step 2: Randomly initialize the positions of all pollen grains x_i^0 (i = 1,…,n), find out the pollen grain individuals with the best fitness in the initial situation, and record the current position and fitness of the individuals;

Step 3: Start the iteration, and update the positions of all pollen grains. First generate a random number *rand* between [0,1] and if $Rand > p$, the pollen grains are pollinated globally according to Eq. 7; otherwise, the pollen grains are locally pollinated according to Eq. 9. After the position update of all pollen grains is completed in this iteration, the current optimal position and fitness of each pollen grain are calculated, and the position and fitness of the optimal individual in the population are updated;

Step 4: Go to Step 5 when it reaches the maximum number of iterations, otherwise go to Step 3 and start the next iteration;

Step 5: Get the global optimal solution and the target function value at this time.

3.2 Chaos optimization strategy

The principle of chaotic search: first, the variable of chaotic space is mapped to the solution space of the problem by some rules, and then the solution space is traversed by virtue of the chaotic feature. There are many chaotic mapping models. The research shows that the chaotic sequence obtained by using logical self-mapping function is better than Logistic mapping. Therefore, the logical self-mapping function is used in this study to construct chaotic sequences to realize chaotic traversal, and the following expressions are given:

$$cx_j^{k+1} = 1 - 2 \times (cx_j^k)^2 \tag{10}$$

where, cx_j^k is the j-th dimensional component of a chaotic variable, k is the number of iteration steps, $cx_j^0 \in (-1,1), j = 1,2,\ldots,d$. When $cx_j^k \neq 0$, the chaotic sequence obtained by logical self-mapping model possesses dynamic characteristics of chaos and is sensitive to initial values. The algorithm steps are as follows:

Step 1: Set $k = 0$, the decision variable x_j^k is mapped to the chaotic variable cx_j^k between –1 and 1 according to Eq. 11.

$$cx_j^k = 2 \times \frac{x_j^k - x_{\min,j}}{x_{\max,j} - x_{\min,j}} - 1 \tag{11}$$

where $j = 1,2,\ldots,d$, $x_{\max,j}$ and $x_{\min,j}$ are the upper boundary and the lower boundary of the j-th dimensional variable search, respectively.

Step 2: According to cx_j^k, the carrier operation is performed with Eq. 4, and the calculation iterates to the next generation of the chaotic variable cx_j^{k+1}.

Step 3: The chaotic variable cx_j^{k+1} is converted into decision variable x_j^{k+1} according to Eq. 12.

$$x_j^{k+1} = \frac{1}{2} \times [(x_{\max,j} - x_{\min,j}) \times cx_j^{k+1} + x_{\max,j} + x_{\min,j}] \tag{12}$$

where $j = 1,2,\ldots,d$.

Step 4: The performance analysis and evaluation on the newly generated solution are carried out according to decision variable cx_j^{k+1}. If the chaotic search has reached the preset limit or the new solution is superior to the initial solution, the new solution is output as the result of the chaotic search, otherwise, set $k = k + 1$ and return to Step 2.

3.3 Boundary buffer factor

In the process of optimization with flower pollination algorithm, the position of pollen grains is likely to break through its boundary value. In this case, the traditional method is to limit the position of pollen grains to the interval $[-x_{\max}, x_{\max}]$. The advantages of this method are simple operation and small computation, but the boundary compulsory processing method is unfavorable to the convergence of the algorithm, resulting in a lot of errors. In order to solve this problem, the method of dynamically changing interval is adopted, the size of the boundary is scaled according to the situation, the out-of-boundary grains are buffered by the boundary buffer factor and processed according to the out-of-boundary position of the pollen grains. The specific operation is as follows:

When $x_{ij}(t) < a_j$, $x_{ij}(t)$ can be represented as:

$$x_{ij}(t) = a_j((1 - \text{sgn}a_jL) + \text{sgn}a_jL(|v_{ij}(t)|/v_{j\max})rand) \quad (13)$$

When $x_{ij}(t) > b_j$, $x_{ij}(t)$ can be represented as:

$$x_{ij}(t) = b_j((1 + \text{sgn}b_jL) - \text{sgn}b_jL(|v_{ij}(t)|/v_{j\max})rand) \quad (14)$$

In Eq. 13 and Eq. 14, sgn is symbolic functions, $L \in [0,1]$, a_j and b_j are lower and upper limits of the grain in the j-th dimension, respectively. From Eq. 13 and Eq. 14, the boundary buffer factors are treated according to the actual situation of pollen grains, and the actual situation and movement of pollen grains are fully considered. The simulation results show that the boundary buffer factor can effectively improve the convergence speed and optimization accuracy of pollen pollination algorithm.

3.4 Specific steps of solution

The specific steps of optimizing the location problem of intelligent logistics distribution center by using the improved flower pollination algorithm are as follows:

Step 1: Initialize basic parameters of the algorithm. Set the number of pollen grains as n, transition probability as P, the maximum value of the number of iterations as $MaxT$, and the largest number of iteration steps for chaotic search as $MaxC$.

Step 2: Randomly initialize the positions of all pollen grains x_i^0 (i = 1,…,n), find out the pollen grain individuals with the best fitness in the current pollen grain set, and record the current position and fitness of the individuals;

Step 3: Start the iteration, and update the positions of all pollen grains. First generate a random number $rand$ between [0,1] and if $rand > p$, the pollen grains are pollinated globally according to Eq. 1; otherwise, the pollen grains are locally pollinated according to Eq. 3. After the position update of all pollen grains is completed in this iteration, the current optimal position and fitness of each pollen grain are calculated, and the position and fitness of the optimal individual in the population are updated;

Step 4: The individuals of pollen grains are evaluated, the n individuals(in percentage) with the best fitness are selected as the best pollen set, and the remaining 1 – n poorer individuals (in percentage) are selected and replaced with randomly generated individuals of pollen grains.

Step 5: The current optimal solution of the individual after chaos optimization is obtained, and the search region is dynamically shrunk by Eqs. 7 and 8.

Step 6: Turn to Step 7 when the number of iterations is equal to $MaxT$ or reaches the set search accuracy; otherwise, turn to Step 3 to perform the next iteration;

Step 7: Output the global optimal solution and the corresponding objective function value at this time.

4. Convergence analysis of chaotic flower pollination algorithm

In the chaotic flower pollination algorithm, the pollination process of flowers is the premise of convergence of the algorithm, the long-distance search and Lévy flight mechanism of pollinators in the cross-pollination guarantee the convergence stability and global optimization, the self-pollination enhances the local optimization ability of the algorithm, and the chaos strategy enhances the swing of the algorithm near the local solution, and reduces search range, gets rid of the local interference, and accelerates the search speed.

If the component x_i (i = 1,2,…,n) of the pollen set $X = (x_1,x_2,…,x_n)$ is represented by the τ dimensional binary coded string, the code string can take 2^τ discrete values, which is equivalent to the 2^τ discrete values dividing the defined interval $[Lb_i, Ub_i]$, with the accuracy of $\varepsilon = (Ub_i - Lb_i)/2^\tau$. Therefore, according to this characteristic, the convergence of the chaotic flower pollination algorithm can be analyzed by real coding. If the search accuracy of the solution space is assumed to be ε, the solution space M can be understood as a discrete space, its size is $|M| = \prod_{i=1}^{n}(Ub_i - Lb_i)/\varepsilon$, where each point $x \in M$ is a pollen grain, so set its fitness $F = f(x)$ is a function of fitness, obviously $|F| \leq M$, so it can be written as $F = \{F^1, F^2, \ldots F^{|F|}\}$, where $F^1 > F^2 > \ldots > F^{|F|}$. According to the different fitness degree, the search space M can be divided into different non-empty subset $\{M^i\}$, which is defined as:

$$M^i = \{x | x \in M, f(x) = F^i\} \tag{15}$$

where $i = 1,2,\ldots,|F|$, then:

$$\begin{gathered} \sum_{i=1}^{|F|} |M^i| = |M|; M^i \neq \varphi, \forall i \in \{1,2,\ldots,|F|\}; \\ M^i \cap M^j = \varphi; M^i \cap M^j = \varphi, \forall i \neq j; U_{i=1}^{|F|} M^i = M \end{gathered} \tag{16}$$

For any two elements $x_i \in M^i$ and $x_j \in M^j$,

$$f(x_i) \begin{cases} > f(x_j), & i < j \\ = f(x_j), & i = j \\ < f(x_j), & i > j \end{cases} \tag{17}$$

It is easy to know that F^1 can be considered a global optimal solution F^*, and that individuals with fitness equal to F^* should be in M^1. In the iterative process of the algorithm, the number N of pollen populations remains unchanged, $p = \{x_1, x_2, \ldots, x_N\}$. Set P as a set, containing all populations, and since the pollen grains in the population are allowed to be the same, the number of possible populations is:

$$|\beta| = C_{|M|+N-1}^{N} \tag{18}$$

Then, in order to judge the quality of the populations, the fitness function of p can be defined as:

$$F(P) = \max\{f(x_i) | i = 1,2,\ldots,N\} \tag{19}$$

Similarly, according to the different fitness, P can be divided into $|F|$ non-empty sub-sets ($i = 1,2,\ldots,|F|$), where $|P^i|$ represents the size of the set P^i. The set P^1 includes all populations with fitness F^1.

Set p^{ij} ($i = 1,2,\ldots,|F|$; $j = 1,2,\ldots,|P^i|$) represents the j-th population of p^i. Under the action of the evolution operator, the probability $\Pr_{ij,kl}$ of p^{ij} transiting to p^{kl} represents the transition probability of any one of the populations from p^{ij} to p^k, and $\Pr_{i,k}$ represents the transition probability of any one of the populations from p^i to p^k, then:

$$\Pr_{ij,k} = \sum_{i=1}^{|p^k|} \Pr_{ij,kl}; \quad \sum_{k=1}^{|F|} \Pr_{ij,k} = 1; \quad \Pr_{i,k} \geq \Pr_{ij,k} \tag{20}$$

A square matrix $A \in R^{n \times n}$ is called:

- a nonnegative matrix, if $a_{ij} \geq 0, \forall i,j \in \{1,2,\ldots,n\}$;
- a primitive matrix, if A is nonnegative, and there is an integer $k \geq 1$, so $A^k > 0$;
- a random matrix, if A is nonnegative and $\sum_{j=1}^{n} a_{ij} = 1$, then $\forall i \in \{1,2,\ldots,n\}$;
- a reducible random matrix can perform a row-column transformation of the same form to obtain $A = \begin{bmatrix} S & 0 \\ R & T \end{bmatrix}$, where s is the m-order primitive matrix, and R and T≠0.

<u>*Definition 1*</u>**:** If an evolutionary algorithm converges to a global optimal solution, its sufficient and necessary conditions are:

$$\lim_{t \to \infty} \Pr\{f(P^t) = F^*\} = 1 \tag{21}$$

where, Pr represents probability and P^t represents the t-th generation population.

<u>*Theorem 1*</u>: Set $\Pr = \begin{bmatrix} S & 0 \\ R & T \end{bmatrix}$ is a reducible random matrix, S is the m-order primitive matrix and R and T≠0, then

$$\Pr^{\infty} = \lim_{n \to \infty} \Pr^n = \lim_{k \to \infty} \begin{bmatrix} S^n & 0 \\ \sum_{i=0}^{n-1} T^i R S^{n-i} & T^n \end{bmatrix} = \begin{pmatrix} S^{\infty} & 0 \\ R^{\infty} & 0 \end{pmatrix} \tag{22}$$

Obviously, $\Pr^{\infty}$ is a stable random matrix and $\Pr^{\infty} = [1,1,\ldots,1]^T[p_1, p_2, \ldots, p_n]$, $\sum_{j=1}^{n} p_{ij} = 1$, and $p_j = \lim_{k\to\infty} p_{ij}^{(k)} \geq 0$, when $p_j \begin{cases} > 0 & (1 \leq j \leq m) \\ = 0 & (m+1 \leq j \leq n) \end{cases}$.

Theorem 2: In the chaotic flower pollination algorithm, for $\forall i, k \in \{1,2,\ldots,|F|\}$,

$$\Pr_{i,k} \begin{cases} > 0, & k \leq i \\ = 0, & k > i \end{cases} \tag{23}$$

Prove: First, for $\forall P^{ij} \in p^i (j = 1,2,\ldots,|p^i|)$, $\exists X^* = (x_1^*, x_2^*, \ldots, x_n^*) \in p^{ij}$, so f(x*) = F^i, $\forall p^{kl} \in p^k$, $k = 1,2,\ldots,|F|$, $l = 1,2,\ldots,|p^k|$, $\exists X' = (x_1', x_2', \ldots, x_n') \in p^{kl}$ and f(X') = F^k. Under the action of evolutionary operators, p^{ij} is transited to be p^{kl}, if p^{ij} of the t-th generation is evolved to be p^{kl} of the $(t + 1)$-th generation, and for convenience, record them as L^t and L^{t+1}. The chaotic flower pollination algorithm has the optimization strategy and the historical optimal solution F* is preserved in each pollination process. Therefore, the optimal position of pollen is F* at L^t, and under the condition that the population remains unchanged, the next generation L^{t+1} is produced through pollination behavior, and the best pollen grain F' in L^{t+1} is compared with F*, if F' is better than F*, the preserved historical optimal solution is replaced by F', otherwise unchanged, then:

$$f(L^{t+1}) \geq f(L^t) \Rightarrow k \leq i \Rightarrow \forall k > i,\ \Pr_{ij,kl} = 0 \Rightarrow \forall k > i,\ \Pr_{ij,k} = \sum_{l=1}^{|p^k|} \Pr_{ij,kl} = 0 \Rightarrow$$

$\forall k > i$ and $\Pr_{i,k} = 0$.

Secondly, in each pollination process, the algorithm is iterated and the optimal solution is found. The behavior selection is made according to the principle of the fastest or better speed. Therefore, set individual X', with the fitness of f(X') = F^k, $k \leq i$, and X' has r components $(x_1', x_2', \ldots, x_r')$ different from X*, the probability of x' generated by x* through Levy flight operator is $\Pr = (1 - \frac{1}{n})^{(n-r)} \prod_{i=1}^{r} \phi(x_i^* - x_i') > 0$, where φ is the probability density function for levy distribution. So the probability of p^{ij} evolving to be p^k is greater than zero, so $\forall k \leq i$ and $\Pr_{i,k} \geq \Pr_{ij,k} > 0$.

Theorem 3: The chaotic flower pollination algorithm is globally convergent.

Prove: Each $\Pr_i$, i = 1,2,...,$|F|$ can be viewed as a state on a time-aligned finite Markov chain. According to Theorem 1, the transition probability matrix of the algorithm is expressed as:

$$\Pr = \begin{bmatrix} \Pr_{1,1} & 0 & \cdots & 0 \\ \Pr_{2,1} & \Pr_{2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \Pr_{|F|,1} & \Pr_{|F|,2} & \cdots & \Pr_{|F|,|F|} \end{bmatrix} = \begin{bmatrix} S & 0 \\ R & T \end{bmatrix}$$

where $R > 0$, $T \neq 0$ and $S = 1$.
According to Theorem 1,

$$\Pr^{\infty} = \lim_{k\to\infty} \Pr^k = \lim_{k\to\infty} \begin{bmatrix} S^k & 0 \\ \sum_{i=0}^{k-1} T^i R C^{k-i} & T^k \end{bmatrix} = \begin{bmatrix} S^{\infty} & 0 \\ R^{\infty} & 0 \end{bmatrix}$$

where $S^{\infty} = 1$ and $R^{\infty} = (1,1,\ldots,1)^T$, so $\Pr^{\infty}$ is a stable random matrix, and

$$\Pr^{\infty} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix}$$

Therefore, regardless of the fitness state of the current population, it will converge to the optimal fitness state with probability 1 after infinite evolution, so the improved flower pollination algorithm is globally convergent.

5. Results and discussion

In this section, the results of optimization of the proposed method, genetic algorithm and standard particle swarm optimization algorithm are compared in Matlab environment to prove the effectiveness and feasibility of the proposed method. During the experiment, the coordinates of 31 cities are collected, and the position coordinates of each demand point and the material demand are given in Table 1. The material demand in the table is the standardized value and does not represent the actual value. The dimensionless process is applied to 31 logistics points. The dimensionless process uses the quotient of the mean value of each factor variable to obtain dimensionless data.

The parameters of flower pollination algorithm are as follows: p = 0.8, $beta$ = 1.5, the number of chaos iterations k = 10, and the ratio of selecting the best pollen set is n = 20. The improved flower pollination algorithm is used to optimize the location model of the intelligent logistics distribution center. The point value is set to be 6, which indicates that 6 logistics distribution centers are selected from 31 demand points.

In the simulation experiment, the final data is the average result after running the program for 100 times, so as to reduce the error. The convergence curve of the improved flower pollination algorithm is shown in Fig. 1.

The location scheme of the intelligent logistics distribution center is shown in Fig. 2. The box represents the distribution center, and the dot represents the demand point. The connection between the box and the dot indicates that the logistics distribution center is responsible for the distribution of the materials at a demand point. The location model of the intelligent logistics distribution center is optimized by the method proposed in this study, and the location scheme is [22 17 6 29 6 11].

In the process of solving the location optimization of the intelligent logistics distribution center with the improved flower pollination algorithm, chaos optimization and boundary buffer factor are added to effectively avoid the algorithm falling into local optimization, so that the convergence speed and optimization accuracy of the algorithm are obviously improved, reflecting that the flower pollination algorithm needs no accurate description of the problem in solving the practical problem, but can quickly obtain the optimal solution within a certain limit. Table 2 shows the performance comparison between the proposed scheme, genetic algorithm and standard particle swarm optimization (PSO) in optimizing the location model of intelligent logistics distribution center.

Table 1 Position of demand point and material demand

No.	Coordinates	Demand	No.	Coordinates	Demand
1	(2378,3172)	50	17	(4462,3641)	90
2	(1689,2623)	60	18	(2685,3252)	80
3	(4973,4582)	60	19	(4891,3431)	50
4	(2732,3452)	40	20	(2541,3964)	40
5	(1367,2473)	70	21	(4571,3115)	90
6	(3874,3994)	70	22	(2714,3574)	60
7	(4953,2363)	30	23	(4325,3683)	70
8	(3852,3372)	60	24	(4232,3231)	60
9	(2454,3068)	60	25	(2232,3462)	50
10	(2122,1374)	80	26	(4436,3285)	50
11	(4132,3879)	40	27	(3340,4112)	30
12	(2244,3896)	30	28	(1203,2183)	60
13	(4564,1598)	30	29	(4632,3461)	80
14	(3933,4212)	80	30	(2262,3365)	50
15	(3284,1475)	30	31	(3422,4316)	70
16	(3433,2756)	50			

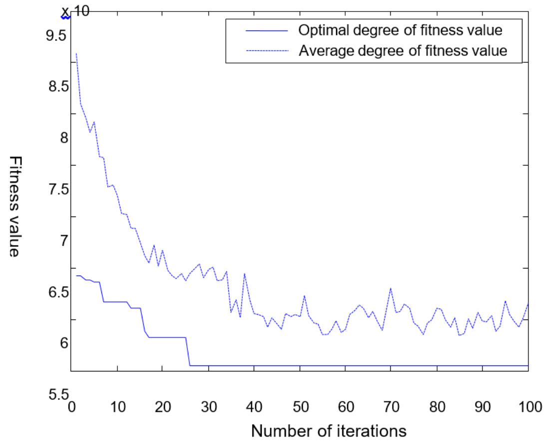


Fig. 1 Convergence curve

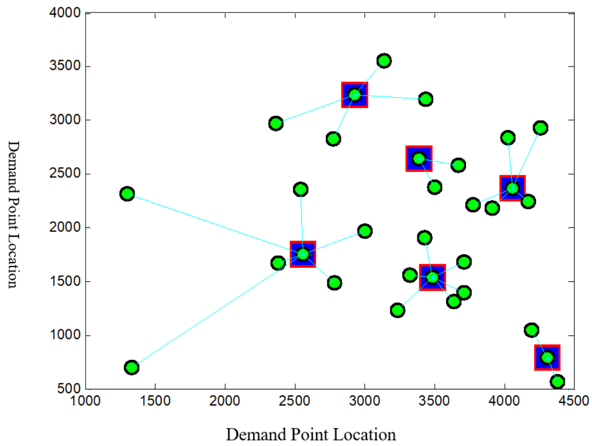


Fig. 2 Location scheme

Table 2 Performance comparison of several algorithms

Algorithm	Location scheme	Average distribution cost	Number of iterations	Running time/s
GA	22 17 6 29 6 11	1892	110	36.3
PSO	22 17 6 29 6 11	1363	60	15.5
Improved FPA	22 17 6 29 6 11	1012	26	11

Table 2 shows that when genetic algorithm, standard particle swarm optimization algorithm and the method in this study optimize the location model of the intelligent logistics distribution center, the optimal scheme can be obtained, The average delivery cost required by the genetic algorithm is 1892, the number of its iterations is 110, and its running time is 36.3 seconds. The average delivery cost required by the standard particle swarm optimization algorithm is 1363, the number of its iterations is 60, and its running time is 15.5 seconds. The average delivery cost required by the improved flower pollination algorithm is 1012, the number of its iterations is 26, and its running time is 11 seconds. Therefore, the performance of the improved flower pollination algorithm is generally better than the other two algorithms. In order to further verify the universality of the improved flower pollination algorithm in solving such problems, the number of logistics demand points is changed, and simulation experiments are carried out on different initial data respectively. The simulation results show that when the number of logistics demand points is large, the advantages of the algorithm presented in this study are more obvious. At the same time, we also find that the flower pollination algorithm has the advantages of self-organizing ability, distributed operation and positive feedback, and can perceive the change of surrounding road condition parameters in real time in complex road network. It is especially suitable for the path optimization problem of intelligent logistics distribution center under multi-constraint conditions.

6. Conclusion

The optimal location model of intelligent logistics distribution centers is a non-convex and non-smooth nonlinear model with complex constraints, which belongs to NP-hard problem. This study uses the intelligent optimization method to the path optimization problem of the intelligent logistics distribution center, and proposes an improved flower pollination optimization algorithm. According to the actual characteristics of the whole logistics network system, a location model of the intelligent logistics distribution center is constructed. According to the biological characteristics of flower pollination, the realization principle of the flower pollination algorithm is described from the mechanism, and the convergence of the algorithm is analyzed by real coding method. The optimization performance of the algorithm is analyzed in detail, and the biological model and theoretical basis of the algorithm are given.

In order to avoid the local optimization and improve the convergence speed and optimization accuracy of the algorithm, this study introduces chaos optimization and boundary buffer factor.

By means of logical self-mapping function, the pollen grains are disturbed by chaos, which makes the pollen grain set without mutation mechanism have strong self-adaptive ability. Meanwhile, the size of the boundary is dynamically scaled according to the actual situation, and the boundary buffer factor is used to buffer the pollen grains crossing the boundary. The experimental results show that compared with other methods, the method proposed in this study is more effective in solving the location optimization of the intelligent logistics distribution center, and can quickly and accurately find the best logistics distribution center for demand points.

Acknowledgement

This research is supported by Humanities and Social Sciences project of the Ministry of Education of China (No.17YJCZH062).

References

[1] Feng, F., Pang, Y., Lodewijks, G., Li, W. (2017). Collaborative framework of an intelligent agent system for efficient logistics transport planning, *Computers & Industrial Engineering*, Vol. 112, 551-567, doi: 10.1016/j.cie.2016.12.044.

[2] Kirch, M., Poenicke, O., Richter, K. (2017). RFID in logistics and production – Applications, research and visions for smart logistics zones, *Procedia Engineering*, Vol. 178, 526-533, doi: 10.1016/j.proeng.2017.01.101.

[3] Yang, L., Ji, X., Gao, Z., Li, K. (2007). Logistics distribution centers location problem and algorithm under fuzzy environment, *Journal of Computational and Applied Mathematics*, Vol. 208, No. 2, 303-315, doi: 10.1016/j.cam.2006.09.015.

[4] Tang, X., Lehuédé, F., Péton, O. (2016). Location of distribution centers in a multi-period collaborative distribution network, *Electronic Notes in Discrete Mathematics*, Vol. 52, 293-300, doi: 10.1016/j.endm.2016.03.039.

[5] Loree, N., Aros-Vera, F. (2018). Points of distribution location and inventory management model for post-disaster humanitarian logistics, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 116, 1-24, doi: 10.1016/j.tre.2018.05.003.

[6] Gutjahr, W.J., Dzubur, N. (2016). Bi-objective bilevel optimization of distribution center locations considering user equilibria, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 85, 1-22, doi: 10.1016/j.tre.2015.11.001.

[7] Holzapfel, A., Kuhn, H., Sternbeck, M.G. (2018). Product allocation to different types of distribution center in retail logistics networks, *European Journal of Operational Research*, Vol. 264, No. 3, 948-966, doi: 10.1016/j.ejor.2016.09.013.

[8] Tan, J., Jiang, G., Wang, Z. (2017). Evolutionary game of information sharing on supply chain network based on memory genetic algorithm, *Journal Européen des Systèmes Automatisés*, Vol. 50, No. 4-6, 507-519, doi: 10.3166/JESA.50.507-519.

[9] Pham, T.Y., Ma, H.M., Yeo, G.T. (2017). Application of fuzzy delphi TOPSIS to locate logistics centers in Vietnam: The logisticians' perspective, *The Asian Journal of Shipping and Logistics*, Vol. 33, No. 4 211-219, doi: 10.1016/j.ajsl.2017.12.004.

[10] He, Y., Wang, X., Lin, Y., Zhou, F., Zhou, L. (2017). Sustainable decision making for joint distribution center location choice, *Transportation Research Part D: Transport and Environment,* Vol. 55, 202-216, doi: 10.1016/j.trd.2017.07.001.

[11] Sun, H., Gao, Z., Wu, J. (2008). A bi-level programming model and solution algorithm for the location of logistics distribution centers, *Applied Mathematical Modelling*, Vol. 32, No. 4, 610-616, doi: 10.1016/j.apm.2007.02.007.

[12] Hua, X., Hu, X., Yuan, W. (2016). Research optimization on logistics distribution center location based on adaptive particle swarm algorithm, *Optik*, Vol. 127, No. 20, 8443-8450, doi: 10.1016/j.ijleo.2016.06.032.

[13] Malik, S., Kumari, A., Agrawal, S. (2015). Selection of locations of collection centers for reverse logistics using GTMA, *Materials Today: Proceedings*, Vol. 2, No. 4-5, 2538-2547, doi: 10.1016/j.matpr.2015.07.199.

[14] Coelho, E.K.F., Mateus, G.R. (2017). A capacitated plant location model for reverse logistics activities, *Journal of Cleaner Production*, Vol. 167, 1165-1176, doi: 10.1016/j.jclepro.2017.07.238.

[15] Zhao, L., Li, H., Li, M., Sun, Y., Hu, Q., Mao, S., Li, J., Xue, J. (2018). Location selection of intra-city distribution hubs in the metro-integrated logistics system, *Tunnelling and Underground Space Technology*, Vol. 80, 246-256, doi: 10.1016/j.tust.2018.06.024.

[16] Vieira, J.G.V., Toso, M.R., da Silva, J.E.A.R., Ribeiro, P.C.C. (2017). An AHP-based framework for logistics operations in distribution centres, *International Journal of Production Economics*, Vol. 187, 246-259, doi: 10.1016/j.ijpe.2017.03.001.

[17] Gao, M., Shen, J., Jiang, J. (2018). Visual tracking using improved flower pollination algorithm, *Optik*, Vol. 156, 522-529, doi: 10.1016/j.ijleo.2017.11.155.

[18] Salgotra, R., Singh, U. (2017). Application of mutation operators to flower pollination algorithm, *Expert Systems with Applications*, Vol. 79, 112-129, doi: 10.1016/j.eswa.2017.02.035.

[19] Yang, X.-S., Karamanoglu, M., He, X. (2014). Flower pollination algorithm: A novel approach for multiobjective optimization, *Engineering Optimization*, Vol. 46, No. 9, 1222-1237, doi: 10.1080/0305215X.2013.832237.
[20] Wang, C.L., Li, S.W. (2018). Hybrid fruit fly optimization algorithm for solving multi-compartment vehicle routing problem in intelligent logistics, *Advances in Production Engineering & Management*, Vol. 13, No. 4, 466-478, doi: 10.14743/apem2018.4.304.
[21] Hu, W., Hou, Y., Tian, L., Li, Y. (2015). Selection of logistics distribution center location for SDN enterprises, *Journal of Management Analytics*, Vol. 2, No. 3, 202-215, doi: 10.1080/23270012.2015.1077481.
[22] Wu, Z. (2014). Optimization of distribution route selection based on particle swarm algorithm, *International Journal of Simulation Modelling*, Vol. 13, No. 2, 230-242, doi: 10.2507/IJSIMM13(2)CO9.
[23] McFarlane, D., Giannikas, V., Lu, W. (2016). Intelligent logistics: Involving the customer, *Computers in Industry*, Vol. 81, 105-115, doi: 10.1016/j.compind.2015.10.002.

APEM journal

Advances in Production Engineering & Management
Volume 14 | Number 2 | June 2019 | pp 189–200
https://doi.org/10.14743/apem2019.2.321

ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

Impact of cooperation uncertainty on the robustness of manufacturing service system

Liang, P.P.[a,*], **Li, C.W.**[b,*]

[a]School of Economics and Management, Anqing Normal University, Anqing, P.R. China
[b]School of Economics, Huaibei Normal University, Huaibei, P.R. China

A B S T R A C T

The cooperation between enterprises is actually at a certain risk of interruption, which has a significant impact on the robustness of manufacturing service system (MSS). Evaluating MSS' robustness is integral to production and service provisioning, and thus the influence mechanism should be clearly revealed for assisting professionals in the company in improving the robust performance. In this paper, we present an effective methodology for explicating the impact of cooperation uncertainty on the robustness of MSS from a complex system standpoint. This methodology characterizes MSS as a topological network consisting of serval service subsystems, and constructs the measure metrics system of which the validity and applicability are proved theoretically from the dimension of structure and performance. Furthermore, it simulates the cooperation interruption from four different scenarios with algorithms, and finally takes an elevator manufacturing service network as the case to illustrate this novel methodology. The simulation findings suggest that identifying the critical paths in MSS and standardizing the cooperation mechanism within and among core manufacturing service principals outperform the other measures in improving the robustness of MSS.

A R T I C L E   I N F O

Keywords:
Manufacturing service system;
Complex system;
Robustness;
Robustness metric;
Cooperation uncertainty

**Corresponding author:*
cwli2008@163.com
(Li, C.W.)
614lpp@163.com
(Liang, P.P.)

Article history:
Received 16 November 2018
Revised 21 April 2019
Accepted 6 May 2019

1. Introduction

The area of manufacturing service management has gained increased attention in recent years [1]. Under a strong trend towards servitization, no enterprise can complete all business process from the initial acceptance of customer orders to the final provision of products or services. Accordingly, meeting the specific service requirements can't do without the efficient synergy of multiple service principals, which together constitute a Service-oriented Manufacturing Network (SMN) [2]. From a complex system standpoint, MSS can be defined as an alliance system with a temporary interest bargain, formed through the dynamic coupling of all heterogeneous enterprises in SMN. Compared to the traditional manufacturing, this new pattern emphasizes the synergy between service principals. But as a greatest obstacle of synergy, the loosely coupling with or between enterprises has been in an urgent need of attention, which may severely disrupt the robust operation of MSS [3]. In practice, these loosely coupling phenomena are particularly prominent and common. Inconsistency of collaborative manufacturing, interactive interruption of information system, termination of customer participation and uncertainty of partnership, especially in a service-oriented logic market, will lead to a huge fluctuation in the structure and performance of MSS.

Unexpected changes make the cooperation have a certain risk in reality. Whether cooperation continues and gains benefit or not, they depend not only on the complementarity of technologies and resources and the risk of technologies and market, but also on the relationship management between partners. It's certain that the cooperation uncertainty have an effect on the robustness of MSS, reflected in service quality, service cost and service response time. As such, adopting what kinds of measures to guarantee system robustness requires us to reveal the influence mechanism in different scenarios. The findings can provide the professionals in the company with a basic theoretical support for the follow-up network planning and construction of cooperative management mechanism. The related researchers mostly focus on the specific design of contract and the revelation and perfection of relational governance mechanism [4, 5], or focus on the robust operation process from the technology control perspective [6], and rarely pay attention to the impact of cooperation uncertainty on the robustness of MSS.

On account of the fact that MSS is a complex system, this paper focuses on how MSS' robustness is affected by cooperation uncertainty and taking what kinds of measures to guarantee system robustness. The remainder of this paper is in the following format. In Section 2, an overview of the relevant researches is presented. Section 3 discusses serval typical manifestations of cooperation uncertainty in the service-oriented environment, while Section 4 develops the model based on complex network theory and proposes four different simulation strategies with algorithms. Section 5 conducts case analysis and simulation. The concluding remarks and comment relating to future research directions are provided in Section 6.

2. Literature review

2.1 Cooperation uncertainty

Cooperation can be defined as a kind of joint action of reciprocity and mutual benefit from an economic perspective, what both sides concern is the benefits of joint action [7]. High satisfactions, timely and accurate distribution, faster speed of development, have been proved to be the critical sources of competitive advantage in a service-oriented logic market [8, 9]. Ring *et al.* found that cooperation mainly faces two types of risk, respectively the future environment risks and cooperation risks [10]. Future environment risks come from the uncertainty of internal and external environment, such as the changes in the market environment, the enterprise strategy adjustment, the uncertainty of natural and artificial calamities, while cooperation risks are mainly embodied in two aspects: relationship risk and performance risk. The former refers to the possibility that the partners' incomplete cooperative behaviours are harmful to cooperation, such as various opportunistic behaviours, information asymmetry, and unanticipated benefit which makes the phenomenon of moral hazard and adverse selection occurrence frequently [11]. The latter mainly refers to that the uncertainty of cooperation still exists with the partners' complete cooperative behaviours by reason of the unfulfilled expected returns. In addition, the cases of terminating voluntarily cooperation relationships are often occurred in business practice.

The general description of uncertainty is the probability of the value of a variable or of the occurrence of an event, and also is unpredictable in advance. As long as there is cooperation, the probability of the occurrence of cooperation interruption exists. In order to minimize the impact of uncertainty, the correlation between the evolution of organization cooperation pattern and uncertainty [12], uncertainty in R&D (Technology) innovation cooperation [13], the optimization of job-shop scheduling [14], and the impact of market uncertainty on the cooperation behaviours and performances [15], had gotten a lot of attention. Through sorting out the relevant researchers, studying cooperation uncertainty from a holistic or a complex system standpoint still lacks.

2.2 Robust operation

The researches on robust operation in management field concentrate mainly on supply chain system and manufacturing system. Of this, supply chain system observes and studies the situations of parameter, time-lag and exogenous destabilization [16, 17], and manufacturing system follows with interests of uncertain situations of resource-constrained [18], time-constrained

[19], changing demand [20]. More related researches of robust management mainly discuss how to guarantee the robustness of product and manufacturing process by technology control [21], framework design [22], and robust system development [23]. Most are achieved by using state-space analysis method or structured analysis tools. Although these methods have the shortcomings of high computational complexity in the context of dynamic systems, they do improve the ability of working smoothly in accordance with robustness measure and robustness analysis under an uncertain environment.

A commonly-held definition for robustness in manufacturing system is: "*the ability of maintain working smoothly within an acceptable range under the expected or unexpected changing*" [24]. As mentioned above, the available literature offers up a wide range of robustness metrics. Some focus mainly on graph theory concepts such as node degree, connectivity, network efficiency, clustering coefficient and two-terminal reliability [25, 26], while others consider the services supported by R value, or stability, or elasticity, or net variation [27, 28]. The selection of specific robustness metrics is usually in accordance with the characteristics of research object, which is proved be efficient.

3. Cooperation uncertainties in MSS

3.1 Uncertainty of R&D cooperation

The key difference between cooperative R&D and traditional R&D organized by a single enterprise lies in the uncertainty of partner. It is clear that an important feature of service-oriented manufacturing mode is the provision of personalized products and services for customers, which is not only dependent on its own power, but also the cooperation of upstream and downstream partners and the participation of customers. With the exception of the inherent uncertainty of market and technology, there still exists the uncertainty in behaviours arising from trading costs in the process of cooperation. In practice, the appearance of behaviour uncertainty such as leakage of knowledge, hitchhiking and ripping off, may result in the final termination of cooperation, furthermore, would have a direct effect on the manufacturing process and the response time to customers' demands, which deviates from the service philosophy of this new manufacturing pattern.

3.2 Uncertainty of collaborative manufacturing

Collaborative manufacturing is the core content of service-oriented manufacturing, mainly reflected in three aspects: synergism of all departments and information systems, collaborative manufacturing between all sub factories and collaborative manufacturing based on the whole supply chain. These factors such as complex relationships of structure, interest paradox between nodes and information asymmetry, would make the entire process of collaborative manufacturing with greater uncertainty. For instance, there will be a Butterfly effect caused by collaboration uncertainty, the lack of effective communication and coordination among departments, sub factories, and supply chain partners. Also, there will be a Matthew effect arising from purchasing uncertainty, information opaque and individual consummate interest, while running uncertainty and multipoint concurrency of producing tasks would lead to a Bottleneck effect of resources. In addition, there still exists much purchasing and producing uncertainty caused by natural or man-made disasters.

3.3 Uncertainty of cooperative marketing

To enhance the quick response ability to market demand, service-oriented manufacturing enterprise needs to integrate the transversal and longitudinal superior resources, and strengthen cooperative marketing among manufacturing enterprise, salesman and the third party logistics enterprises (e.g., cooperative marketing among Apple, eBay and USPS or among Huawei, Tmall and SF-express). Nevertheless, the ability of marketing partners, communication and coordination, lack of credibility, unrealistic expectations, etc., would cause greater uncertainty in cooperative marketing and further may lead to the final cooperation termination that has an unpredictable loss.

4. Robustness measurement

4.1 Model setup of system topology structure

As shown in Fig. 1, MSS can be divided into three subsystems: service production subsystem, production service subsystem and customer management subsystem. Each subsystem covers a cluster of similar or complementary work, formed through dynamic coupling of all the isomeric and heteroid enterprises in SMN, which has the main characteristics of a complex system. Compared to traditional manufacturing network, SMN has a wider range of source choices, and its structure has changed from tree structure to multi-loop network structure. Similarly, the relationship between network organization and synergetic effect is not a simple linear one. For the reasons mentioned above, it's more suitable and more efficient to use network analysis theory, method and tools to study the robustness of MSS.

As such, we can analyze the impact of cooperation uncertainty on the robustness of MSS based on complex network theory. Complex network is the abstract representation of a complex system, so we can regard the cooperative enterprises as the network nodes and the collaborative relationships as the network edges. Hence, the manufacturing service network can be expressed by the undirected network graph G(V, E). $V = \{v_1, v_2, v_3, \dots v_n\}$ and $E = \{(v_i, v_j), i, j = 1,2, \dots, n\}$, respectively, denote the node set and the edge set of SMN. In addition, we use $\mathrm{W} = (\mathrm{w}_{ij})_{n\times n}$ to represent the adjacency matrix of network. If there exists collaborative relationship between network nodes v_i and v_j, $w_{ij} = 1$, else $w_{ij} = 0$.

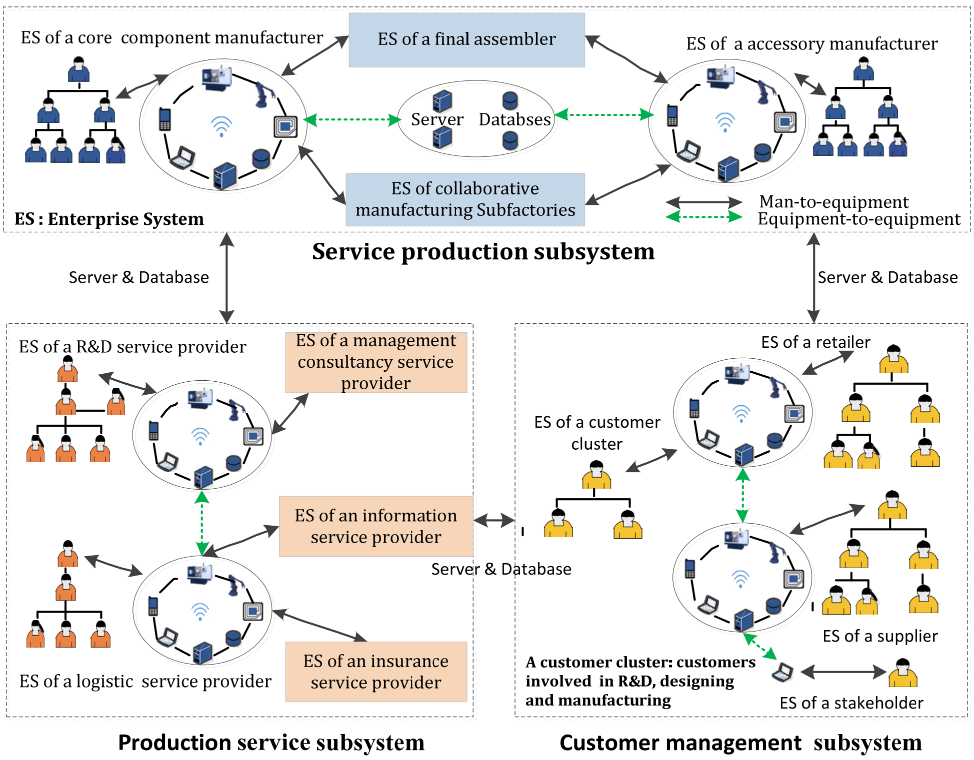


Fig. 1 A framework of MSS

4.2 Robustness metrics

In terms of MSS, there are a large number of random, fuzzy and uncertain factors, which may lead to a decline in collaboration among service principals and be seriously likely to a disruption of partnerships and also a loss of its structural functions. Before introducing our methodology, it's important to define the robustness of MSS, which is: "*the ability to maintain its basic structure and performance under the termination of collaboration relationships caused by random or targeted factors*". Accordingly, learning the robustness ought to be from the structural and performance dimensions.

Structural robustness. The definition of structural robustness is: "*the ability to resist the decline of network connectivity caused by the inactivation of service nodes*". We choose natural connectivity as the structural robustness metric, firstly proposed by Wu *et al.* in 2010. It's easy to understand that the higher the redundancy of alternative routes is, the better the connectivity of network structure and the structural robustness will be. Especially in a service-oriented market environment, the main way to deal with the interference of uncertainties is to enhance the elasticity of MSS, while the guarantee of elasticity depends on the proper redundancy of system structure. Taking account of the complexity of MSS, the existing mature structural robustness metrics, such like maximum connected graph, node connectivity, algebraic connectivity, either have the limitation to sensitivity decline of node inactivation or edge blocking, or are difficult to distinguish the difference of the minimum degree, and also ignore the factor that there are a lot of peripheral nodes in SMN.

Natural connectivity can be used to describe the redundancy of alternative routes between vertices by quantizing the weighted sum of numbers of closed walks, and also has a clear physical meaning. As such, with redundancy as the breakthrough point, the structural robustness of MSS can be effectively characterized by natural connectivity, which can be represented by the average eigenvalue of the network graph adjacency matrix [29]:

$$\bar{\lambda} = \ln\left(\frac{1}{n}\sum_{i=1}^{n} e^{\lambda_i}\right)$$

In which n and λ_i respectively represent the node number and the characteristic roots of adjacency matrix $\mathrm{A}_{\mathrm{n*n}}$ of graph G. In order to eliminate the impact of network size on natural connectivity, we can make the normalization treatment in the following:

$$\tilde{\lambda} = \frac{\bar{\lambda}}{n - \ln n}$$

Performance robustness. The definition of performance robustness is: "*the ability to resist the decline of system performance caused by the inactivation of service nodes*". We choose endurance to describe the ability of an organism to withstand an adverse situation in order to remain active for a certain period of time, which is a time-dependent property. Compared with the traditional performance robustness metrics, it is more concerned with the frequent perturbations of low-scale network elements (nodes or edges) in practice. In order to clear the reason for choosing endurance as the performance robustness metric, we firstly give the definition of endurance ξ as follows [30]:

$$\xi(a,b) = \frac{\sum_{n=a+1}^{b} A(a,n)}{\sum_{n=1}^{b-a} n}, \qquad b > a \tag{1}$$

$$A(p,q) = \sum_{n=p}^{q-1} \frac{C(n) + C(n+1)}{2}, \qquad q > p \tag{2}$$

a, b, p, q represent the proportion of the removed network elements, and $C(n)$ is a normalized function that represents the value of service parameters when removing n% network elements. Obviously, the endurance value is normalized over the interval [0, 1]. When the selected service parameters (e.g. network efficiency) are inversely proportional to the number of the removed network elements, $\xi = 1$ means the best performance robustness, whereas $\xi = 0$ means the non-existence of robustness. In contrast, ξ=0 and 1 express a opposite meaning of the former when the service parameters (e.g. blocking rate) are proportional to the number of the removed network elements.

Theorem:

(i) if $C(n)$ is a decreasing function of n, then$\xi(a,b) \geq A(a,b)/(b-a)$;

(ii) if $C(n)$ is a increasing function of n, then $\xi(a,b) \leq A(a,b)/(b-a)$.

Proof:

(i) For $A(a,b)=\sum_{n=a}^{b-1}\frac{C(n)+C(n+1)}{2}$, and $C(n)$ is a decreasing function of n, we can obtain that:

$$\frac{A(a,a+2)}{a+2-a}=\frac{C(a)+C(a+2)}{4}+\frac{C(a+1)}{2}\leq\frac{C(a)}{2}+\frac{C(a+1)}{2}=\frac{A(a,a+1)}{a+1-a}$$

Furthermore, we can obtain the following conclusion:

$$\frac{A(a,a+1)}{1}\geq\frac{A(a,a+2)}{2}\geq\cdots\geq\frac{A(a,b)}{b-a}$$

Also for $\xi(a,b)=\frac{\sum_{n=a+1}^{b}A(a,n)}{\sum_{n=1}^{b-a}n}$, then

$$\xi(a,b)=\frac{A(a,a+1)+A(a,a+2)+\cdots+A(a,b)}{1+2+\cdots+(b-a)}\geq\frac{\frac{A(a,b)}{b-a}+\frac{2A(a,b)}{b-a}+\cdots+\frac{(b-a)A(a,b)}{b-a}}{1+2+\cdots+(b-a)}=\frac{A(a,b)}{b-a}$$

That is $\xi(a,b)\geq A(a,b)/(b-a)$, prove up.

(ii) The proof is similar to (i), so the process is omitted.

According to the theorem (i), we know that performance robustness measured with endurance is higher than that measured with service parameters. Because of the decreasing property of $C(n)$, the larger the value of ξ is, the better the performance robustness will be. In reality, the probability of a large-scale interruption of cooperation is low, while the probability of a low-proportion interruption is high. To these, MSS has a good tolerance for the fully expectation of the above frequent uncertainties, and it is more in line with reality to assess the performance robustness of MSS with endurance.

It can also be seen from the definition that endurance has a strong compatibility with some common performance robustness metrics, such like service quality, service reliability, network efficiency and blocking rate. When discussing other complex networks with different features, we can take those metrics as the QoS parameters. Furthermore, the evaluation of performance robustness should be not only in a specific instant of time, but also in a period. Therefore, the above facts make it more suitable and more objective as the performance robustness metric.

A simple example is presented below to explain how endurance can be computed. We take network efficiency as the QoS parameter and assume that $C(n)$ satisfies the changing rule in Fig. 2.

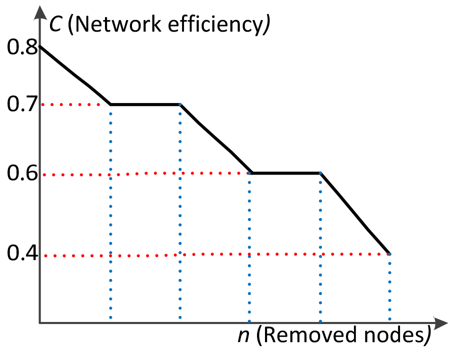


Fig. 2 The function diagram of *C* and *n*

From Fig. 2, the endurance can be calculated as depicted from the following formulas. Finally the performance robustness of such a network is of $\xi=0.68$.

$$A(0,1)=\sum_{n=0}^{0}\frac{C(n)+C(n+1)}{2}=\frac{0.8+0.7}{2}=0.75$$

$$A(0,2)=1.45, A(0,3)=2.1, A(0,4)=2.7, A(0,5)=3.2$$

$$\xi(0,5)=\frac{\sum_{n=1}^{5}A(0,n)}{\sum_{n=1}^{5}n}=\frac{A(0,1)+A(0,2)+A(0,3)+A(0,4)+A(0,5)}{1+2+3+4+5}=0.68$$

If we use network efficiency to measure the robustness, the normalized value is $A(0,5)/5 = 0.64$. This value results to be worse than the one computed by endurance because the nature of endurance is related to the fact that the network tolerates better those failures which have a lower probability of occurring. Consequently, it's further proved that taking endurance as the performance robustness metric is appropriate.

4.3 Simulation strategy and algorithm

As mentioned above, there are a lot of uncertain factors in MSS, which may cause some service principals to fail to complete their role functions, and also lead to a termination of collaboration relationships. In this paper, we divide the factors into two categories: random cooperation interruption and targeted cooperation interruption. Random cooperation interruptions mainly refer to the random uncertainties among the internal and external service principals, such like the cooperation termination caused by natural disasters or man-made accidents. Targeted cooperation interruptions generally refer to the subjective and purposeful termination, such like competitors poaching and unexpected returns. Depending on where the interruption happened, we will discuss two typical kinds of cooperation interruptions: interruption occurring in the interior of nodes and interruption occurring in the edges among nodes, which can be respectively simulated by deleting nodes and deleting edges. The specific strategies and algorithm are as follows:

Strategy 1: If random cooperation interruptions occur in the interior of MSS' nodes, then select network nodes randomly and delete these nodes.

Strategy 2: If random cooperation interruptions occur in the edges among nodes, then select network edges randomly and delete these edges.

Strategy 3: If targeted cooperation interruptions occur in the interior of MSS' nodes, then select and delete the nodes with higher node degree. (The reason for this is that node degree reflects the importance of nodes in the network. The nodes with higher node degree often have the position of "central point", which face a more external and internal inferences for its strong position in cooperation).

Strategy 4: If targeted cooperation interruptions occur in the edges among nodes, then select network edges according to the descending order of the product ($d_i d_j$) of node degree. The bigger the value of $d_i d_j$ is, the greater the probability of being attacked is, whereas d_i and d_j denote the node degree value of joint nodes.

Table 1 Simulation algorithms with strategy 1 and 3

Algorithm1	
Input: w_{ij}, n (i, j←1, 2⋯n) and the maximum percentage of removed nodes N	$\tilde{\lambda} \leftarrow \widetilde{\lambda_1}, \xi \leftarrow \xi_1$;
Output: $\tilde{\lambda}$, ξ ,the adjacency matrix W, and all nodes degree $D_{1\times n}$	**else if** strategy 3
begin{	$[r,c]$ ←find(max(D))
for k from 1 to N	**for** m from 1 to k
if strategy 1	$w(r:) \leftarrow [], w(:c) \leftarrow []$;
$B_{1\times k} \leftarrow$ randperm(N, k);	$W \leftarrow (w_{ij})_{(n-m)\times(n-m)}$
for m from 1 to k	computing $\widetilde{\lambda_1}$ and ξ_1
$w(B_m:) \leftarrow [], w(:B_m) \leftarrow []$;	$\tilde{\lambda} \leftarrow \widetilde{\lambda_1}, \xi \leftarrow \xi_1$;
end for	**end for**
computing the new value of natural connectivity $\widetilde{\lambda_1}$ and endurance ξ_1	**end else if**
	end if
	} **end** of the algorithm1

Table 2 Simulation algorithms with strategy 2 and 4

Algorithm2
1: **Input**: w_{ij}, n (i, j←1, 2⋯n) and the maximum percentage of removed edges M
2: **Output**: $\tilde{\lambda}, \xi, W, D_{1\times n}$
3: **begin**{
4: **for** k from 1 to M
5: B←find(W), $[a,b] \leftarrow$ size(B);
6: **for** m from 1 to a
7: **if** B(m,1)=B(m,2)
8: $\mathrm{B}(m:) \leftarrow []; B \leftarrow (B)_{(a-m)\times 2}$;
9: **else**
10: $B \leftarrow (B)_{a\times 2}$;
11: **end if**
12: $[r,c] \leftarrow$ size(B);
13: **end for**
14: **If** Strage2
14: $C_{1\times k} \leftarrow$ randperm(r, k);
15: **for** d from 1to k
16: $m \leftarrow C(1,d), m \leftarrow B(m,2), W(\mathrm{h},\mathrm{m}) \leftarrow 0, W(\mathrm{m},\mathrm{h}) \leftarrow 0; W \leftarrow (w_{ij})_{(n-k)\times(n-k)}$;
17: computing $\widetilde{\lambda_1}$ and ξ_1
18: $\tilde{\lambda} \leftarrow \widetilde{\lambda_1}, \xi \leftarrow \xi_1$;
19: **end for**
20: **else if** Strage 4
21 **for** i from 1 to r
22: $D1(1,i) \leftarrow D(1,B(i,1)) * D(1,B(i,2))$;
23: **end for**
24: r1 ←find(max(D1));
25: $h \leftarrow B(r1,1), m \leftarrow B(r1,2), W(\mathrm{h},m) \leftarrow 0, W(m,\mathrm{h}) \leftarrow 0; W \leftarrow (w_{ij})_{(n-k)\times(n-k)}$
26: computing $\widetilde{\lambda_1}$ and ξ_1
27: $\tilde{\lambda} \leftarrow \widetilde{\lambda_1}, \xi \leftarrow \xi_1$;
28: **end else if**
29: **end if**
30: **end for**
31: } **end** of the algorithm2

5. Case study

In the early stage, our team conducted an in-depth survey of 17 manufacturing enterprises in the Pearl River Delta region of China, mainly from luminaries, furniture, equipment, electronics and other industries. In the electronics and equipment industry, we found that the alliance with a temporary interest bargain is more common and faces a higher degree of cooperation uncertainty caused by non-standard coordination, discordant job schedule and asymmetric information. Considering the characteristics of high technology content, high added value and high degree of association, equipment manufacturing industry has a wider potential and space of service transformation, and its new value creation pattern is more typical and representative.

R is a joint venture company specializing in producing and designing elevator. Since 2012, this company has begun the pace of transformation to servitization, called"2.5 strategy", and has formed a complex and huge manufacturing service network in the last five years. Nevertheless, the cooperation interruptions resulting in a poor service quality and a high service cost often occur. Based on his own data statistics, R pays about ¥19 million per year on average for these uncertain expenditures.

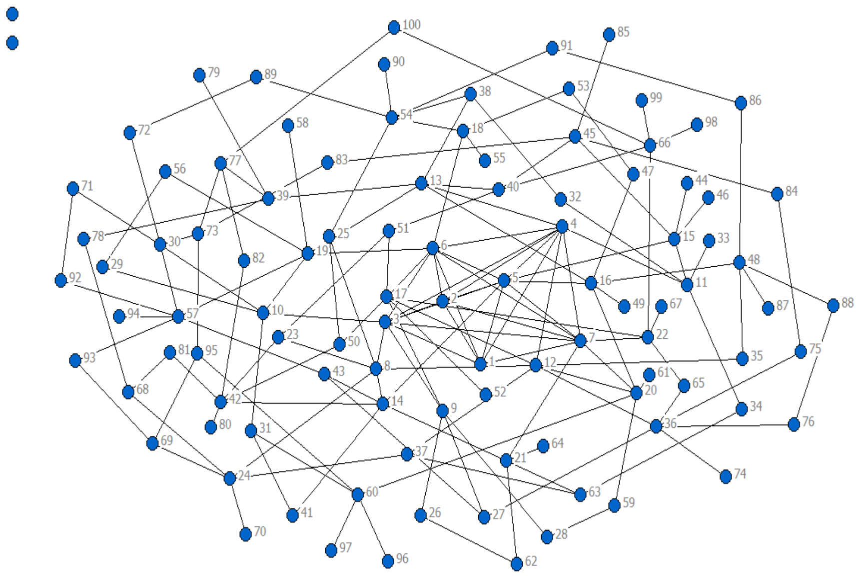


Fig. 3 The topology graph of MSN

At present, MSS has five manufacturing factories and two R&D centers in China, respectively described by nodes 1-7 in the topology graph. The others mainly refer to the primary partners and the secondary partners including service providers, outsourcing providers, suppliers and major product sellers. According to the modularized process flow and data flow of business, we can determine the collaborative relationship network among service principals, and further construct the adjacency matrix of MSN using the method described in 4.1. The following topology diagram (Fig. 3) is drawn with the NetDraw tool, which contains 100 nodes and 176 edges.

5.1 Simulation of structural robustness

In accordance with the type of cooperation interruption and the location, the 0-50 % node inactivation rate is simulated in turn. As shown in Fig. 4, the structural robustness of the initial network is 0.0245. Seen from Fig. 4, the black scatter curve fluctuates in a certain range when some nodes are attacked; however, the natural connectivity presents a downward trend with the wave band on the whole. The fluctuation reflects the randomness of node inactivation, and the upward or downward fluctuation depends on the node importance of random deletions.

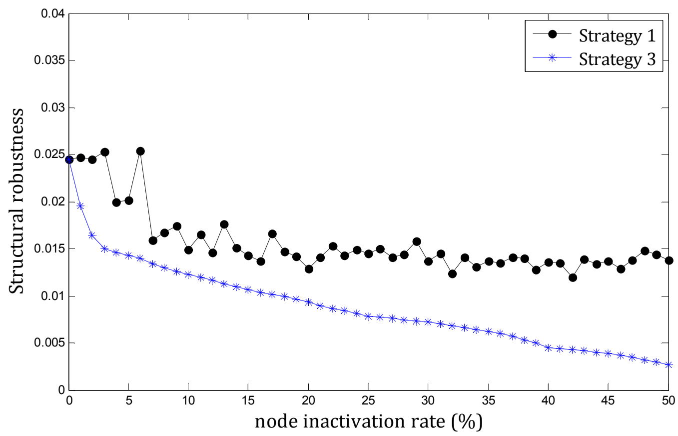


Fig. 4 Simulation results of structural robustness with strategy 1 and 3

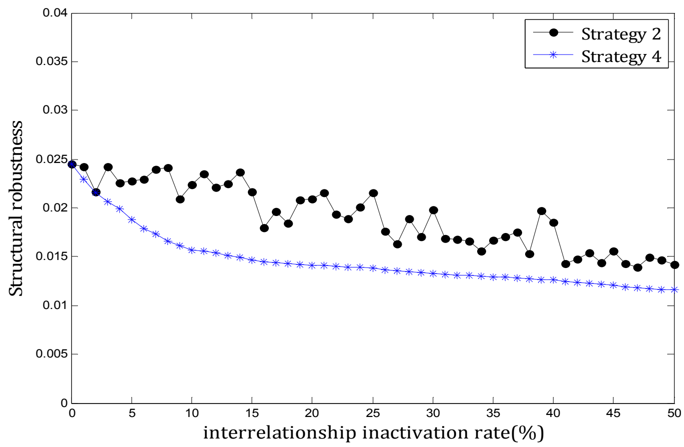


Fig. 5 Simulation results of structural robustness with strategy 2 and 4

In Strategy 1, when the node inactivation rate reached 8 %, the structural robustness had a reduction of 35.1 % than the initial state, and when the node inactivation rate reached 50 %, it only decreased by 43.67 %. These imply that MSS shows an obvious vulnerability to some core nodes, and may has a threshold for the occurrence of random cooperation interruptions in the interior of nodes, around where the structural robustness is very sensitive to node inactivation. Similarly in Strategy 3 and 4, the dramatic decline of structural robustness in the early stage with targeted cooperation interruptions also illustrates this phenomenon. What's different is that the threshold in the situation of random cooperation interruptions is different from that in the situation of targeted cooperation interruptions, which are just right for the correspondence to the two critical values of percolation theory under random failure and targeted failure. The phenomenon that the blue star curve was always below the black dot curve shows MSS has a certain resistance to random cooperation interruptions compared to targeted cooperation interruptions, and its structural robustness is slightly stronger.

5.2 Simulation of performance robustness

As you can see from Fig. 6 and Fig. 7, the performance robustness of MSS has a strong resistance to random cooperation interruptions. The reasons are chiefly as follows: one is due to the existence of a great many peripheral nodes, of which the inactivation may be beneficial to the improvement of network performance. This is consistent with many real networks, such as supply chain network and collaborative innovation network. The other is due to the tight partnerships between core manufacturers and major partners, who constitute a non-chain network structure like the relationship network formed by nodes 1-7 in Fig. 3. In particular, when the random factors lead to the interruptions within the node subsystem, MSS shows a more stable performance. The reason mainly lies in two aspects: 1) Random interruptions occur at the fringe nodes, which result in the deactivation of fringe nodes and improving the network efficiency. 2) Compared with targeted interruptions, random interruptions have little effect on the relation network consisting of core nodes with their high synergy. Although the interruptions of one or more interaction relations occur, the network efficiency shows a relatively stable trend on the whole as depicted in the black dot curve of Fig. 6.

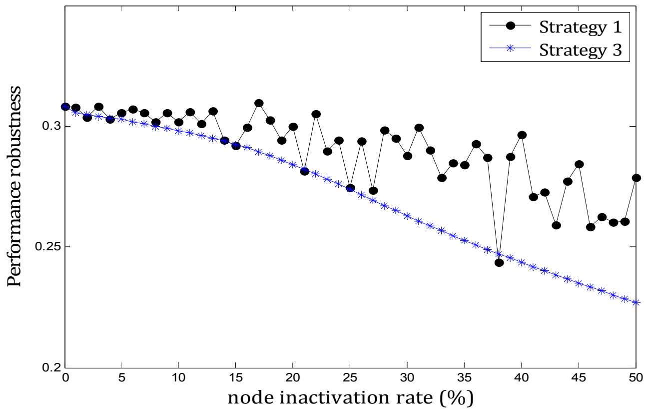


Fig. 6 Simulation results of performance robustness with strategy 1 and 3

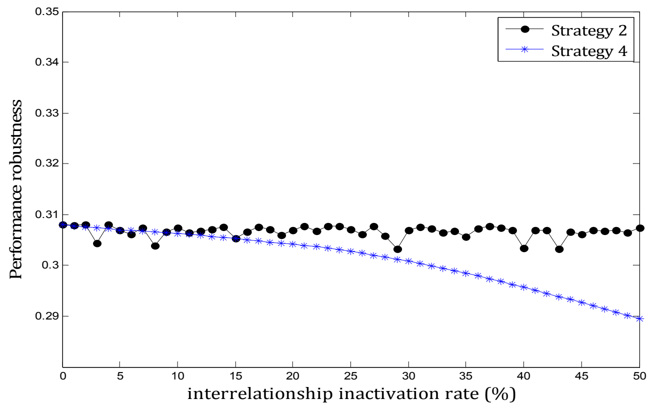


Fig. 7 Simulation results of performance robustness with strategy 2 and 4

From Fig. 6, we also find that the wave amplitude of the black dot curve is large at some time. For example, when the node inactivation rates are respectively 22 %, 39 % and 46 %, MSS' performance robustness severally had a reduction of 6.26 %, 15.18 %, 9.2 % than the previous state. This is because that the node inactivation rate of the simulation algorithm is based on the original network nodes every time instead of accumulating the node inactivation rate based on the previous network that has been attacked. Therefore, the randomness of node deletion may lead to the inactivation of some core nodes and even lead to the emergence of a sharp decline in performance robustness.

6. Conclusion

The downward fluctuation in the robustness of MSS caused by cooperation uncertainty will result in the degradation of service quality, the increase of service cost and the delay of service response time, so it's crucial for the company to guarantee the stability of internal and external cooperation relationships. Aiming at the problems that the loose coupling character of cooperation, we analyze the impact of cooperation uncertainty on the robustness of MSS. The findings show that: (1) natural connectivity and endurance can effectively measure the changes of structure and performance. (2) The structural robustness of MSS has an obvious vulnerability to the inactivation of core node subsystems, and also has a little stronger resistance to random cooperation interruption compared to targeted cooperation interruption. In addition, there should be a threshold respectively in the situation of random cooperation interruption and targeted cooperation interruption, which can be used as the judgement conditions for structure collapse. (3) The performance of MSS shows a stronger resistance to random cooperation interruptions, which verifies the importance of large scale presence of terminal enterprise, and of the synergy of core manufacturing enterprises and service providers. Accordingly, the performance robustness is more sensitive to the inactivation of nodes compared to the inactivation of interrelationships among nodes.

Therefore, the companies should first identify the critical paths of MSS, and then standardize the cooperation mechanism within and between core organizations, which can contribute the most to minimizing cooperation uncertainty. Besides, we can take the following measures to enhance the activity of network nodes, such like adjusting operation mode, strengthening informationization construction and building an emergency response mechanism. Simultaneously, managing customer relationship, establishing new collaborative relationships with those normative members, and searching for new partners with potentiality are also feasible to maintain the activity of nodes and edges of the critical paths. Compared to other measures, the governance of partnership and customer relationship can reduce the impact of cooperation uncertainty on the robustness with lower possible costs. The follow-up studies can be focused on the critical threshold in the situation of inactivation of interrelationship, robust operation optimization and service network planning.

Acknowledgement

This work is partly supported by Natural Science Foundation of China (grant NO. 71571072), Research Foundation of Education Bureau of Anhui Province (grant NO. SK2017A0317) and Natural Science Foundation for Youths of Anhui Province (grant NO.1808085QG230).

References

[1] Zhao, Y., Zhang, W., Liu, D., Bao, F., Tian, L. (2017). Service implementation in manufacturing firms: The role of service-orientated human resource management practices and demand-side search, *Management Decision*, Vol. 55, No. 4, 648-661, doi: 10.1108/MD-12-2015-0605.

[2] Huang, S., Zeng, S., Fan, Y., Huang, G.Q. (2011). Optimal service selection and composition for service-oriented manufacturing network, *International Journal of Computer Integrated Manufacturing*, Vol. 24, No. 5, 416-430, doi: 10.1080/0951192X.2010.511657.

[3] Almaktoom, A.T., Krishan, K.K., Wang, P., Alsobhi, S. (2014). Assurance of system service level robustness in complex supply chain networks, *International Journal of Advanced Manufacturing Technology*, Vol. 74, No. 1-4, 445-460, doi: 10.1007/s00170-014-5883-7.

[4] Buzacott, J.A., Peng, H.S. (2012). Contract design for risk sharing partnerships in manufacturing, *European Journal of Operational Research,* Vol. 218, No. 3, 656-666, doi: 10.1016/j.ejor.2011.12.008.

[5] Olander, H., Hurmelinna-Laukkanen, P., Blomqvist, K., Ritala, P. (2010). The dynamics of relational and contractual governance mechanisms in knowledge sharing of R&D projects, *Knowledge and Process Management*, Vol. 17, No. 4, 188-204, doi: 10.1002/kpm.356.
[6] Zhang, W., Shi, Y., Yang, M., Gu, X., Tang, R., Pan, X. (2017). Ecosystem evolution mechanism of manufacturing service system by service providers, *International Journal of Production Research*, Vol. 55, No. 12, 3542-3558, doi: 10.1080/00207543.2017.1305131.
[7] Clements, K.C., Stephens, D.W. (1995). Testing models of non-kin cooperation: Mutualism and the Prisoner's Dilemma, *Animal Behaviour,* Vol. 50, No. 2, 527-535, doi: 10.1006/anbe.1995.0267.
[8] Schmenner, R.W. (2009). Manufacturing, service and their integration: Some history and theory, *International Journal of Operations & Production Management*, Vol. 29, No. 5, 431-443, doi: 10.1108/01443570910953577.
[9] Liang, P., Sun, Y., Pan, Y. (2015). The co-evolution mechanism and stability analysis of service-oriented manufacturing enterprise system, *Journal of Industrial Engineering and Management*, Vol. 8, No. 5, 1475-1490, doi: 10.3926/jiem.1457.
[10] Smith Ring, P., van de Ven, A.H. (1992). Structuring cooperative relationships between organizations, *Strategic Management Journal*, Vol. 13, No. 7, 483-498, doi: 10.1002/smj.4250130702.
[11] Cheng, J.-H. (2011). Inter-organizational relationships and knowledge sharing in green supply chains – Moderating by relational benefits and guanxi, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 47, No. 6, 837-849, doi: 10.1016/j.tre.2010.12.008.
[12] Buston, P.M., Balshine, S. (2007). Cooperating in the face of uncertainty: A consistent framework for understanding the evolution of cooperation, *Behavioural Process*, Vol. 76, No. 2, 152-159, doi: 10.1016/j.beproc.2007.01.020.
[13] Bilgiç, T., Güllü, R. (2016). Innovation race under revenue and technology uncertainty of heterogeneous firms where the winner does not take all, *IIE Transactions*, Vol. 48, No. 6, 527-540, doi: 10.1080/0740817X.2015.1110651.
[14] Chen, Q., Deng, L.F., Wang, H.M. (2018). Optimization of multi-task job-shop scheduling based on uncertainty theory algorithm, *International Journal of Simulation Modelling*, Vol. 17, No. 3, 543-552, doi: 10.2507/IJSIMM17(3)CO14.
[15] Hua, Z., Li, S. (2008). Impacts of demand uncertainty on retailer's dominance and manufacturing-retailer supply chain cooperation, *Omega*, Vol. 36, No. 5, 697-714, doi: 10.1016/j.omega.2006.02.005.
[16] Boukas, E.K. (2006). Manufacturing system: LMI approach, *IEEE Transactions on Automatic Control*, Vol. 51, No. 6, 1014-1018, doi: 10.1109/TAC.2006.876945.
[17] Huang, X.-Y., Yan, N.-N., Guo, H.-F. (2007). An H_∞ control method of the bullwhip effect for a class of supply chain system, *International Journal of Production Research*, Vol. 45, No. 1, 207-226, doi: 10.1080/00207540600678912.
[18] Liu, G.Y., Li, Z.W., Barkaoui, K., Al-Ahmari, A.M. (2013). Robustness of deadlock control for a class of Petri nets with unreliable resources, *Information Sciences*, Vol. 235, 259-279, doi: 10.1016/j.ins.2013.01.003.
[19] Mhalla, A., Jerbi, N., Dutilleul, S.C., Craye, E., Benrejeb, M. (2013). Robust control strategies facing disturbances in manufacturing workshops with time constrains, *Computers & Industrial Engineering*, Vol. 65, No. 2, 268-276, doi: 10.1016/j.cie.2013.03.003.
[20] Chica, M., Cordón, Ó., Damas, S., Bautista, J. (2013). A robustness information and visualization model for time and space assembly line balancing under uncertain demand, *International Journal of Production Economics*, Vol. 145, No. 2, 761-772, doi: 10.1016/j.ijpe.2013.05.030.
[21] Boorla, S.M., Howard, T.J. (2016). Production monitoring system for understanding product robustness, *Advances in Production Engineering & Management*, Vol. 11, No. 3, 159-172, doi: 10.14743/apem2016.3.217.
[22] Suri, R., Otto, K. (2001). Manufacturing system robustness through integrated modeling, *Journal of Mechanical Design*, Vol. 123, No. 4, 630-636, doi: 10.1115/1.1402129.
[23] Kusumoto, S., Kitamura, A., Nakamura, M. (2012). Robust scheduling for plural lot size manufacturing, In: *Proceedings of ASME/ISCIE International Symposium on Flexible Automation*, St. Louis, Missouri, USA, 359-364, doi: 10.1115/ISFA2012-7244.
[24] Mondal, S.C., Ray, P.K., Maiti, J. (2014). Modelling robustness for manufacturing processes: A critical review, *International Journal of Production Research*, Vol. 52, No. 2, 521-538, doi: 10.1080/00207543.2013.837588.
[25] Watts, D.J., Strogatz, S.H. (1998). Collective dynamics of 'small-world' networks, *Nature*, Vol. 393, No. 6684, 440-442, doi: 10.1038/30918.
[26] Dong, J., Horvath, S. (2007). Understanding network concepts in modules, *BMC Systems Biology*, Vol. 1, No. 1, Start page: 24, doi: 10.1186/1752-0509-1-24.
[27] Neumayer, S., Modiano, E. (2010). Network reliability with geographically correlated failures, In: *Proceedings of 29th IEEE INFOCOM Conference on Information Communications*, San Diego, California, USA, 1658-1666, doi: 10.1109/INFCOM.2010.5461984.
[28] Boorla, S.M., Troldtoft, M.E., Eifler, T., Howard, T.J. (2017). Quantifying the robustness of process manufacturing concept – A medical product case study, *Advances in Production Engineering & Management*, Vol. 12, No. 2, 127-138, doi: 10.14743/apem2017.2.245.
[29] Wu, J., Mauricio, B., Tan, Y.-J., Deng, H.-Z. (2010). Natural connectivity of complex networks, *Chinese Physics Letters,* Vol. 27, No. 7, doi: 10.1088/0256-307X/27/7/078902.
[30] Manzano, M., Calle, E., Torres-Padrosa, V., Segovia, J., Harle, D. (2013). Endurance: A new robustness measure for complex networks under multiple failure scenarios, *Computer Networks*, Vol. 57, No. 17, 3641-3653, doi: 10.1016/j.comnet.2013.08.011.

APEM journal

Advances in Production Engineering & Management
Volume 14 | Number 2 | June 2019 | pp 201–212
https://doi.org/10.14743/apem2019.2.322

ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

Time-dependent and bi-objective vehicle routing problem with time windows

Zhao, P.X.[a,*], Luo, W.H.[a], Han, X.[a]

[a]School of Management, Shandong University, Jinan, Shandong, P.R. China

A B S T R A C T

The optimization of bi-objective vehicle routing problem has become a research hotspot in recent years. In this paper, a time-dependent and bi-objective vehicle routing problem with time windows (TD-BO-VRPTW) is proposed, which is a new extension of classical vehicle routing problem. Time-dependency is presented for the situation that vehicle's travel speed is affected by its departure time and the distance between two customers. The total transportation costs and time costs are two objectives optimized simultaneously through constructing a bi-objective mixed integer linear programming model. To deal with this problem, the non-dominated sorting genetic algorithm II (NSGA-II) is adopted to obtain the Pareto optimal solution set. In the numerical examples, the RC108 from Solomon's benchmark set is employed and the results in the Pareto front show the efficiency of NSGA-II for the TD-BO-VRPTW. To further test the performance of this algorithm, two objectives are optimized separately and then the sum of two objectives is also optimized. Through comparing these results with solutions in the Pareto front, it can be concluded that the algorithm is reliable, and the results in Pareto front are competitive because there is a trade-off between two objectives.

A R T I C L E   I N F O

Keywords:
Vehicle routing problem;
Time-dependency;
Bi-objective optimization;
Time windows;
Pareto optimal solutions;
Evolutionary algorithms;
NSGA-II algorithm

**Corresponding author:*
pxzhao@sdu.edu.cn
(Zhao, P.X.)

Article history:
Received 14 February 2019
Revised 10 May 2019
Accepted 17 May 2019

1. Introduction

The Vehicle Routing Problem (VRP) was first proposed by Dantzig and Ramser in 1959 [1]. It is defined on a network, in which the points represent the depot and customers and the weights on the arcs stand for the distance between two points. The depot is the start point and end point of each route, and each customer has a demand that must be served by a vehicle. Each vehicle serves the customers along its route and the sum of all demands on this route cannot exceed the maximum capacity. VRP aims to find the optimal solution and thus the objective function is optimized and the corresponding constraints are met when each vehicle travels along its own route. Zhu *et al*. proposed an improved hybrid algorithm based on heuristic to solve VRP under multi-depots condition [2]. Chiang and Hsu considered that VRP combines the characteristics of two typical NP-hard optimization problems including traveling salesman problem (TSP) and bin packing problem (BPP) [3]. Similar to VRP, TSP aims to find the optimal route, but it only considers the situation when the number of vehicles is equal to one**.** BPP assigns customers to different vehicles to minimize the number of vehicles and satisfies all the constraints at the same time, but it does not consider the customer service priorities.

The vehicle routing problem with time windows (VRPTW), which was first introduced by Solomon in 1987, is an extension of VRP [4]. In the VRPTW, each customer is assigned with a time window who must be served within the time window. Cao *et al*. solved the vehicle routing problem with multiple fuzzy time windows using an improved wolf pack algorithm [5]. Wu *et al*. proposed a co-evolutionary based algorithm and applied it to solve VRPTW [6]. Yu *et al*. proposed an improved branch-and-price (BAP) to handle the VRPTW with heterogeneous fleet [7]. Miranda and Conceição proposed an extension of VRPTW in which the travel time and service time are stochastic and solved it by metaheuristic [8]. In the present study, we regard the time window as a soft constraint. If the vehicle arrives outside the time window, no matter earlier or later, it should pay for the penalty cost. Many real-world applications can be considered as the VRPTW, such as postal delivery, train and bus scheduling as well as waste collection. VRPTW is a typical NP-hard combinatorial optimization problem. Due to its practical value and high computational complexity, it has become a research hotspot in operational research and management science.

The bi-objective vehicle routing problem (BOVRP) is another extension of VRP. The classical way to solve BOVRP is to add the two objectives together and solve it as a single-objective optimization problem. This method is essentially to search for a single-objective optimal solution. Another way is to find the Pareto optimal solution set, which is a trade-off between two objectives. In this set, solutions are not going to be worse than any other solutions on both objective function values simultaneously. Geiger was one of the first researchers solving multi-objective vehicle routing problem by using Pareto approaches [9]. Then Barán and Schaerer proposed a multi-objective ant colony algorithm to minimize the number of vehicles, travel distance and service time simultaneously [10]. Tan *et al*. put forward a hybrid multi-objective evolutionary algorithm (HMOEA) to deal with BOVRP, in which the two objectives are number of vehicles and travel distance [11]. Ombuki *et al*. designed a multi-objective genetic algorithm to minimize the number of vehicles and travel distance [12]. Garcia-Najera and Bullinaria proposed a multi-objective evolutionary algorithm (MOEA), which is characterized by considering individual similarity in the selection process [13]. Chiang and Hsu proposed a knowledge-based evolutionary algorithm to solve bi-objective vehicle routing problem with time windows, minimizing the number of vehicles and travel distance [3]. Qi *et al*. proposed a memetic algorithm based on decomposition and applied it to solve multi-objective VRPTW [14]. Iqbal *et al*. proposed a new model for multi-objective VRP, and solved it on the basis of local search [15]. Some researchers considered the minimization of time window constraint violations as an objective. For example, Xu *et al*. optimized the constraint violation and two other objectives by an Or-opt NSGA-II [16]. Castro *et al*. considered the constraint violation of vehicle capacity and time window as two objectives [17].

The time-dependent vehicle routing problem (TDVRP) was initially proposed by Malandraki and Daskin to capture the congestion in a traffic network [18]. Hill and Benton constructed a model for time-dependent conditions, which lays the foundation of many studies [19]. Malandraki and Dial solved TSP under time-dependent conditions by dynamic programming [20]. However, all of the above studies disrespect the First-In-First-Out (FIFO) principle, indicating that it is possible that a vehicle departing later arrives earlier. Ichoua *et al*. proposed a time-dependent model in which travel speed is a step function and travel time is a piecewise linear function [21]. Fleishmann *et al*. proposed a method to construct time-dependent model in which the travel time was smoothed by a step function [22]. These methods respect the FIFO principle. These two methods are widely used in studies afterwards. Then, Huang *et al*. proposed the conception of path flexibility for TDVRP and solved it by CPLEX [23]. Çimen and Soysal extended the conception of TDVRP to stochastic conditions and proposed a heuristic to solve it [24].

Based on previous research, we combine the bi-objective optimization, the time-dependent travel time model and the time window as well as innovatively propose the time-dependent and bi-objective vehicle routing problem with time windows (TD-BO-VRPTW). In the remainder of this paper, we first give a detailed description of the TD-BO-VRPTW in Section 2. In Section 3, we introduce the notion of NSGA-II. Our main results are given in Section 4. Finally, Section 5 contains a brief summary.

2. Problem description and model construction

This section mainly includes the notion of time-dependent travel time model, the definition of parameters and decision variables, assumptions and mathematical model.

2.1 Time-dependent travel time model

As mentioned previously, Malandraki and Daskin first proposed the conception of time-dependent travel time model [18]. Fig. 1 shows the relationship between travel time and departure time when the distance is 1, presenting that the travel time between two points is a step function of departure time. The main feature of this model is that it does not respect the FIFO principle. For example, a vehicle leaving at $t_1 = 1$ will arrive at $t_2 = 4$, while a vehicle leaving at $t_3 = 2$ will arrive at $t_4 = 3$. Hill and Benton proposed another time-dependent model in which travel speed is a step function of the departure time, as shown in Fig. 2. This model still disrespects the FIFO principle [19]. In order to give a better description of time-dependency, adjustment of travel speed should be taken into account when vehicle travels beyond the bounds of the time intervals. Ichoua *et al.* proposed a time-dependent travel time model in which travel speed is a step function and travel time is a piecewise linear function, as shown in Fig. 3 and Fig. 4 [20].

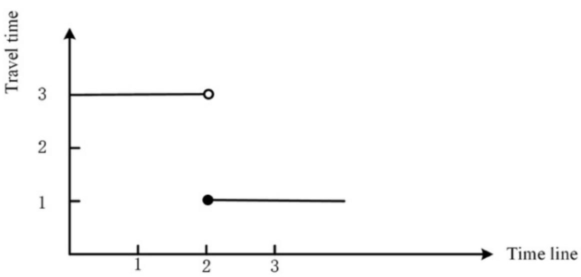


Fig. 1 Time-dependent model I

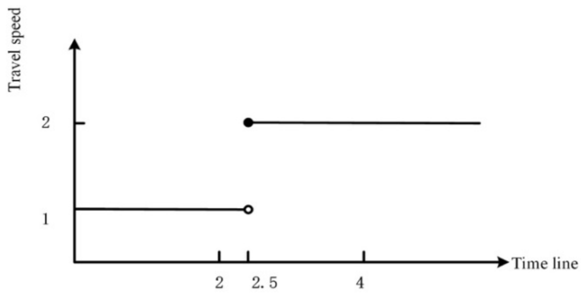


Fig. 2 Time-dependent model II

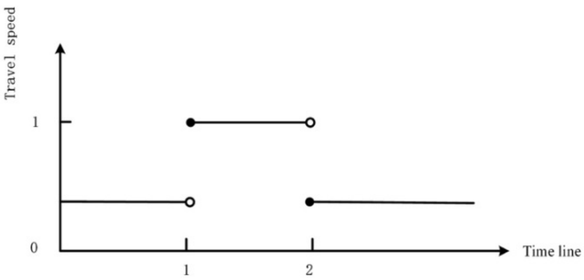


Fig. 3 Time-dependent model III: Speed

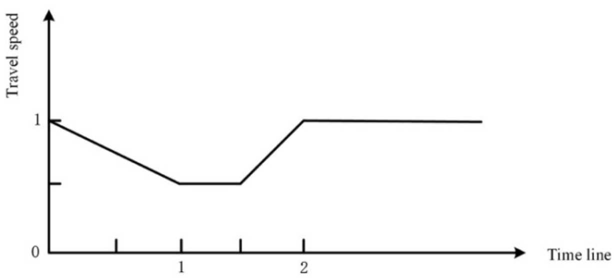


Fig. 4 Time-dependent model III: Time

Fig. 3 shows the relationship between travel speed and departure time, and Fig. 4. provides an example of travel time function for an arc of length 1. For a non-time-dependent and symmetrical road model, travel time of an arc (v_i, v_j) can be simply expressed as $t_{ij} = d_{ij}/v_{ij}$, where v_{ij} is the average travel speed on arc (v_i, v_j) and d_{ij} is the distance of arc (v_i, v_j). However, in time-dependent model, travel speed changes with departure time. Suppose the departure time is t_0, and travel speed is a step function of departure time, wherein speed of the k-th time period is v_{ij}^k, corresponding to the departure time interval for $[t_k, t_{k+1}]$. Then, the process of calculation is expressed as follows:

Step 1: Calculate the value of k according to $t_0 \in [t_k, t_{k+1}]$, and get the corresponding v_{ij}^k, assign $d_{ij} \to d$, calculate $t' = t_0 + d/v_{ij}^k$.

Step 2: If $t' \leq t_{k+1}$, then output arrival time t' and travel time $t' - t_0$, end the loop. Otherwise, go to Step3.

Step 3: Update $d_{ij} - v_{ij}^k(t_{k+1} - t_0) \rightarrow d$, calculate $t' = t_{k+1} + d/v_{ij}^{k+1}$,update $k + 1 \rightarrow k$, return to Step 2.

This time-dependent model respects the FIFO principle. Compared with the former two time-dependent models, it can better capture congestion in a traffic network. Therefore, we choose this model for further study.

2.2 Assumptions and symbolic descriptions

The following assumptions will be needed throughout the paper:

- Travel speed is variable, and there are multiple time intervals. Speed in a certain time interval is constant.
- Each customer is assigned with a time window and vehicles arriving in advance or beyond the deadline must pay for penalty cost.
- Each customer must be served exactly once by one vehicle.
- Each vehicle must depart from the depot and return to the depot after completing the delivery task.
- Vehicles are homogeneous, and the constraints of maximum driving distance and maximum driving time are not considered.

Based on the above assumptions, the present study aims to minimize transportation costs and time costs. Transportation costs include depreciation costs and fuel costs, and time costs include waiting costs and penalty costs. Since the problem involved in the current work is essentially a derivative model of classical VRPTW, it can be described by the traditional network fluid formula system as follows [25]:

Let $G = (V, D)$ represents a complete graph, where V= $\{v_0,v_1,...,v_{n+1}\}$ is the point set. Therefore, the subset C=$\{v_1,v_2,...,v_{n+1}\}$ are the customers. Then, define the node sets V_0=$\{v_0,v_1,...,v_n\}$ and V_{n+1}=$\{v_1,v_2,...,v_{n+1}\}$, and arc set D=$\{(v_i, v_j)|\ (v_i, v_j)|v_i \in V_0,v_j \in V_{n+1},i \neq j\}$. Each vehicle $k \in K$ has a maximum capacity Q. Besides, each customer $i \in C$ has a demand $q_i \geq 0$, a service time $s_i \geq 0$ and a time window $[e_i,l_i]$. The depot also has a time window $[0, T_{max}]$. The distance of arc $(v_i, v_j) \in D$ is d_{ij}. c^f is the vehicle's fuel cost per unit distance, and c^d is the vehicle's depreciation cost. p_e is the waiting cost per unit time, and p_l is the penalty cost per unit time.

2.3 Mathematical model

First, we define two decision variables:

- Binary decision variable $x_{ij}^k, \forall k \in K, \forall i \in V_0, \forall j \in V_{n+1}$.

$$x_{ij}^k = \begin{cases} 1 \text{ If vehicle } k \text{ travels from customer } i \text{ to customer } j \\ \\ 0 \text{ Otherwise} \end{cases}$$

- Non-negative real decision variable y_i^k: The departure time after vehicle k has served customer i.

According to the previous time-dependent model, travel speed on arc (v_i, v_j) is a function of the departure time $v_{ij}(y_i)$ and then travel time can be denoted by d_{ij} and $v_{ij}(y_i)$, i.e. $t_{ij}(d_{ij}, v_{ij}(y_i))$. Travel speed function $v_{ij}(y_i)$ and travel time function $t_{ij}(d_{ij}, v_{ij}(y_i))$ are a step function and a piecewise linear function, respectively. Based on this notation, the TD-BO-VRPTW can be constructed as the following bi-objective mixed integer linear programming model:
Objective functions:

$$minf_1 = \sum_i \sum_j \sum_k c^f x_{ij}^k d_{ij} + c^d \sum_k \sum_j x_{0j}^k \tag{1}$$

$$minf_2 = p_e \sum_i max(e_i - y_i^k + s_i, 0) + p_l \sum_i max(\ y_i^k - s_i - l_i, 0) \tag{2}$$

Constrains:

$$\sum_{i\in C}\sum_{j\in V_{n+1}} q_i x_{ij}^k \leq Q, \forall k \in K \tag{3}$$

$$\sum_{k\in K}\sum_{j\in V_{n+1}} x_{ij}^k = 1, \forall\, i \in C \tag{4}$$

$$\sum_{i\in V_0} x_{ih}^k - \sum_{j\in V_{n+1}} x_{hj}^k = 0, \forall h \in C, \forall k \in K \tag{5}$$

$$x_{i0}^k = x_{n+1,i}^k = 0, \forall i \in C, \forall k \in K \tag{6}$$

$$\sum_{i\in C} x_{0i}^k = \sum_{i\in C} x_{i,n+1}^k \leq 1, \forall k \in K \tag{7}$$

$$\left(y_i^k + t_{ij}\left(d_{ij}, v_{ij}\left(y_i^k\right)\right)\right) x_{ij}^k \leq y_j^k - s_j, \forall i \in V_0, \forall j \in V_{n+1} \forall k \in K \tag{8}$$

$$x_{ij}^k \in \{0,1\}, \forall i \in V_0, \forall j \in V_{n+1} \forall k \in K \tag{9}$$

$$y_i^k \geq 0, \forall i \in V, \forall k \in K \tag{10}$$

The objective function (Eq. 1) minimizes the transportation costs, and the objective function (Eq. 2) minimizes the time costs. Expressions (Eq. 3) are the maximum capacity constraints, indicating that for each route, the sum of all demands cannot exceed the maximum capacity. Constraints (Eq. 4) ensure that each customer is visited exactly once by one vehicle. Constraints (Eq. 5) and (Eq. 6) are the vehicle flow conservation constraints. Constraints (Eq. 7) stipulate that each vehicle can be used at most once. Expressions (Eq. 8) are the arrival time constraint, indicating that a vehicle must arrive at the next customer earlier than its service start time. Expressions (Eq. 9) and (Eq. 10) are variable specification constraints.

3. Used methods

Recently, the non-dominated sorting genetic algorithm II (NSGA-II) proposed by Deb has been extensively responded and applied [26]. The selection process of individuals is improved by employing the elite strategy, density value estimation strategy and fast non-dominated sorting strategy. This genetic algorithm has been widely used in various fields. It not only ensures the value of the optimal solution and obtains the Pareto front, but also reduces the algorithm complexity and ensures the efficiency of the algorithm. It has become a classical algorithm in the field of multi-objective optimization. Therefore, this algorithm is adopted to solve the problem. This section mainly introduces integer linear programming, NSGA-II and its related conceptions.

3.1 Introduction to integer linear programming and NSGA-II

Integer linear programming indicates that the values of some or all decision variables in linear programming are integers. Integer linear programming is mostly NP-hard. Therefore, when the scale of the problem is large, it will be difficult to solve the problem. Integer linear programming is an important issue in operations research. In recent decades, numerous scholars have studied the solution to solve this problem. The most commonly used method is the heuristic algorithm, whose principle is to find a feasible solution within an acceptable range. The advantage is that the convergence speed is fast, and the disadvantage is that it may fall into local optimum. Common heuristic algorithms include genetic algorithm, tabu search, simulated annealing, etc.

The NSGA-II algorithm was proposed based on genetic algorithm. The non-dominated sorting approach is put forward in order to give every solution in the population a non-domination rank.

The non-domination rank starts at 1 and increases by 1 each round. In each round of the loop, every unsorted individual p in the population is compared with all the remaining unsorted individuals to determine whether the individual p dominates all the other unsorted individuals. If so, individual p is assigned with the current non-domination rank. Finally, all individuals in the population are assigned with a non-domination rank.

Additionally, crowding distance is also defined in NSGA-II to get an estimate of the density of solutions surrounding a particular solution. In terms of bi-objective optimization, calculate the average distance of two points on either side of a particular point along each objective. This value can be regarded as the sum of the adjacent two sides of a rectangle formed by two adjacent solutions on both sides of a particular solution as vertices. As presented in Fig. 5, the crowding distance of the th i-th solution is equal to half the perimeter of the rectangle formed by its two adjacent solutions as vertices, as shown by dashed line.

Each solution is assigned with a crowding distance by calculation. For individuals with the same non-dominated rank, the larger the crowding distance, the better the diversity. Therefore, it should be preferred in the selection process. Based on the definition of non-domination rank i_{rank} and crowding distance $i_{distance}$, two solutions i and j in the population can be compared using the comparison operator:

$i \prec j$
if: $i_{rank} < j_{rank}$
or: $i_{rank} = j_{rank}$ and $i_{distance} > j_{distance}$

That is, in the selection process of the algorithm, it is preferred when a solution has a lower non-domination rank. Otherwise when the two solutions have the same non-dominated rank, we prefer the solution with larger crowding distance due to its better diversity.

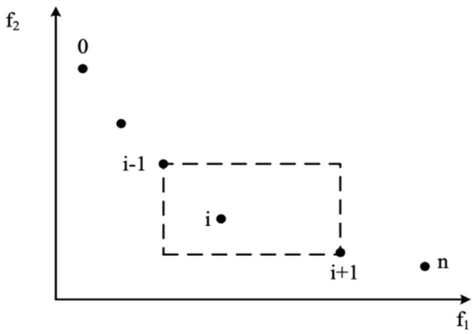


Fig. 5 Crowding distance

3.2 Main loop

The main loop of NSGA-II can be found in Fig. 6.

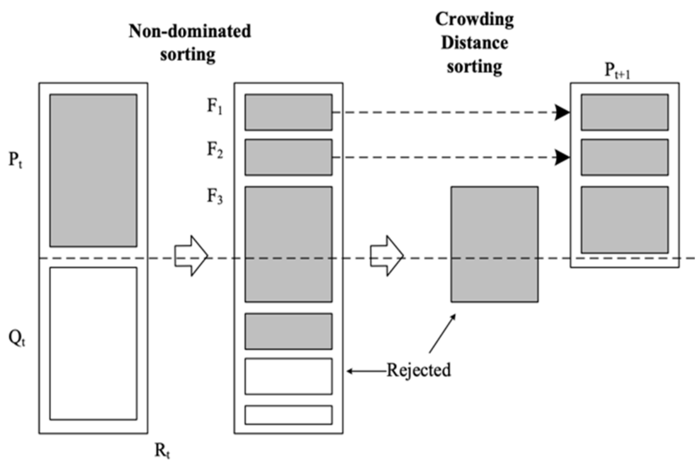


Fig. 6 Main loop

First, generate the initial population P_0, which contains N solutions. Then, the population is sorted based on domination, and each solution is assigned with a non-domination rank. An offspring population Q_0 of size N is created by selection, crossover and mutation operators. These two populations are combined to get $R_0 = P_0 \cup Q_0$. Finally, the population R_0 is trimmed based on non-domination rank and crowding distance to ensure that the population is the same size at the beginning of each iteration.

4. Results and discussion

Solomon's benchmark set is widely used in VRP, which consists of 56 instances categorized into 6 sets. Each instance includes 1 depot and 100 customers. According to the spatial distribution of customers, these instances can be divided into 3 categories including C (clustered), R (random) and RC (mixed). It is assumed that the travel speed of all arcs conforms to a unified time-dependent step function. The time window of depot $[0, T_{max}]$ ($T_{max} = 240$) is divided into five intervals on average, and the average speed of each interval is [1, 1.6, 1.05, 1.6, 1]. In this paper, RC108 is selected to test the performance of the algorithm. In this instance, customers are clustered and randomly distributed, the maximum capacity of vehicle Q is 200, and the service time s_i is 10.

All the experiments are based on MATLAB R2017a, and the algorithm parameters are set as follows: Population size $Popsize = 100$, maximum number of iterations $IterMax = 500$, crossover probability $p_c = 0.5$ and mutation probability $p_m = 0.1$. Vehicle's fuel cost per unit tance $c^f = 0.5$, depreciation cost $c^d = 50$, the time cost coefficient is $p_e = 0.5$ and $p_l = 5$. Fig. 7 represents the first non-dominated front, namely Pareto front. There are 27 solutions in the Pareto optimal set. Table 1 shows the detailed data.

The results in Fig. 7 show that these two objectives f_1 and f_2 are not positively correlated. If the two objectives change in the same direction, some solutions will be better on both objectives and thus the number of Pareto solutions becomes small. There are 27 solutions in Pareto front, indicating that the two objectives are contradictory to some extent. For one objective to become better, it is necessary to take the other to become worse.

To further compare the performance of the algorithm, we use single-objective optimization algorithm to optimize f_1 and f_2 respectively. We use the classical genetic algorithm to carry out this experiment. The relevant parameters are set as before. First, f_1 is optimized and the final optimization result is shown in Fig. 8. The left figure represents the optimal routes, where the square and circles represent the depot and customers respectively, and the right figure stands for the change of the objective function f_1 with the number of iterations. Fig. 9 is obtained by optimizing f_2. The detailed route of each vehicle and the values of f_1 and f_2 are shown in Table 2 and Table 3. Finally, f_1 and f_2 are added together as one objective and optimized. The results are shown in Table 4 and Fig. 10.

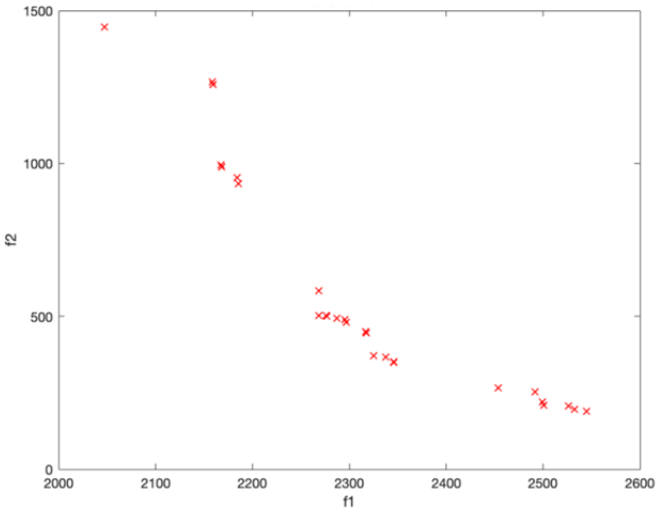


Fig. 7 Pareto front

Table 1 Detailed data of Pareto front

Solution No.	Values of Pareto optimal solutions	
	f_1	f_2
1	2047.2581	1445.5884
2	2158.6574	1266.6068
3	2159.2772	1258.6648
4	2167.7613	996.1304
5	2168.3811	988.1884
6	2184.0555	954.5300
7	2185.5761	934.0495
8	2268.4119	583.7399
9	2268.8158	502.9485
10	2276.2644	502.6524
11	2276.8878	499.6613
12	2287.2728	494.8850
13	2295.2586	489.3503
14	2296.7619	480.1851
15	2316.6997	450.1220
16	2317.8669	444.9339
17	2325.2467	370.6427
18	2337.3527	367.2576
19	2345.8875	352.2281
20	2345.8913	348.4370
21	2453.7982	265.2944
22	2492.1418	253.8482
23	2499.1596	220.4622
24	2500.9599	208.9021
25	2526.6566	207.1993
26	2532.5440	195.3103
27	2545.1208	190.0310

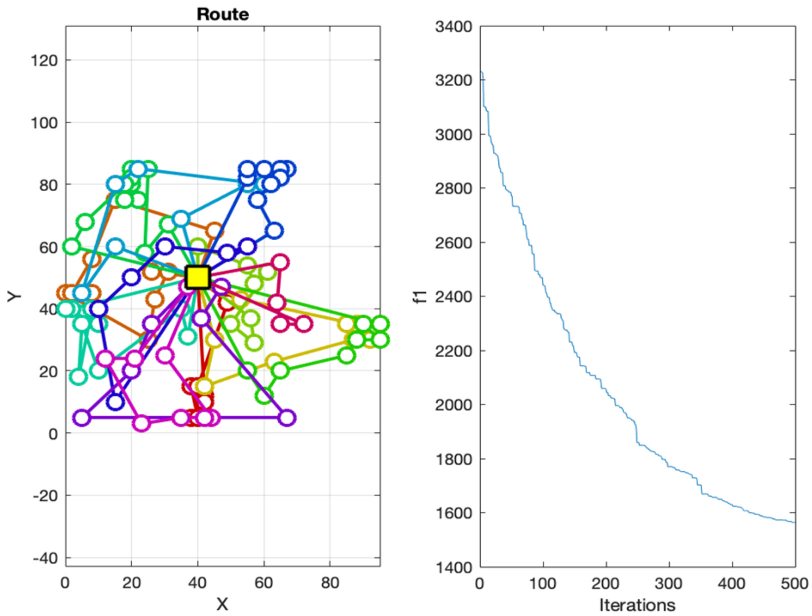


Fig. 8 Results when optimizing f_1

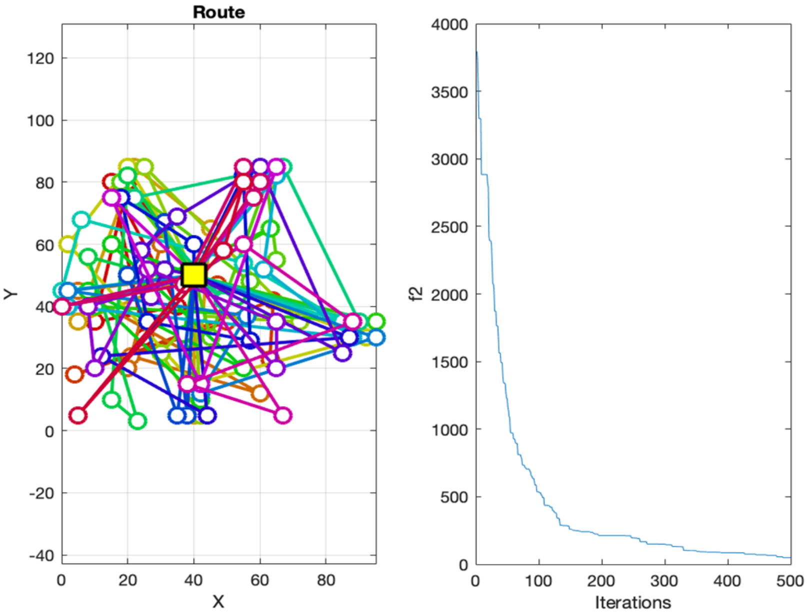


Fig. 9 Results when optimizing f_2

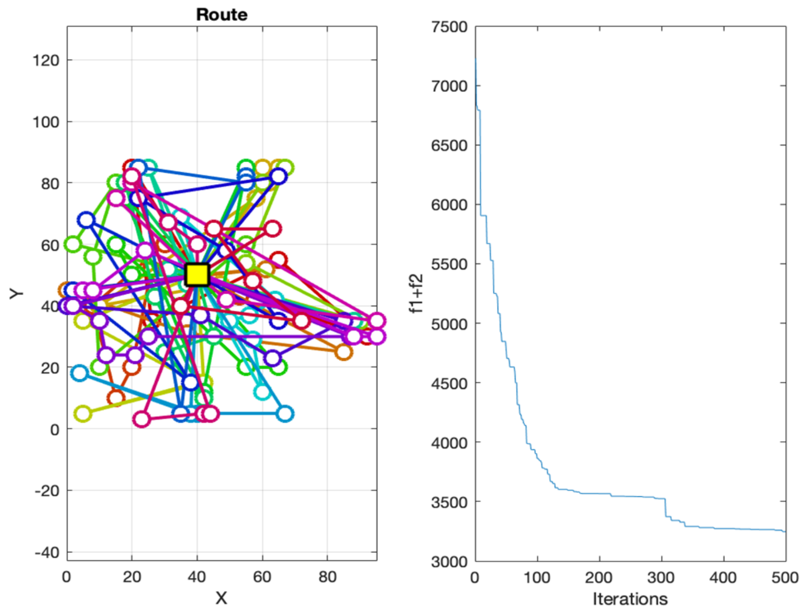


Fig. 10 Results when optimizing $f_1 + f_2$

Table 2 Detailed data of optimal solution when optimizing f_1

Vehicle No.	Optimal routes and the value of f_1 and f_2 when optimizing f_1
1	0-49-19-23-21-24-22-91-0
2	0-98-69-82-52-17-12-47-78-7-61-0
3	0-92-31-34-28-32-85-20-64-0
4	0-68-96-94-93-56-84-95-0
5	0-51-76-63-33-30-26-29-27-0
6	0-100-88-1-5-4-45-2-6-46-79-73-0
7	0-15-16-9-11-97-13-59-65-83-0
8	0-60-14-8-3-39-42-70-0
9	0-44-43-40-35-36-37-38-41-72-0
10	0-58-10-53-55-81-54-0
11	0-99-74-75-89-66-80-0
12	0-57-18-48-25-77-87-86-90-0
13	0-50-62-67-71-0
f_1	1564.0496
f_2	3909.1340

Table 3 Detailed data of optimal solution when optimizing f_2

Vehicle No.	Optimal routes and the value of f_1 and f_2 when optimizing f_2
1	0-52-8-65-9-4-0
2	0-85-67-97-56-80-0
3	0-14-76-86-74-55-0
4	0-21-61-3-13-0
5	0-5-73-22-28-0
6	0-48-57-66-50-1-0
7	0-64-38-71-94-46-0
8	0-27-10-72-0
9	0-83-19-12-51-60-0
10	0-92-78-77-58-45-0
11	0-29-35-2-91-0
12	0-17-79-96-34-0
13	0-15-30-93-37-0
14	0-47-95-49-26-0
15	0-23-100-53-25-0
16	0-44-84-99-6-68-0
17	0-18-87-32-0
18	0-69-82-88-70-40-33-0
19	0-11-59-98-62-63-0
20	0-7-20-36-0
21	0-89-24-31-54-0
22	0-16-90-43-41-39-0
23	0-75-81-42-0
f_1	3572.6748
f_2	48.6028

Table 4 Detailed data of optimal solution when optimizing $f_1 + f_2$

Vehicle No.	Optimal routes and the value of f_1 and f_2 when optimizing $f_1 + f_2$
1	0-28-71-99-5-0
2	0-92-98-74-58-11-55-0
3	0-33-83-17-93-0
4	0-36-38-40-41-10-80-0
5	0-13-22-75-20-0
6	0-39-35-96-29-0
7	0-73-8-78-59-54-0
8	0-53-60-51-63-0
9	0-6-49-43-0
10	0-64-57-19-82-46-0
11	0-31-69-90-1-0
12	0-76-84-56-95-67-70-0
13	0-89-21-97-23-0
14	0-44-42-3-25-0
15	0-79-24-47-0
16	0-62-81-2-37-0
17	0-16-66-85-34-0
18	0-15-9-87-86-52-32-0
19	0-14-12-88-30-26-0
20	0-7-4-27-91-0
21	0-77-48-18-45-100-68-0
22	0-72-61-94-50-65-0
f_1	3115.7205
f_2	132.0221

According to Table 1, Table 2 and Table 3, when f_1 is optimized, the optimal solution $f_1 = 1564.0496$ and the corresponding $f_2 = 3909.1340$; when f_2 is optimized, the result is $f_1 = 3572.6748$ and $f_2 = 48.6028$. Obviously, when optimization is carried out with one objective, the other objective will be significantly worse, which is consistent with the previous analysis that f_1 and f_2 are contradictory. For example, when f_1 is optimized, the value of optimal solution is 1564.0496, which is better than all the solutions in Pareto front. However, the corresponding f_2 is 3909.1340, which is worse than the worst in Pareto front, i.e. 1445.5884. The same goes for the situation when f_2 is optimized. From Fig. 8 and Fig. 9, we can also find that when f_1 is optimized, the routes with shorter distance is preferred. When f_2 is optimized, the routes with shorter interval between time windows of customers are preferred. Above all, it can be concluded that not as extreme as the single-objective optimal solution, bi-objective optimization is a trade-off between two objectives.

Additionally, based on Table 1 and Table 4, when we regard f_1 and f_2 as one objective, the value of optimal solution is $f_1 = 3115.7205$ and $f_2 = 132.0221$, which is close to the No. 1 solution in Pareto front. Moreover, 132.0221 is better than any other solutions in Pareto front while 3115.7205 is worse than the all. This is because the impact of f_1 and f_2 on the total cost is not the same. When we add the two objectives together, the algorithm will give priority to optimizing the objective that has a greater impact on total cost, namely f_2. We can also see that no matter the objective is f_1 or f_2 or $f_1 + f_2$, the optimal solution cannot dominate any solutions in the Pareto front, indicating that the optimal solutions obtained by NSGA-II is competitive. From the above analysis, we can conclude that the algorithm adopted in this paper is effective, and the results obtained by Pareto approach are credible.

5. Conclusion

The VRP has become a classical optimization problem in recent decades due to its high computational complexity and practical value. In the present study, a time-dependent and bi-objective vehicle routing problem with time windows (TD-BO-VRPTW) is proposed based on the existing research, and a bi-objective mathematical model is formulated. The calculation process of travel time under time-dependent conditions and the main loop of bi-objective optimization algorithm are given. In the numerical experiment, an instance in Solomon benchmark set is selected to test the performance of our algorithm. The objectives are to minimize transportation costs and time costs. Meanwhile, two objectives are optimized separately and are added together for optimization for making the comparison. It is shown that the results of bi-objective optimization are competitive.

Further research is needed. For example, there are many complex factors that need to be considered, such as nonhomogeneous fleet and multiple depots. The definition of time-dependency can also be extended to random or regional time-dependency, etc. In addition, related studies have demonstrated that NSGA-II is less efficient when the number of objective increases and there are some difficulties in balancing convergence and diversity preservation. Future research can test the algorithm performance of different multi-objective optimization algorithms. Moreover, in the case of an emergency, there may be path failure and reconnection or other effects, which are also the focus of future research.

References

[1] Dantzig, G.B., Ramser, J.H. (1959). The truck dispatching problem, *Management Science*, Vol. 6, No. 1, 80-91, doi: 10.1287/mnsc.6.1.80.

[2] Zhu, X.N., Yan, R., Zhang, Q. (2015). A promoted hybrid heuristic algorithm for two-dimensional multi-depots vehicle routing problem, *International Journal of Simulation Modelling,* Vol. 14, No. 3, 499-510, doi: 10.2507/IJSIMM14(3)CO11.

[3] Chiang, T.-C., Hsu, W.-H. (2014). A knowledge-based evolutionary algorithm for the multiobjective vehicle routing problem with time windows, *Computers & Operations Research*, Vol. 45, 25-37, doi: 10.1016/j.cor.2013.11.014.

[4] Solomon, M.M. (1987). Algorithm for the vehicle routing and scheduling problems with time window constraints, *Operations Research*, Vol. 35, No. 2, 254-265, doi: 10.1287/opre.35.2.254.

[5] Cao, Q.K., Yang, K.W., Ren, X.Y. (2017). Vehicle routing optimization with multiple fuzzy time windows based on improved wolf pack algorithm, *Advances in Production Engineering & Management*, Vol. 12, No. 4, 401-411, doi: 10.14743/apem2017.4.267.

[6] Wu, D.Q., Dong, M., Li, H.Y., Li, F. (2016). Vehicle routing problem with time windows using multi-objective co-evolutionary approach, *International Journal of Simulation Modelling*, Vol. 15, No. 4, 742-753, doi: 10.2507/IJSIMM15(4)CO19.

[7] Yu, Y., Wang, S., Wang, J., Huang, M. (2019). A branch-and-price algorithm for the heterogeneous fleet green vehicle routing problem with time windows, *Transportation Research Part B: Methodological*, Vol. 122, 511-527, doi: 10.1016/j.trb.2019.03.009.

[8] Miranda, D.M., Conceição, S.V. (2016). The vehicle routing problem with hard time windows and stochastic travel and service time, *Expert Systems with Applications*, Vol. 64, 104-116, doi: 10.1016/j.eswa.2016.07.022.
[9] Geiger, M.J. (2008). Genetic algorithms for multiple objective vehicle routing, arXiv:0809.0416, from *https://arxiv.org/abs/0809.0416*, accessed February 1, 2019.
[10] Barán, B., Schaerer, M. (2003). A multiobjective ant colony system for vehicle routing problem with time windows, In: *Proceedings of the 21st IASTED International Conference on Applied Informatics*, Innsbruck, Austria, 97-102.
[11] Tan, K.C., Chew, Y.H., Lee, L.H. (2006). A hybrid multiobjective evolutionary algorithm for solving vehicle routing problem with time windows, *Computational Optimization and Applications*, Vol. 34, 115-151, doi: 10.1007/s10589-005-3070-3.
[12] Ombuki, B., Ross, B.J., Hanshar, F. (2006). Multi-objective genetic algorithms for vehicle routing problem with time windows, *Applied Intelligence*, Vol. 24, No. 1, 17-30, doi: 10.1007/s10489-006-6926-z.
[13] Garcia-Najera, A., Bullinaria, J.A. (2011). An improved multi-objective evolutionary algorithm for the vehicle routing problem with time windows, *Computers & Operations Research*, Vol. 38, No. 1, 287-300, doi: 10.1016/j.cor.2010.05.004.
[14] Qi, Y., Hou, Z., Li, H., Huang, J., Li, X. (2015). A decomposition based memetic algorithm for multi-objective vehicle routing problem with time windows, *Computers & Operations Research,* Vol. 62, 61-77, doi: 10.1016/j.cor.2015.04.009.
[15] Iqbal, S., Kaykobad, M., Rahman, M.S. (2015). Solving the multi-objective vehicle routing problem with soft time windows with the help of bees, *Swarm and Evolutionary Computation,* Vol. 24, 50-64, doi: 10.1016/j.swevo.2015.06.001.
[16] Xu, H., Fan, W., Wei, T., Yu, L. (2008). An Or-opt NSGA-II algorithm for multi-objective vehicle routing problem with time windows, In: *Proceedings of 2008 IEEE International Conference on Automation Science and Engineering,* Arlington, USA, 310-314, doi: 10.1109/coase.2008.4626505.
[17] Castro, J.P., Landa-Silva, D., Pérez, J.A.M. (2009). Exploring feasible and infeasible regions in the vehicle routing problem with time windows using a multi-objective particle swarm optimization approach, In: Krasnogor, N., Melián-Batista, M.B., Pérez, J.A.M., Moreno-Vega, J.M., Pelta, D.A. (eds.), *Nature inspired cooperative strategies for optimization, (NICSO 2008), Studies in Computational Intelligence,* Vol. 236, Springer, Berlin, Heidelberg, Germany, 103-114, doi: 10.1007/978-3-642-03211-0_9.
[18] Malandraki, C., Daskin, M.S. (1992). Time dependent vehicle routing problems: Formulations, properties and heuristic algorithms, *Transportation Science*, Vol. 26, No. 3, 185-200, doi: 10.1287/trsc.26.3.185.
[19] Hill, A.V., Benton, W.C. (1992). Modelling intra-city time-dependent travel speeds for vehicle scheduling problems, *Journal of the Operational Research Society*, Vol. 43, No. 4, 343-351, doi: 10.1038/sj/jors/0430406.
[20] Malandraki, C., Dial, R.B. (1996). A restricted dynamic programming heuristic algorithm for the time dependent traveling salesman problem, *European Journal of Operational Research*, Vol. 90, No. 1, 45-55, doi: 10.1016/0377-2217(94)00299-1.
[21] Ichoua, S., Gendreau, M., Potvin, J.-Y. (2003). Vehicle dispatching with time-dependent travel times, *European Journal of Operational Research*, Vol. 144, No. 2, 379-396, doi: 10.1016/s0377-2217(02)00147-9.
[22] Fleischmann, B., Gietz, M., Gnutzmann, S. (2004). Time-varying travel times in vehicle routing, *Transportation Science*, Vol. 38, No. 2, 160-173, doi: 10.1287/trsc.1030.0062.
[23] Huang, Y., Zhao, L., Van Woensel, T., Gross, J.-P. (2017). Time-dependent vehicle routing problem with path flexibility, *Transportation Research Part B: Methodological*, Vol. 95, 169-195, doi: 10.1016/j.trb.2016.10.013.
[24] Çimen, M., Soysal, M. (2017). Time-dependent green vehicle routing problem with stochastic vehicle speeds: An approximate dynamic programming algorithm, *Transportation Research Part D: Transport and Environment*, Vol. 54, 82-98, doi: 10.1016/j.trd.2017.04.016.
[25] Reil, S., Bortfeldt, A., Mönch, L. (2018). Heuristics for vehicle routing problems with backhauls, time windows, and 3D loading constraints, *European Journal of Operational Research*, Vol. 266, No. 3, 877-894, doi: 10.1016/j.ejor.2017.10.029.
[26] Deb, K., Pratap, A., Agarwal, S., Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE Transactions on Evolutionary Computation*, Vol. 6, No. 2, 182-197, doi: 10.1109/4235.996017.

APEM journal
Advances in Production Engineering & Management
Volume 14 | Number 2 | June 2019 | pp 213–224
https://doi.org/10.14743/apem2019.2.323
ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

Effect of purchasing and marketing integration on new product development speed: The moderating role of environmental dynamism

González-Zapatero, C.[a,*], González-Benito, J.[a], Lannelongue, G.[a]

[a]Universidad de Salamanca, IME Instituto Multidisciplinar de Empresa, Campus Miguel de Unamuno, Salamanca, Spain

A B S T R A C T

The increasing relevance of purchasing-marketing functional integration (PMFI) has drawn scholars' attention in recent years. However, more empirical research is still needed that adopts a contingent approach and studies the differentiated role each of these two functions plays in PMFI. Based on Information Processing Theory, the two flows of information that PMFI requires, from marketing to purchasing and vice versa, are used as a PMFI proxy. The study posits a positive impact of these two information flows on a typical NPD performance indicator, namely, its speed, and a positive moderation of environmental dynamism on that effect. Data from 141 Spanish firms are used to conduct a moderated multiple regression analysis to test these effects, showing that the marketing information impact of NPD speed is positive regardless of the level of environmental dynamism. However, the effect of purchasing information on NPD speed is positive when the rate of environmental dynamism is medium or high, but negative when it is low. These results will help managers to assess when each one of these flows should be promoted. Above all, they stress the need to control for possible asymmetries in the role the different functions play in functional integration.

A R T I C L E I N F O

Keywords:
Product development;
Innovation speed;
Purchasing and marketing integration;
Environmental dynamism;
Information processing theory

**Corresponding author:*
carmengz@usal.es
(González-Zapatero, C.)

Article history:
Received 22 June 2018
Revised 24 April 2019
Accepted 9 May 2019

1. Introduction

Higher competition in today's markets has drawn attention to the need firms have to be market-oriented [1]. This need to integrate the marketing function with the rest of the organisation has also been extrapolated to the particular context of new product development (NPD). Sundry studies have analysed marketing integration with different functions during NPD, such as R&D [2], or manufacturing [3].

However, the benefits of integrating the marketing function with the purchasing function in particular have only recently received significant attention [4-9]. The reason for this delay may well be that purchasing has traditionally been considered a merely administrative function, subject to the dictates of other functional areas [9, 10]. Nevertheless, different factors have revealed that not only is a firm's dependence on exigent markets greater than ever, but so is its dependence on external resources for production and innovation. For example, firms now generally rely more on outsourcing non-core activities [4]. The competition for supplies is also greater than ever in a global economy [4]. All these trends have led scholars to present purchasing-marketing functional integration (PMFI) as the necessary internal complement to supply chain integration initiatives [7, 8].

Nevertheless, PMFI is still in need of theoretical and empirical development. Firstly, very few empirical articles support recent conceptual contributions. Secondly, the literature has reported that the roles customers and suppliers play in increasing NPD success are undeniable. NPD benefits from the integration initiatives of external customers and suppliers [11]. However, the necessary internal integration of the two functions representing them within companies (purchasing and marketing), and complementing these external integrations, has scarcely been addressed in the NPD context. Thirdly, there is a need to separately analyse PMFI's impact on different business objectives (e.g., quality, costs, and speed), as extant studies rely on economic performance measures or aggregated measures that combine different operational objectives [8, 12]. Fourthly, recent decades have witnessed a development in contingent thinking on strategic management that has also been reflected within the context of NPD [12, 13]. There is a scarcity of empirical studies that analyse PMFI under different circumstances (e.g., stable vs. dynamic markets, radical vs. incremental innovation). The contingent approach helps to explain when PMFI acquires more relevance. Fifthly, there is a call for studies that use different theoretical lenses to explain the functional integration phenomenon [14]. Finally, scholars also signal the need to conduct studies that focus on analysing functional integration per se, or integration components (e.g., information sharing or communication) instead of studies that focus on analysing practices that could enable this functional integration (e.g., functional diversity, physical proximity, computer-aided design instruments, common rewards, etc.) [14, 15]. This study will refer to these practices or drivers as "integrating mechanisms" [2].

The particular barriers these two functions encounter [9, 10] and the variability in the way purchasing integration has been conceptualised and operationalised in previous research make it difficult to simply extrapolate their results to this particular purchasing-marketing case. This paper therefore seeks to expand previous PMFI by helping to fill the abovementioned gaps. Consequently, it focuses on the NPD level instead of the company level; and it adopts theoretical background, specifically Information Processing Theory (IPT) to analyse PMFI by focusing directly on its main component, information shared and understood, rather than on PMFI's possible drivers. This research differs from similar prior studies by separately analysing the two information flows involved in PMFI: from marketing to purchasing and vice versa. It also differs from previous studies by conducting a contingent empirical analysis of these two flows and their effect on a typical performance indicator, namely, speed [2, 12, 13, 16], in different environments in terms of dynamism.

The rest of the article is structured as follows: Section 2 includes the study's theoretical underpinnings. Section 3 introduces an explanatory model of the impact that the purchasing information shared and understood with/by marketing has on NPD speed, as well as of the capacity environmental dynamism has to moderate that relationship. Section 4 presents the methodology used to test the hypotheses. Section 5 presents the results of the analysis. Section 6 discusses the results' implications and summarises the study's conclusions.

2. Theoretical background

2.1 Information processing theory

IPT was adapted by Galbraith [17] from developmental psychology. IPT [17, 18] considers that companies are open social systems that subdivide into specialised subunits as they grow [18]. These subunits become interdependent when they undertake mainstream tasks requiring their particular expertise. This specialisation and interdependency inform the myopic perspectives and limited rationality of functional decision-makers, causing uncertainty. Companies should therefore deal with the uncertainty derived from both the environment they are exposed to and the mainstream tasks involving interaction across specialised subunits. To reduce this uncertainty, organisations become information processing systems. "Information processing refers to the gathering, interpreting, and synthesis of information in the context of organisational decision making" [17, p. 614]. The greater inter-unit task interdependence and the more unstable the environment, the greater the need to process information [18].

IPT represents an ideal theoretical framework for our work for several reasons: Firstly, the specific unit of analysis, NPD, adapts perfectly to IPT assumptions. NPD is an interdependent task requiring different specialised functional areas or subunits (e.g., marketing, R&D, manufacturing, purchasing, etc.) to interact with each other during different stages, such as the selection of ideas, appraisal and approval of the product's final prototype, production, and distribution. Each one of these interactions affects its outcomes. Their interdependence during this process generates uncertainty. They are also subject to external uncertainty, as to succeed they need to adapt to unstable markets and suppliers. NPD therefore requires information processing.

This leads us to the second reason for relying on IPT. Information Processing occurs within and between subunits of functional areas, being described therefore as an integrating concept [17, 18]. This argument explains our use of information exchange as a proxy for functional integration. This notion is reinforced by the fact that information exchange or interaction has traditionally been seen as one of the essential components of functional integration [19-22].

Thirdly, IPT development and implications will help to explain the relationship between the variables under study: purchasing-marketing information exchange, NPD speed, and dynamism. The relationship between information processing and delivery times has been explained by IPT [18, 23], as has the relationship between information processing and dynamic environments [18]. The proposed model (Section 3) further develops these arguments.

Previous studies on functional integration have also relied on IPT [20-22], reinforcing the validity of this approach. Each one of these contributions focuses on different performance indicators and levels of analysis, providing different insights into the functional integration phenomenon and complementing each other. This paper expands this approach by providing a novel application of IPT. Specifically, it considers that information processing requires several flows of information that should each be analysed separately.

2.2 The moderating role of dynamism

The contradictory results often obtained when analysing the effect that one variable has over another has led scholars to formulate contingent hypotheses. This contingent approach has also been applied to the analysis of NPD success factors, such as NPD speed. Along these lines, Carbonell and Rodriguez [12], for example, have investigated how the effect of different factors on NPD speed depends on the product's technological complexity. Swink [12] has studied how the capacity different mechanisms have to streamline NPD varies depending on whether there is an accelerating intention in their use.

Studies like these indicate that both operating conditions and environmental conditions moderate the effect that different variables have on NPD speed. Dynamism has been one of the main variables used in the literature to characterise this environment. This attribute refers to the presence of instability, turbulence, volatility, or degree of change [24, 25]. This dynamism has also been related to functional integration and the need for interfunctional communication. In more dynamic or changing environments, more anomalies are recorded, and it is more difficult to predict what is going to happen next, and so greater uncertainty is therefore generated [24]. This uncertainty limits the possibility of normalising processes [18]. In such circumstances, an organisation needs to resort to other coordination and integration mechanisms, such as lateral relations [18]. It may therefore be concluded that this dynamism is a relevant factor to be considered when analysing the relationship between shared information and the outcomes of NPD processes.

3. Model and hypotheses

IPT establishes a relationship between information processing and delivery time, stating that when an organisation's information processing capacity is insufficient to cope with the uncertainty it faces, a possible solution is to extend completion dates [17]. Consequently, increasing information processing will reduce delivery times. IPT explains the nature of the impact information processing has on cycle time, whereby information processing reduces the bounded rationality specialised subunits are subject to, leading to a common understanding [22, 23]. This enables more robust and optimal decisions [17], avoiding workarounds and redundancies, which in turn avoid time-wasting and reduce the cycle time [22, 23].

This paper applies this reasoning to the purchasing and marketing case during NPD. These two functions' exposure to environmental uncertainty is high, as they depend on the unknown reactions of external agents such as customers, competitors, or suppliers. If purchasing assimilates marketing information, such as customer preferences, purchasers can start the provisioning process from the early stages of NPD, even when product specifications have not been fully determined. This practice will also allow the earlier booking of manufacturing capacity in suppliers' production plants and obtaining better delivery dates. Understanding market preferences will avoid wasting time in searching for and evaluating supplier options that are not aligned with them.

Despite the time taken up by inter-functional communication, the absence of this information will impede proper coordination between the two functions' requirements and capabilities, causing delays due to the need to conduct reviews as the NPD process advances. This reasoning leads us to formulate the following hypothesis:

H1: The extent to which information related to the marketing function is shared and understood with/by the purchasing function is positively related to the speed of the NPD process.

Likewise, having more knowledge on suppliers' requirements and capabilities from the very first stages of the process will enable marketing to promote product options along the different stages of the NPD process, which being more closely aligned with customers' expectations will be more consistent with those resources and outside capabilities. Such requirements may be taken into account at the prototype testing stage. Swink and Song [26] have reported that including questions related to manufacturing issues in market surveys ensures a closer alignment between manufacturing and marketing. The same happens with suppliers' manufacturing processes. This better alignment from the very first stages will subsequently avoid provisioning problems or having to change a product's specifications, with the ensuing delay this entails. During the production stage, if marketing staff better understand the impact decisions on volume and variety have on the supplier's production process (e.g., adding another colour to a component's production line), they will be more likely to propose more closely aligned product specifications [26]. In a word, a greater understanding of the resources and capabilities available in the suppliers' market enables the marketing function to investigate and promote NPDs based on them in a speedy and direct manner, avoiding having to review negotiations, with the ensuing waste of time.

H2: The extent to which information related to the purchasing function is shared and understood with/by the marketing function is positively related to the speed of the NPD process.

Secondly, the model postulates that the abovementioned effects are moderated by environmental dynamism. In those environments that record more changes in technologies, suppliers, customers or competitors, there will be a greater need to vary a product's specifications in order to adapt to the new situation, more decisions will have to be made during NPD, and there will be more uncertainty over what is going to happen next and over the implications for other functions. According to IPT, the greater the uncertainty, the greater the need for processing information for making more robust decisions that pre-empt unexpected problems [17, 18]. Applying this reasoning to the logic of our two first hypotheses, it may be deduced that purchasing's greater assimilation of marketing information, as well as the opposite, will be more useful for avoiding workarounds and redundancies in more dynamic and changing environments. The literature on purchasing-marketing integration has also reported a greater need for communication between purchasing and marketing in order to save time in more agile environments [7].

In contrast, within a context in which there are fewer changes in the environment, there will in turn be fewer changes in the specifications of NPDs, and fewer decisions to be made. Having more information on purchases within this context will take more time than can be saved via more aligned decisions. Based on this reasoning, the two following hypotheses are formulated.

H3: Environmental dynamism positively moderates the relationship between the extent to which information on the marketing function is shared and understood with/by the purchasing function and the speed of the NPD process.

H4: Environmental dynamism positively moderates the relationship between the extent to which information on the purchasing function is shared and understood with/by the marketing function and the speed of the NPD process.

4. Methodology

4.1 Data

To test our model we selected a population 535 firms from the "top ten technological areas in terms of patent application" as per the 2013 ranking of the OEPM (Spanish Office of Patents and Trademarks). These firms are expected to be intensively involved in NPD, which should ensure good levels of knowledge regarding the concepts being studied. Specifically, we selected the following sectors: Fabricated Metal Products, except Machinery and Transportation Equipment (SIC 34), Transportation Equipment (SIC 37), and Electronic and other Electrical Equipment and Components, except Computer Equipment (SIC 36). From these three sectors, we retained those firms with more than 50 employees. The higher the number of employees in a firm, the higher the possibility that that firm is organised into specialised functions. Firms organised into specialised subunits are more likely to need internal integration for successful NPD. Selecting bigger firms from the initial population will ensure further good levels of knowledge on our model's variables.

A questionnaire was sent to these companies, with 197 being returned: 56 of them were unusable because of the high rate of missing data. Out of the remaining 141 surveys, five had a missing value. To complete the missing value in these five incomplete surveys, respondents were contacted by email or phone. This initiative allowed completing another two questionnaires, with the one missing value in the remaining three surveys being replaced by the mean score. This therefore provided a total of 141 surveys for our analysis, at a response rate of 26 %. The sampling error was ± 7.08 %, with a 95 % level of confidence. An ANOVA analysis was conducted to test for differences between the sample and the population, and between the first and the last 25 respondents regarding two demographic variables (number of employees, total assets). Not significant differences were found which indicates that non-response bias is not a problem in our sample.

Purchasing managers were selected as key informants. Key informants should be knowledgeable about the matters being studied and willing to discuss them (Kumar *et al.*, 1993). For this study's particular purposes, purchasing managers were considered competent to assess both information flows, as they have a long tradition of receiving requirements from other functions and providing them with information [8, 10]. While the marketing function also has an extended culture and tradition for relaying commercial information to other subunits, the literature has reported that the tradition of receiving purchasing information may be weaker [9,10]. Hence, they may not realise they are missing useful supplier-related information. Purchasers' incentives are frequently related to the fulfilment of suppliers' pre-scheduled delivery dates, which means they are well informed on this variable. They were considered to be in a position to report the environment's degree of dynamism during a particular NPD that has already been completed. The response rate and low levels of missing data confirm their validity as informants.

Using a single data source involves the risk of common method bias. Although there are some statistical procedures for dismissing this bias, Podsakoff *et al.* [27] state that they are not suitable when formative constructs are an integral part of the study. This study uses formative constructs as explained in Section 4.2. In such cases, they recommend relying on different procedural remedies when designing the questionnaire. We have therefore applied their recommendations for counterbalancing the order of the dependent and independent variable questions in the survey and separating the variables' measurement using different response formats. Once the answers had been obtained, a Harman Test [28] was conducted to control for this bias. This test revealed that the variables in our analysis do not load in a single factor, but in many different ones; five of them record an eigenvalue higher than 1, and each one of them accounts for a low percentage of the variance. We may therefore conclude that common method bias is not a significant problem in our sample.

4.2 Metrics

Respondents were asked to answer the survey questions referring to a specific NPD that met the following criteria: (1) the selected NPD's specific process had to have already finished and information on its level of success had to be available (e.g., sales, market share, etc.), and (2) the selected NPD's supply activities had to be managed by the respondent's team. In order to avoid receiving surveys referring solely to successful NPDs, respondents were asked to select the last NPD that met these two criteria. To ensure all the answers refer to the selected NPD, respondents were asked to describe the product at the beginning of the survey.

Based on IPT, functional integration was measured through the degree of information shared and understood by these two functions. The four items for its measurement (see Table 1) were adapted from Schoenherr and Swink [21], and inspired by the work of Gupta *et al*. [29].

An Exploratory Factor Analysis (SPSS, 23) and a Confirmatory Factor Analysis (AMOS, 23) were conducted to confirm that purchasing-marketing information assimilation split into these two dimensions. These EFA and CFA analyses are shown in Table 1. CFA results reveal adequate levels of the fit indicators -$\chi 2$/df, GFI, AGFI, TLI, CFI, and RMSA - according to generally accepted reference values [30, 31]. The indicators of reliability, Cronbach's Alpha and Composite Reliability, also recorded recommended values [31]. Convergent validity was confirmed by the levels of the standardised coefficients, as well as by their significance. Discriminant validity is also satisfactory, as each construct's average variance extracted (AVE) is higher than the square of its estimated correlation with the other constructs.

The scales used for measuring both the speed of the NPD process and environmental dynamism fulfilled the assumptions of the formative constructs [32, 33]. We built their metrics by computing the mean of the scores assigned to the different items used for measuring each one of them, as featured in Tables 2 and 3. This scale makes interpretation and replication easier.

NPD speed was measured through a multivariate construct. The heads of purchasing were asked to use a Likert-type scale to rate the speed, related to the schedule (from 1 – well below to 7 – well above), of five specific stages of the NPD process. Relative measures of NPD speed allow comparing the results for very different types of products, and are frequently used by researchers [12, 13, 34, 35]. Scheduled NPD time is often used as a yardstick in these relative measures [34], and it was selected for this study because purchasing managers are often rewarded for meeting NPD deadlines and are familiar with this information. When these deadlines are set by companies, they already take into account competition lead times. Finally, to ensure respondents used the same stages when scoring NPD speed, as Chen *et al*. [35] recommend, the respondents were asked to rate the speed of five specific stages of the NPD process.

Environmental dynamism was measured through a multivariate construct, asking the purchasing managers to rate the extent to which a series of statements matched their own experience using a Likert-type scale from 1 (not at all) to 7 (completely/intensely). These statements were adapted from the studies by González Benito *et al*. [25] and Miller and Friesen [24].

As explained earlier, we have intentionally avoided the inclusion of integrating mechanisms in our analysis. However, to control for alternative explanations, other possible antecedents of NPD speed have been included in the study: (1) Participant's exclusive dedication to NPD. Following Carbonell and Rodríguez [13], the heads of purchasing were asked to use a Likert-type scale from 1 (not at all) to 7 (fully) to rate the extent to which the participants in the NPD process were dedicated full-time to the project. (2) Participants' experience in similar NPD processes. Again, following Carbonell and Rodríguez [13], the purchasing managers were asked to use a Likert-type scale from 1 (not at all) to 7 (fully) to rate the selected participants' experience in NPD. (3) Clarity of objectives. Based on Swink [12], the heads of purchasing were asked to use a Likert-type scale from 1 (not at all) to 7 (fully) to rate the extent to which clear and explicit objectives had been set for the selected NPD process. (4) Firm size. This data was retrieved from the SABI database.

Table 1 Information shared and understood: Exploratory and confirmatory factor analyses

		Exploratory factor analysis		Confirmatory factor analysis	
		F1	F2	F1	F2
Marketing information shared and understood	Purchasing managers/professionals received enough commercial information (e.g., product strengths and weaknesses, trends, market threats and opportunities) to efficiently play their role during the NPD process.	0.903	0.147	0.796	
	Purchasing managers/professionals reached a high degree of understanding on the commercial implications of the decisions made during the NPD process.	0.886	0.194	0.832	
Purchasing information shared and understood	Marketing/Commercial managers/professionals received enough purchasing information (e.g., available suppliers, materials and components, costs, quality, deliveries) to efficiently play their role during the NPD process.	0.099	0.905		0.675
	Marketing/Commercial managers/professionals reached a high degree of understanding on the purchasing implications of the decisions made during the NPD process.	0.250	0.854		0.901
		Varimax rotation explained variance 81.98 %		X^2/g.l. = 0.645 GFI = 0.998 AGFI = 0.977 TLI = 1.017 CFI = 1 RMSA = 0	
		Cronbach's alpha:		0.797	0.756
		Composite reliability:		0.797	0.772
		AVE		0.663	0.634
		Squared correlation estimates:			0.228
		F2: Purchasing information shared		0.228	

Table 2 Speed of NPD

Speed of NPD	Speed, related to schedule, in brainstorming and screening stage
	Speed, related to schedule, in design stage
	Speed, related to schedule, in final prototype testing stage
	Speed, related to schedule, in production stage
	Speed, related to schedule, in transportation stage (from suppliers to point of sale)

Table 3 Dynamism

Dynamism	Consumer/customer preferences have changed very often.
	The technology and/or design trends in our industry have changed very often.
	Our key competitors' commercial strategies and actions have changed very often (campaigns, promotions, new openings, products, etc.).
	Our main suppliers' actions have changed very often (qualities, prices, timings, and conditions of service, etc.).

4.3 Analysis

In order to verify the hypotheses considered in the conceptual model, multiple regression analysis has been conducted for the direct effects, and moderated multiple regression (MMR) analysis for the moderating effects [31]. These two techniques are among the most popular ones for testing direct effects and interaction effects in social sciences. Estimating interaction effects using MMR involves creating different models, or regressions. The first model includes only the control variable, the second adds the independent variable, the third one the moderator, and finally, the fourth one adds the interaction between the independent variable and the moderator.

5. Results and discussion

The correlations and descriptive statistics for the variables in our study are included in Table 4. Table 5 shows the results of four different models used to test the effect on NPD speed of the marketing information shared and understood with/by purchasing, as well as the moderation of environmental dynamism on that effect.

Table 4 Mean, standard deviation, and Pearson correlation coefficients

	Mean	S.D.	1	2	3	4	5	6	7	8
1. Exclusive dedication to the NPD	3.60	1.99	1							
2. Experience in NPD	5.59	1.36	0.272**	1						
3. Clear objectives of the NPD	4.99	1.56	0.246**	0.261**	1					
4. Firm Size	2519	24503	0.115	0.095	0.115	1				
5. Marketing information shared and understood	4.50	1.42	0.045	0.258**	0.291**	0.120	1			
6. Purchasing information shared and understood	4.82	1.21	-0.157+	0.043	0.201*	0.080	0.374**	1		
6. Dynamism	4.65	1.08	0.136	0.050	-0.008	0.164+	-0.004	0.047	1	
7. Speed of NPD	4.66	0.89	0.041	0.295**	0.257**	0.089	0.289**	0.135	0.076	1

Note: +p < 0.1 *p < 0.05; **p < 0.01; ***p < 0.001. Pearson Correlation Coefficients (bilateral).

Table 5 Effect of 'Marketing information shared and understood' on the 'Speed of NPD', moderated by environmental 'Dynamism'

	Model 1	Model 2	Model 3	Model 4
Exclusive dedication to the NPD	-0.074	-0.057	-0.069	-0.070
Experience in NPD	0.257**	0.219*	0.218*	0.220*
Clear objectives of the NPD	0.195*	0.147+	0.153+	0.155+
Firm Size	0.051	0.036	0.025	0.016
Marketing information shared and understood.		0.185*	0.186*	0.008
Dynamism			0.074	-0.066
Marketing information shared and understood x Dynamism				0.229
R2	0.126	0.156	0.161	0.162
Δ R2	0.126**	0.030*	0.005	0.001
F	4.905**	4.980***	4.281**	3.677**

+p < 0.1; *p < 0.05; **p < 0.01; ***p < 0.001

Table 6 Effect of 'Purchasing information shared and understood' on the 'Speed of NPD', moderated by environmental 'Dynamism'

	Model 1	Model 2	Model 3	Model 4
Exclusive dedication to the NPD	-0.074	-0.057	-0.068	-0.084
Experience in NPD	0.257*	0.254*	0.253**	0.245**
Clear objectives of the NPD	0.195*	0.177*	0.183*	0.170*
Firm size	0.051	0.046	0.036	0.009
Purchasing information shared and understood.		0.076	0.071	-0.794*
Dynamism			0.067	-0.819*
Purchasing information shared and understood x Dynamism				1.286*
R2	0.126	0.131	0.136	0.170
Δ R2	0.126**	0.005	0.004	0.035*
F	4.905**	4.083**	3.504**	3.895**

+p < 0.1; *p < 0.05; **p < 0.01; ***p < 0.001

Model 2 reveals that the effect that purchasing's assimilation of more marketing information has on NPD speed is positive and significant. This result confirms our first hypothesis (H1). However, Model 4 reveals that the effect of the interaction between environmental dynamism and purchasing's assimilation of NPD marketing information is not significant, at least insofar as our sample is concerned. Our third hypothesis (H3) is not therefore confirmed.

Tables 6 shows the equivalent models for the opposite flow of information, the one related to supplier aspects that go from purchasing to marketing. In this case, Model 2 shows that the effect of this second flow of information on NPD speed is not significant. This means our second hypothesis (H2) is not confirmed in our sample. Nevertheless, Model 4 shows how the interaction between the purchasing information assimilated by marketing and environmental dynamism do indeed have a significant impact on the speed of the process. We may therefore accept our fourth hypothesis (H4): environmental dynamism moderates the impact that the purchasing information assimilated by marketing has on NPD speed.

Fig. 1 shows an interaction plot revealing the nature of the relationship between the purchasing information assimilated by marketing and NPD speed in three different contexts (when environmental dynamism is low, medium and high). It is clear that the relationship between more purchasing information shared and understood with/by marketing and the speed of the NPD process is positive when the dynamism is medium or high, confirming our model's logic. Only when the dynamism is low does sharing more purchasing information have a negative impact on speed.

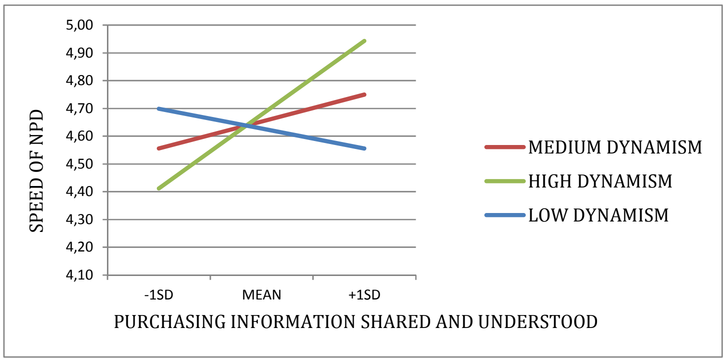


Fig. 1 Effect of the interaction between 'Purchasing information shared and understood' and environmental 'Dynamism' on 'Speed of NPD'

6. Conclusion and implications

The literature has recently stressed the need to integrate the purchasing function specifically with the marketing function [4-9]. The need to develop integrated systems that take into account supplier and customer integration during NPD has also been stressed [11]. However, further theoretical and empirical research in this field is needed. This study has contributed to its state-of-the-art by empirically studying how environmental dynamism moderates the impact that each one of the information flows PMFI requires has on NPD speed. The results show that these two flows behave differently. Marketing information assimilated by purchasing accelerates NPD regardless of the level of environmental dynamism. In contrast, purchasing information assimilated by marketing accelerates NPD only when environmental dynamism is medium or high, but when it is low it actually reduces NPD speed.

This research has straightforward managerial applications. It invites managers to assess whether NPD speed is relevant for their companies to compete, and what their environmental degree of dynamism is. It shows which flow of information should be promoted in each context, and it indirectly invites an evaluation to be made of the integrating mechanisms that should be used to promote each flow. González-Zapatero *et al*. [36] have reported that certain traditional integrating mechanisms do not foster both flows of information. Such was the case of the mere physical proximity of the two functions. Perhaps because of the lesser tradition in sharing purchasing information, this co-location simply promoted the flow of information from marketing to purchasing, but not the other way around. Companies deciding to adopt a concurrent engineering approach to NPD could also find our research findings useful. Concurrent engineering involves considering different function requirements and preferences from the beginning of the NPD process [37], especially those of the manufacturing function [16]. Taking into account, for instance, the requirements of the production function without ensuring a good marketing-purchasing integration may delay the NPD process, especially in more dynamic environments. The production function may spend time finding good alignments with marketing or R&D requirements and resources, but then when they try to obtain the supplies needed to start manufacturing, they may discover that these supplies are not available. Purchasing-marketing integration is therefore an important link in a concurrent engineering approach. This study's findings are also useful for SMEs, especially in transactional economies (e.g., Serbia). In these economies, NPDs are mainly pull market-driven [38]. This research paper may help them to see the potential of including a supplier's voice in NPD by fostering PMFI, especially as environments become more dynamic.

This study also has important implications for academics. Firstly, this work adds to that of other scholars who highlight the importance of addressing functional integration itself directly through the study of its different components, instead of studying it indirectly through different integrating mechanisms. This direct analysis shows each component's true potential. However, studying FI directly poses a challenge because, as scholars recurrently decry, there is a lack of consensus about the concept itself [13, 18]. In this paper, we have relied on IPT to study information processing as a proxy for PMFI. Our proposal draws attention to the fact that information processing implies several flows and should be taken into account both in defining and operationalising the concept. Further developments of the concept of functional integration [13] will enrich this analysis. Secondly, there is a pressing need to conduct more contingent studies. Some of the most recent conceptual studies on the purchasing-marketing link report the expediency of managing this link in a different way depending on the context in which a company operates [6,5]. Wagner and Eggert [5], for example, contend that the purchasing-marketing link should be managed differently depending on a company's level of dependence on its customers and suppliers. The margin of time for managing this dependence may also have a bearing on how to do so, as this work suggests. Thirdly, the NPD process requires integrating other functions (R&D, marketing, production, logistics, etc.). The activities that all these functions undertake during the NPD process may be organized into different sequences [37, 39, 40]. Some scholars have stressed that different sequences, also labelled product development process (PDP) architectures [40], may have a greater impact on NPD performance than others. This paper's findings should be taken into account by academics studying the efficacy of different NPD or (PDP) architectures.

This research has certain limitations; for instance, the scales used to measure the variables are perceptual. Although these scales rely on extant literature, and although the Harman test [28] shows that common method bias does not seem to be a significant problem for this analysis, combining different data sources will constitute a useful extension. Including other control variables, such as the degree of dependence on customers or suppliers or the degree of innovation, could also help to explain the purchasing-marketing link. Considering the theoretical implications described above and overcoming these limitations provide a path and orientation for further research. Other developments of this analysis would be to identify possible practices that foster or help to manage the desired information flows. Although some scholars have studied possible PMFI drivers [36], the potential new available technologies have to manage this link remain unexplored. These technologies include big data business analysis, artificial intelligence, the internet of things, and virtually reality. In the era of Big Data the traditional techniques used to process information may be insufficient in some contexts [41-43].

Acknowledgement

This research was funded by the Spanish Government (Ministerio de Economía y Competitividad) through research project ECO2016-76876-R. Funding was also received from the research project SA027U16 financed by the Consejería de Educación de la Junta de Castilla y León (Department of Education of the Autonomous Community of Castile and Leon).

References

[1] Narver, J.C., Slater, S.F. (1990). The effect of a market orientation on business profitability, *Journal of Marketing*, Vol. 54, No. 4, 20-35, doi: 10.1177/002224299005400403.

[2] Leenders, M.A.A.M., Wierenga, B. (2002). The effectiveness of different mechanisms for integrating marketing and R&D, *The Journal of Product Innovation Management*, Vol. 19, No. 4, 305-307, doi: 10.1111/1540-5885.1940305.

[3] Swink, M., Song, M. (2007). Effects of marketing-manufacturing integration on new product development time and competitive advantage, *Journal of Operations Management*, Vol. 25, No. 1, 203-217, doi: 10.1016/j.jom.2006.03.001.

[4] Sheth, J.N., Sharma, A., Iyer, G.R. (2009). Why integrating purchasing with marketing is both inevitable and beneficial, *Industrial Marketing Management*, Vol. 38, No. 8, 865-871, doi: 10.1016/j.indmarman.2008.12.021.

[5] Smirnova, M., Henneberg, S.C., Ashnai, B., Naudé, P., Mouzas, S. (2011). Understanding the role of marketing-purchasing collaboration in industrial markets: The case of Russia, *Industrial Marketing Management,* Vol. 40, No. 1, 54-64, doi: 10.1016/j.indmarman.2010.09.010.

[6] Wagner, S.M., Eggert, A. (2016). Co-management of purchasing and marketing: Why, when and how?, *Industrial Marketing Management*, Vol. 52, 27-36, doi: 10.1016/j.indmarman.2015.07.012.

[7] Toon, M.A., Morgan, R.E., Lindgreen, A., Vanhamme, J., Hingley, M.K. (2016). Processes and integration in the interaction of purchasing and marketing: Considering synergy and symbiosis, *Industrial Marketing Management*, Vol. 52, 74-81, doi: 10.1016/j.indmarman.2015.07.014.

[8] Ziggers, G.W., Henseler, J. (2016). The reinforcing effect of a firm's customer orientation and supply-base orientation on performance, *Industrial Marketing Management*, Vol. 52, 18-26, doi: 10.1016/j.indmarman.2015.07.011.

[9] Matthyssens, P., Bocconcelli, R., Pagano, A., Quintens, L. (2016). Aligning marketing and purchasing for new value creation, *Industrial Marketing Management*, Vol. 52, 60-73, doi: 10.1016/j.indmarman.2015.07.016.

[10] Luzzini, D., Ronchi, S. (2016). Cinderella purchasing transformation: Linking purchasing status to purchasing practices and business performance, *Production Planning & Control*, Vol. 27, No. 10, 787-796, doi: 10.1080/09537287.2015.1137986.

[11] 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 – Technical Gazette,* Vol. 22, No. 1, 191-200, doi: 10.17559/TV-20140623130536.

[12] Swink, M. (2003). Completing projects on-time: How project accelerations affect new product development, *Journal of Engineering and Technology Management,* Vol. 20, No. 4, 319-344, doi: 10.1016/j.jengtecman.2003.08.002.

[13] Carbonell, P., Rodríguez, A.I. (2006). Designing teams for speedy product development: The moderating effect of technological complexity, *Journal of Business Research*, Vol. 59, No. 2, 225-232, doi: 10.1016/j.jbusres.2005.08.002.

[14] Frankel, R., Mollenkopf, D.A. (2015). Cross-functional integration revisited: Exploring the conceptual elephant, *Journal of Business Logistics*, Vol. 36, No. 1, 18-24, doi: 10.1111/jbl.12081.

[15] Turkulainen, V., Ketokivi, M. (2012). Cross-functional integration and performance: What are the real benefits?, *International Journal of Operations & Production Management*, Vol. 32, No. 4, 447-467, doi: 10.1108/01443571211223095.

[16] Miranda González, F.J., Bañegil Palacios, T.M. (2002). The effect of new product development techniques on new product success in Spanish firms, *Industrial Marketing Management,* Vol. 31, No. 3, 261-271, doi: 10.1016/S0019-8501(00)00150-4.

[17] Galbraith, J.R. (1974). Organization design: An information processing view, *INFORMS Journal on Applied Analytics,* Vol. 4, No. 3, 28-36, doi: 10.1287/inte.4.3.28.

[18] Tushman, M.L., Nadler, D.A. (1978). Information processing as an integrating concept in organizational design, *Academy of Management Review*, Vol. 3, No. 3, 613-624, doi: 10.5465/amr.1978.4305791.

[19] Pagell, M. (2004). Understanding the factors that enable and inhibit the integration of operations, purchasing and logistics, *Journal of Operations Management*, Vol. 22, No. 5, 459-487, doi: 10.1016/j.jom.2004.05.008.

[20] Swink, M., Narasimhan, R., Wang, C. (2007). Managing beyond the factory walls: Effects of four types of strategic integration on manufacturing plant performance, *Journal of Operations Management*, Vol. 25, No. 1, 148-164, doi: 10.1016/j.jom.2006.02.006.

[21] Schoenherr, T., Swink, M. (2012). Revisiting the arcs of integration: Cross-validations and extensions, *Journal of Operations Management*, Vol. 30, No. 1-2, 99-115, doi: 10.1016/j.jom.2011.09.001.

[22] Swink, M., Schoenherr, T. (2015). The effects of cross-functional integration on profitability, process efficiency, and asset productivity, *Journal of Business Logistics*, Vol. 36, No. 1, 69-87, doi: 10.1111/jbl.12070.

[23] Hult, G.T.M., Ketchen, D.J., Slater, S.F. (2004). Information processing, knowledge development, and strategic supply chain performance, *Academy of Management Journal*, Vol. 47, No. 2, 241-253, doi: 10.5465/20159575.

[24] Miller, D., Friesen, P.H. (1983). Strategy-making and environment: The third link, *Strategic Management Journ*al, Vol. 4, No. 3, 221-235, doi: 10.1002/smj.4250040304.

[25] González-Benito, J., Reis da Rocha, D., Queiruga, D. (2010). The environment as a determining factor of purchasing and supply strategy: An empirical analysis of Brazilian firms, *International Journal of Production Economics*, Vol. 124, No. 1, 1-10, doi: 10.1016/j.ijpe.2009.08.032.

[26] Swink, M., Song, M., (2007). Effects of marketing-manufacturing integration on new product development time and competitive advantage, *Journal of Operations Management,* Vol. 25, No. 1, 203-217, doi: 10.1016/j.jom.2006.03.001.
[27] Podsakoff, P.M., Organ, D.W. (1986). Self-reports in organizational research: Problems and prospects, *Journal of Management*, Vol. 12, No. 4, 531-544, doi: 10.1177/014920638601200408.
[28] Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y., Podsakoff, N.P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies, *Journal of Applied Psychology*, Vol. 88, No. 5, 879-903, doi: 10.1037/0021-9010.88.5.879.
[29] Gupta, A.K., Raj, S.P., Wilemon, D. (1986). A model for studying R&D – Marketing interface in the product innovation process, *Journal of Marketing*, Vol. 50, No. 2, 7-17, doi: 10.1177/002224298605000201.
[30] Chau, P.Y.K. (1997). Reexamining a model for evaluating information center success using a structural equation modeling approach, *Decision Sciences*, Vol. 28, No. 2, 309-334, doi: 10.1111/j.1540-5915.1997.tb01313.x.
[31] Hair, J.F., Tatham, R.L., Anderson, R.E., Black, W.C. (1998). *Multivariate data analysis,* 5th edition, Prentice-Hall International, New York, USA.
[32] Diamantopoulos, A., Winklhofer, H.M. (2001). Index construction with formative indicators: An alternative to scale development, *Journal of Marketing Research*, Vol. 38, No. 2, 269-277, doi: 10.1509/jmkr.38.2.269.18845.
[33] Jarvis, C.B., MacKenzie, S.B., Podsakoff, P.M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research, *Journal of Consumer Research*, Vol. 30, No. 2, 199-218, doi: 10.1086/376806.
[34] Kessler, E.H., Bierly, P.E. (2002). Is faster really better? An empirical test of the implications of innovation speed, *IEEE Transactions on Engineering Management*, Vol. 49, No. 1, 2-12, doi: 10.1109/17.985742.
[35] Chen, J., Reilly, R.R., Lynn, G.S. (2005). The impacts of speed-to-market on new product success: The moderating effects of uncertainty, *IEEE Transactions on Engineering Management*, Vol. 52, No. 2, 199-212, doi: 10.1109/TEM.2005.844926.
[36] Gonzalez-Zapatero, C., Gonzalez-Benito, J., Lannelongue, G. (2016). Antecedents of functional integration during new product development: The purchasing–marketing link, *Industrial Marketing Management*, Vol. 52, 47-59, doi: 10.1016/j.indmarman.2015.07.015.
[37] Takeuchi, H., Nonaka, I. (1986). The new new product development game, *Harvard Business Review,* Vol. 64, No. 1, 137-146.
[38] Vorkapić, M., Radovanović, F., Ćoćkalo, D., Đorđević, D. (2017). NPD in small manufacturing enterprises in Serbia, *Tehnički Vjesnik – Technical Gazette,* No. 24, No. 1, 327-332, doi: 10.17559/TV-20150807185156.
[39] Urban, G.L., Hauser, J.R., (1993). *Design and marketing of new products,* 2nd edition, Englewood Cliffs, Prentice-Hall International, New York, USA.
[40] Yin, F.P., Gao, Q., Ji, X. (2017). Performance modelling based on value analysis for improving product development process architecture, *Advances in Production Engineering & Management,* Vol. 12, No. 1, 17-28, doi: 10.14743/apem2017.1.236.
[41] Xu, Z., Frankwick, G.L., Ramirez, E. (2016). Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective, *Journal of Business Research,* Vol. 69, No. 5, 1562-1566, doi: 10.1016/j.jbusres.2015.10.017.
[42] Wang, G., Gunasekaran, A., Ngai, E.W.T., Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications, *International Journal of Production Economics,* Vol. 176, 98-110, doi: 10.1016/j.ijpe.2016.03.014.
[43] Yang, Y., Fu, C., Chen, Y.-W., Xu, D.-L., Yang, S.-L. (2016). A belief rule based expert system for predicting consumer preference in new product development, *Knowledge-Based Systems,* Vol. 94, 105-113, doi: 10.1016/j.knosys.2015.11.012.

APEM journal

Advances in Production Engineering & Management
Volume 14 | Number 2 | June 2019 | pp 225–238
https://doi.org/10.14743/apem2019.2.324

ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

A new framework for complexity analysis in international development projects – Results from a Delphi study

Gajić, S.[a,*], **Palčič, I.**[b]

[a]University of Novi Sad, Faculty of Technical Sciences, Novi Sad, Serbia
[b]University of Maribor, Faculty of Mechanical Engineering, Maribor, Slovenia

A B S T R A C T

The main objective of this paper is to develop a framework for characterising project complexity in International Development (ID) projects. Contemporary challenges in ID projects have led to their growth in their complexity, which in recent years has driven researches in recent years to publish numerous papers that deal with this topic, demonstrating its importance in current project management research. Nevertheless, existing literature lacks in generally accepted framework that considers specifics of project complexity in ID projects. Thus, new framework was developed, based on a two-round Delphi survey, building upon existing TOE (technology-organisation-environment) framework with new empirical insights given from the experts in the field of ID projects. The main contribution of the paper is the validation of existing TOE complexity factors, in the context of International Development projects. Additionally, eight new complexity factors were proposed by the experts, and it was concluded that Environmental complexity had the biggest impact on International Development projects. From a managerial perspective, proposed complexity framework can be used for making a complexity footprint, which could indicate the critical areas of the project where complexity could be expected. In addition, the model represents a novel theoretical lens for assessing complexity in ID projects.

A R T I C L E   I N F O

Keywords:
International development projects (ID);
Project management;
Complexity;
Project complexity;
Technology-organisation-environment (TOE);
Delphi study

**Corresponding author:*
gajic.sladjana@uns.ac.rs
(Gajić, S.)

Article history:
Received 20 December 2018
Revised 15 May 2019
Accepted 20 May 2019

1. Introduction

Over the last ten years, interest in project management has significantly grown. It has been reported that about 24 % of the world GDP ($19 trillion) is spent on projects every year [1]. Project failure seems to be the rule in all types of projects [2], and this seems to be particularly true for International Development (ID) projects that have the ultimate objective to reduce poverty or to improve governance and build institutional capacity [3-5]. Contemporary challenges (dynamic and uncertain environment, increasing number of stakeholders) that influence these projects are closely related to the complexities of these projects [6-10]. It has been acknowledged that there is a correlation between project performance and complexity, and still, there is a huge knowledge gap about how complexity relates to project management practice and no widely accepted framework of project complexity in ID projects [11-14].

Numerous papers have been published in the field of project complexity, with intention to explain the relationship between complexity theory and project management [8, 9, 15, 16]. Management of ID projects requires a novel framework for dealing with project complexity due to their specifics in comparison to conventional projects. The framework in this paper includes some of the elements of TOE (technology-organisation-environment) framework, proposed by Bosch-Rekveldt, as well as the complexity model proposed by Vidal and Marle [9, 17]. This paper investigates complexity factors that contribute to overall complexity of ID projects from the perspective of the experts.

ID projects have specific context that increases their overall complexity. These projects cover almost every sector of activity, since they take place at multiple locations and in a different time zone. These are public-sector projects that show cultural complexity, unique context and institutional challenges. ID projects are specific due to their intangible, unique goals, unique ways of organizing, tool-intensity and the large number of stakeholders.

Delphi method has been used in this study to verify the significance of existing complexity factors proposed in the literature and to further update the list of complexity factors in the context of ID projects. Novel, modified complexity framework was developed based on the insights of the Delphi study. We found that experts take as most important the following complexity factors in ID projects: Clarity of goals, Variety of stakeholders' perspectives, Dependencies between tasks, Interface between different disciplines and Dependencies on other stakeholders. Most of the factors with the highest grade are in the group of Environmental complexity, which makes this type of complexity the most important for this type of projects. This conclusion is a unique contribution, in comparison with the results given from the similar studies done on different type of projects (see for example [17]).

The paper is organized as follows. Section 2 reviews recent works on complexity in the field of ID projects and presents the theoretical background of the proposed model. Section 3 introduces Delphi research, including data collection an analysis processes related to the Delphi questionnaires done by practitioners and experts. Section 4 proposes a refined measuring model for complexity in ID projects and shows the significance of this research. The final section presents conclusions, potential implications and limitations of the proposed model.

2. Literature review

2.1 International development projects

International development (ID) projects are the projects that deal mainly with poverty reduction, and cover sectors of agriculture, transportation, water, energy, health, population, education, reform and governance, etc. [18]. These are public projects, funded by donors from developed countries and are implemented in under-developed countries, which bring numerous political and cultural challenges [4].

ID projects have certain similarities with conventional projects: they deliver goods and services, they are limited, temporary and unique endeavours that go through project life cycle; these projects are constrained by the "iron triangle"- time, cost, and use project management standards, tools and techniques for the implementation [3, 19].

Peculiarities of ID projects are often interconnected with their not-for-profit, social, technical, and political nature; they are funded by external donors and have intangible and even conflicting objectives difficult to measure. In addition, they often have more stakeholders in comparison to conventional projects – at least three most important stakeholders: funding agency, implementing agency and the beneficiaries [3], that often have conflicting expectations. ID projects have specific context in which they occur, with numerous political, cultural, legal, social, technical, economic and environmental challenges. Important characteristics of ID projects are optimism bias, planning fallacy, strategic misrepresentation, and they are prone to media scrutiny, intolerance of failure, rigid procedures etc. [2].

Ika *et al*. [4] claim that due to their evident socio-political complexity, ID projects could "fit at the far right end of the spectrum on a continuum from private sector projects, through public sector projects, to international projects".

2.2 Project complexity

Projects can be observed as complex, self-organising systems, with their specifications (requirements and constraints) that undergo design process that is highly social, consisting of hundreds of designers, customers, and other participants [24]. Complexity has been recognized as one of the most important streams in project management research [20]. One of the first problems in understanding project complexity is lack of consensus regarding the definition of complexity in the project context [4, 6, 7, 13-16]. A review of recent papers had confirmed that definitions of complexity continue to be ambiguous [17, 18]. One of the most cited definitions is Baccarini's one defines complexity as "consisting of many varied interrelated parts and can be operationalised in terms of differentiation and interdependency" [21]. Followed by Baccarini's work, Williams explained that complexity consists of "structural complexity" – the number of varied components, "interdependency" – degree of dependence between these components and "uncertainty in goals and means" [22]. Geraldi *et al*. clarified complexity into structural, dynamical, uncertainty, pace and socio-political complexity. Sommer and Loch [23] define complexity as having "two dimensions": system size (the number of influence variables) and the number of interactions among influence variables.

The term "complex" stems from the Latin words cum (together, linked) and plexus (braided, plaited). The Oxford dictionary defines complexity as "consisting of parts" and "intricate, not easily analysed or disentangled." [15].

When defining complexity, it is important to make a distinction between two terms [8]: "complex" and "complicated". Complex systems contain multiple parts with several connections and interactions between the parts and behaviour that is a result of these emergent properties. Complicated systems emerge as the result of complex systems, without the right tools for analysis and management.

Vidal claims that there are two main scientific approaches to complexity [24]:

1) Descriptive complexity – this approach considers complexity as an intrinsic property of a project system,
2) Perceived complexity – complexity as subjective matter.

Bakhshi *et al*. concluded that there are three main schools of thought within the construct of complex projects: the Project Management Institute (PMI) perspective, the "System-of-systems" (SoS) approach, and the complexity theories perspective [25].

Cicmil and colleagues distinguish two different terms when discussing complexity [5]:

- Complexity in projects (how complexity can be manifested in projects),
- Complexity of projects (what factors make projects complex or difficult to manage).

The first stream is mainly theory-driven, and leans on complexity theories [5, 14, 19]. The second stream is practitioner-driven and aims to identify factors of complex projects and the strategies on how organisations can respond to complexity [20-23]. This paper focuses on the second stream.

Lack of consensus in defining project complexity leads to lack of understanding the concept. In this paper, we will accept Vidal's definition of project complexity as "the property of a project which makes it difficult to understand, foresee, and keep under control its overall behaviour, even when given reasonably complete information about the project system" [24].

2.3. Complexity factors in project management – Gathering elements from the literature

Numerous attempts have been made to measure and model project complexity (Appendix A – Complexity measuring in the literature), and most of them attempted to measure complexity quantitatively by focusing on the most important complexity factors [8, 9, 14, 17, 21, 22, 26-28]. In this paper, the classification of complexity factors proposed by Bosch-Rekveldt has been adopted – the TOE framework (Table 1), that includes Technological, Organizational and Environmental complexity factors [17]. It builds upon Baccarini's and Williams' existing complexity frameworks [21, 29].

Table 1 Included elements of TOE framework

Complexity type	Element name
Organisational complexity	Duration, Compatibility of different project management methods and tools, Size in CAPEX, Size of the project team, Number of locations, Resource and skills, Experience with parties involved, Interfaces between different disciplines in ID projects, Number of different nationalities in ID projects, Number of different languages, Cooperation of JV partners, Trust in project team (JV partner), Organisational risks
Technological complexity	Number of goals, Goal alignment, Clarity of goals, Scope largeness, Uncertainties in scope, Quality requirements, Number of tasks, Variety of tasks, Dependencies between tasks, Uncertainty in technical methods, Conflicting norms and standards, Newness of technology, Experience with technology
Environmental complexity	Number of stakeholders (internal and external), Variety of stakeholders' perspectives, Dependencies on other stakeholders, Political influence, Organisational internal support, required local content, Experience in the country of implementation, Stability of project environment, Risks from environment

De Bruijn already categorized complexity factors in three groups: technical complexity, social complexity and organizational complexity. This categorization has been furtherly developed by Jaafari, and also by Xia and Lee, but they have investigated the significance of TOE factors only in large engineering projects [17]. Technical view included technical content of the project, Organizational view included people and organizational aspects of the project, and Environmental view was mainly focused on the influences in the project environment on the complexity.

3. Methodology

3.1 Delphi method

As already mentioned, there is no clear consensus of the researchers on the complexity measures. Several studies have already been done in the area of project complexity, based on statistical calculations or surveys. ID projects are characterised with dynamic environments, numerous stakeholders, customized projects, and exposure to external conditions that often make traditional, research methods unrealistic for this type of research. The main benefit from conducting a qualitative study is validation of a local expression and ability to understand certain phenomena from the inside out [30]. Delphi method is designed to obtain reliable consensus about the topic from a panel of experts by conducting series of questionnaires combined with controlled opinion feedback, and with results of each round being fed into the next round [31]. In the field of management, a modified approach of Delphi method has been used to shape a group consensus about the relative importance of proposed issues [32]. Lindstone and Turoff [33] proposed the following definition: "Delphi may be characterized as a method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem. To accomplish this "structured communication" there is a need to provide the following: some feedback of individual contributions of information and knowledge; some assessment of the group judgment or view; some opportunity for individuals to revise views; and some degree of anonymity for the individual responses." Delphi method is more objective in its outcomes than individual statements, even though the judgments of experts are based on subjective opinions. One of the main advantages of the approach is the fact that direct confrontation of the experts is avoided [32]. Delphi method was widely used in the field of industrial engineering and project management [34-36]. There are several quantitative methodologies that could be used for investigating the relative importance of issues (e.g., Emerging Issues Analysis, Environmental Scanning, Issues Management, Analytical Hierarchical Process). Most of them are future oriented and outline individual opinions. The major advantage of Delphi in comparison to these methodologies is that it is the most prominent of consensus methodologies [37]. Additionally, Delphi approach was selected in this paper to reconcile different opinions between practitioners and experts about the importance of different complexity factors in ID projects. The validity of results was assured by heterogeneity of the panellists and anonymous response format.

Delphi study does not need to include a representative sample of any population. It consists of qualified experts who have a deeper understanding of the selected research issues, which makes the selection of the participants one of the most critical requirements [38]. Existence of bias was reduced in the study by implementation of a well-structured, academically rigorous process, and by selecting qualified experts for participation in the study in accordance with predefined guidelines - preparing a Knowledge Resource Nomination Worksheet.

In this paper two-round "ranking-type" Delphi was used to develop group consensus about the relative importance of complexity factors in ID projects (Figure 1 Delphi study algorithm). Purpose of the research was to develop a ranked list of most important complexity factors for ID projects. Two panels of participants were selected: the first group were academics, and the second group were practitioners – experienced project coordinators in ID projects.

Three-step strategy was adopted as a research program in the paper. Firstly, list of complexity factors that contribute ID projects was identified, based on the literature review (see Appendix A). Secondly, the identified factors were quantitatively tested, in order to verify if selected factors were truly relevant to the experts. Additional factors were proposed by the panellists. Thirdly, the selected factors were again ranked in the second round of the Delphi, and recommendations weremade. Biases are reduced by strategically constructed questionnaires, controlled feedback, detailed analysis of the group response and by two rounds of the research. Iteration is essential factor of any Delphi study. In this paper, iteration involved redistribution of the Delphi survey accompanied with controlled feedback, given to panellists with simple statistical summaries of the responses from the first round. This step, together with preserved anonymity of the participants, eliminates the dominance bias and minimizes the effects of the von Restorff effect [39].

Recent studies provide some theoretical discussion related to the complexity factors in the field of ID and their relative importance [4]. This paper contributes in obtaining a more comprehensive view from the perspective of the two major stakeholders in international development: practitioners and academics.

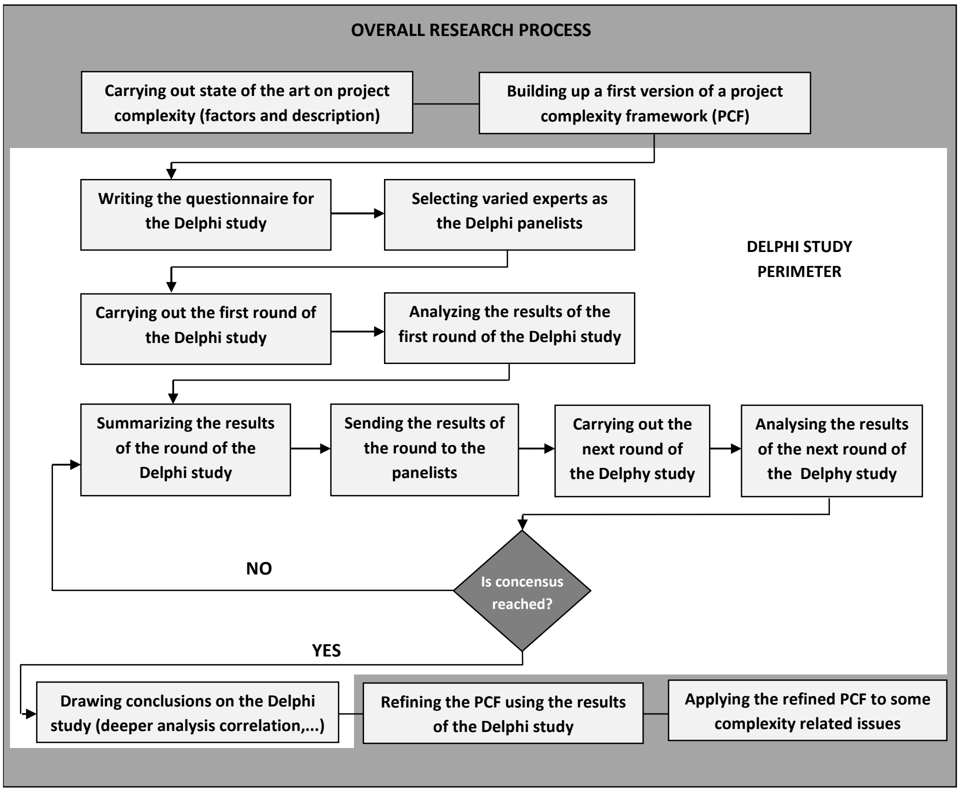


Fig. 1 Delphi study two round algorithm [24]

3.2 Participants selection

In this paper, experts were divided into two panels: academics and practitioners. Heterogenous group of participants allows for a somewhat different perspectives about the selected topic, as well as the comparison of the perspectives of the different stakeholder groups. A panel usually consists of 15 to 30 participants from the same discipline, or five to 10 per category from different professional groupings. Following the recommendations from Delphi literature, there was 11 participants in the first phase of the study, and 7 participants in the second phase of the study.

The following steps were included in selection of the experts for the Delphi study (Table 2) [38]:

Step 1: Preparing a Knowledge Resource Nomination Worksheet (KRNW) – identification of relevant disciplines or skills: academics and practitioners; identifying relevant academic and practitioner literature,
Step 2: Populating KRNW with names,
Step 3: Nominating additional experts by existing contacts,
Step 4: Inviting experts for each panel until the target size of the panel is reached.

The academics were selected based on the literature review of academic and practitioner journals. We have identified experts and asked them to nominate others for inclusion on the list.

They were provided with a brief description of the Delphi study and explanation that we have identified them as the experts on complexity in the field of project management and invited to participate in the study. The practitioners were selected from the base of Erasmus plus project coordinators. Seven academics and four practitioners agreed to participate in the Delphi study. Web was used as the mean for reaching focus organizations. Related literature focused on two most prominent SCI journals in the area of project management (International Journal of Project Management and Project Management Journal) were reviewed in order to identify articles concerning ID projects and complexity. Delphi questionnaire was administrated using e-mail and Survey Monkey software.

Table 2 Knowledge Resource Nomination Worksheet

Disciplines or Skills	Organizations	Related Literature
Academics	European Commission	Academics
Journal List		International Journal of Project management
Practitioners – Erasmus plus coordinators' list		Project Management Journal

4. Results and discussion

4.1 Data collection and analysis method

The Delphi questionnaires were administered using e-mail and the Web. One of the benefits of using these "rapid" media is increasing speed of the turnaround time between questionnaires, which is important factor in the Delphi method [32].

Administration of the questionnaires included the proposed procedure for "ranking-type" Delphi studies by Schmidt [40], that includes the following steps:

- Brainstorming for important factors and validation of the proposed factors,
- Narrowing down the original list to the most important ones,
- Ranking the list of important factors.

4.2 Brainstorming for important complexity factors

In this phase, panellists were asked to rank complexity factors on the five-level Likert scale. These factors belonged to the three main groups of factors: technological, organizational and environmental group. In addition, participants were asked to list additional relevant complexity factors in ID projects with a brief explanation for each factor. Duplicates were removed, new complexity factors were classified, and the terminology of the proposed factors was unified. After this, consolidated lists were sent to participants.

Validation of categorized list of factors - in this phase, panellists were given a list with all the consolidated factors obtained from the first questionnaire, grouped into categories, with brief explanation of each factor, based on the information from the first questionnaire. Furthermore, an exact copy of the responses from the first phase was sent to participants. Panellists were asked to verify that their answers were correctly interpreted and placed them in an appropriate category. According to Schmidt [40], "without this step, there is no basis to claim that a valid, consolidated list has been produced."

New elements were proposed by the experts: (1) Overlap of the project phases, (2) Interdependence of different stakeholders, (3) Diversity of stakeholder expectations, (4) Lack of clarity or consensuses on project benefits among project stakeholders, (5) Variation (1st type of uncertainty), (6) Foreseen uncertainty (2nd type of uncertainty), (7) Unforeseen uncertainty (3rd type of uncertainty), and (8) Chaos (4th type of uncertainty).

TOE factors were rated on a five-level Likert scale by the panellists. Consensus measurement has a pivotal role in Delphi research which could be defined as a gathering around median responses with minimal divergence [41]. Two criteria were selected for consensus measurement: (a) mean > 3 and (b) Interquartile range IQR < 1.

Six elements had the mean less than three, and were excluded from the next phase: (1) Quality requirements, (2) Duration of ID projects, (3) Compatibility of different project management methods and tools, (4) Size in CAPEX, (5) Number of different languages, and (6) Cooperation of Joint Venture partners.

Consensus between two groups of panellists was not reached (IQR>1) for the following elements, and they were rated again in the second phase: (1) Number of goals, (2) Goal alignment, (3) Clarity of goals, (4) Number of tasks, (5) Dependencies between tasks, (6) Conflicting norms and standards, (7) Newness of technology (world-wide), (8) Experience with technology, (9) Resource and skills availability, (10) Interfaces between different disciplines, (11) Trust within the project team (Joint Venture partner), and (12) Experience in the country of implementation.

4.3 Narrowing down the original list to the most important complexity factors

In the second phase, the list of factors was narrowed. Four participants did not proceed with the study; seven panellists remained in the second phase of the study (four academics and three practitioners). The main goal of this phase was to understand the rating of importance of the factors, based on the different perspectives of various stakeholder groups.

All the new factors proposed by the participants had the mean >3, which makes them all significant to complexity of ID projects based on the opinion of the panellists.

In addition, IQR was >1 – consensus was not reached on the following statements:

- Number of different nationalities in ID projects influences project complexity;
- Political influence in ID projects influences project complexity;
- Stability of the project environment (in terms of exchange rates, material pricing etc.) in ID project influences project complexity;
- Interdependence among different stakeholders' influences project complexity;
- Chaos (4th type of uncertainty) influences project complexity.

T-tests measure the difference between two groups of panellists. In the first round, they showed difference on seven answers, and in the second round there is no difference in the answers.

Project coordinators thought that the number of goals, compatibility of different project management methods and tools, as well as the size in CAPEX and required local content in ID projects had more significant effect on project complexity than academics did.

Academics find that variety of stakeholder's perspectives, dependencies on other stakeholders and the political influence had a more significant effect on project complexity than project coordinators did.

Table 3 Wilcoxon test

	Number of goals	Goal alignment	Clarity of goals	Number of tasks	Dependencies between tasks	Conflicting norms and standards	Newness of technology (world-wide)	Experience with technology	Resource and skills availability	Interfaces between different disciplines	Number of different nationalities
Z	-1.300[b]	-1.13[b]	-1.134[c]	-1.633[c]	-1.000[c]	-.816[b]	-0.333[c]	-1.000[c]	-0.577[c]	-1.604[c]	-0.736[c]
Asymp. Sig. (2-tailed)	0.194	0.257	0.257	0.102	0.317	0.414	0.739	0.317	0.564	0.109	0.461

[a]Wilcoxon signed ranks test
[b]Based on positive ranks
[c]Based on negative ranks

Wilcoxon matched-pairs signed-ranks test measures changes in consensus between first and second round of Delphi study. Wilcoxon test has the purpose of measuring stability of the data and helping researchers determine if there was a difference between the data of two Delphi rounds.

All the concepts reached stability (the significance level was set at .05) and thus the Delphi was terminated with two rounds. For 11 questions that were repeated in both rounds, there was no significant statistical difference (Table 3).

4.4 Ranking the list of important factors

The descending order of the top ten weighted measures were found to be Clarity of goals (4.57), Variety of stakeholders' perspectives (4.45), Dependencies between tasks (4.43), Interfaces between different disciplines (4,29), Dependencies on other stakeholders (4.27), Risks from environment in ID project (4.27), Lack of clarity or consensuses on project benefits among project stakeholders (4.14), Unforeseen Uncertainty (4.14), Political influence (4.14), Number of stakeholders (internal and external) (4.09); for all the means see Appendix B.

Most of the factors with the mean higher than three are in the group of Environmental complexity. Based on the research, it is concluded that Environmental complexity has the most significant effect on the composite complexity, in comparison to technical and organizational complexity contribution. It might be concluded that experts think that environmental complexity is the most important of three types of complexity, which is the main contribution of the paper. In addition, it was found that uncertainty significantly influences complexity, based on the new factors proposed by the experts. Four uncertainty types were proposed, based on the classification of Meyer *et al*. – variation, foreseen uncertainty, unforeseen uncertainty and chaos. Particularities of different types of uncertainty require different managerial approach: "Projects in which variation and foreseen uncertainty dominate allow more planning, whereas projects with high levels of unforeseen uncertainty and chaos require a greater emphasis on learning."

Novel, modified complexity framework was developed based on the insights of the Delphi study (Fig. 2).

4.5 Implications of the study results

In the increasingly complex and unpredictable environment of ID projects, understanding complexity is becoming very important for planning, managing and executing strategies. Organizations that are delivering ID projects need to understand and adapt to these changes and include constant feedback from all the stakeholders in all the project phases.

One of the uses of the given complexity framework is creating awareness amongst the different stakeholders about the complexity on the project. Additionally, the framework could be used to access complexity on the project in different project phases and react in accordance to it.

Project managers of ID projects are usually not enough equipped to adequately handle complex projects since they base on their managerial style on traditional project management tools and techniques. Understanding the potential complexities can be a first step to making better strategies to manage ID projects.

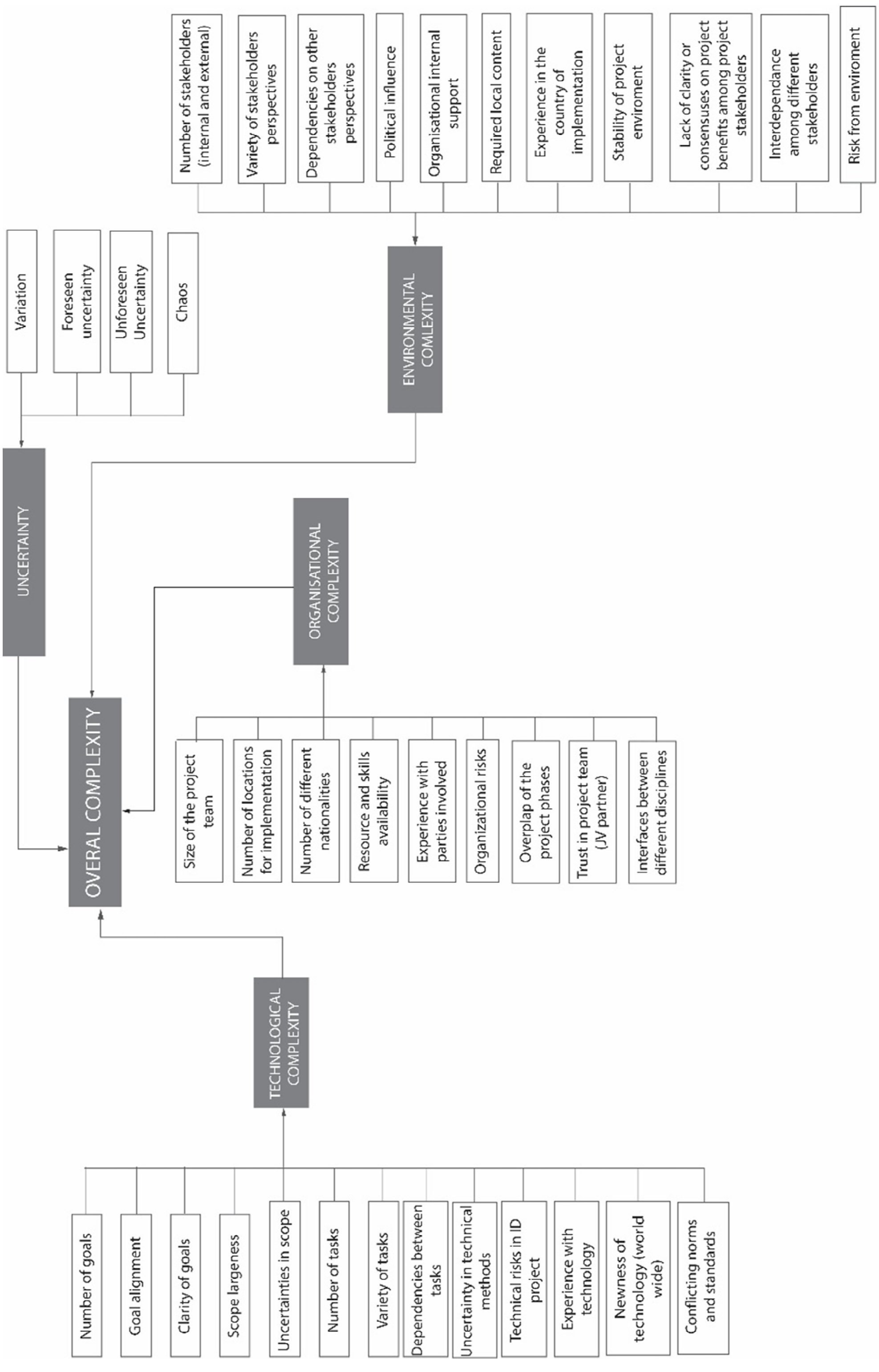


Fig. 2 Modified TOE framework

4.6 Limitations of the study

One of the limitations of the study could be sample size and response rate in the second round of the Delphi study. Future research could widen the results of this study by conducting interviews with non-responders. It would also be useful to investigate opinions from different stakeholders in International Development projects (except from project coordinators) about their perceptions on important complexity factors in ID projects.

The Delphi approach has many drawbacks, including subjective nature of the research, that can sometimes lead to biases. Additionally, method can be considered vulnerable to misrepresentation. The accuracy and reliability of the study are based on subjective judgment of the panellists. In this paper, biases were reduced by strict procedures, iteration, controlled feedback etc. For the validation of the given framework, it would be worthwhile conducting quantitative study that would objectively investigate refined list of complexity factors, in the context of different types of projects.

Additionally, Delphi approach could be supplemented with some other qualitative methodologies like Issues Management or Analytical Hierarchical Process, that would provide greater efficiency of the research.

Lastly, it would be worthwhile investigating how the selected complexity factors influence overall project performance of ID projects, when it comes to quality, time and costs, as well as how to manage the complexity elements in order to increase the chance of project success. In this context, it could be worthwhile investigating what are the competencies of project managers they need to work on complex ID projects.

5. Conclusion

A two-round Delphi survey was conducted to identify which complexity factors have the greatest influence in ID projects. A novel, modified TOE framework was proposed to access complexity in the domain of ID projects. This article is a first step in bridging the knowledge gap toward the development of a theoretically grounded and empirically validated framework of project complexity in ID context.

Inductive approach was used in the combination with literature review. Insights from the two-round Delphi resulted in the modified complexity framework. In total, 37 elements of complexity were identified and grouped in technical, organizational and environmental complexity group (Figure 2). Additional separate group was proposed by the experts – uncertainty category. The major contribution of this paper is investigation of complexity factors in ID sector, which was conducted for the first time. Organizational complexity was usually found to be the most significant type of complexity in previous researches conducted on large infrastructure projects. The Delphi study showed that, from the perspective of experts, Environmental factors were the most important in ID projects. Implications of this for organizations that are implementing ID projects should be to periodically review project objectives and to match project adaptability to the environment. On the other hand, the research methodology proposed in this study can be replicated to the other International Development projects to quantify different kinds of project complexity for improving the decision making and improving their execution performance.

The insights of the study can be used by both the academics and the practitioners. The framework could be used to assess the complexity of the ID project. Existing framework has an objective to contribute to better understanding of project complexity in the context of international development, and identification of complexity areas in specific projects, that could lead to better management of potential risks, as well as improvements in the process of project planning and implementation.

One of the major limitations of the study could be overcome by investigating the framework by employing quantitative approach. In addition, project coordinator's competences that match complexity type could be furtherly investigated in the future.

Acknowledgement

The authors acknowledge the financial support from the Slovenian Research Agency (Research Core Funding No. P2-0190).

References

[1] Ika, L.A., Söderlund, J., Munro, L.T., Landoni, P. (2018). Special issue: When project management meets international development, what can we learn?, *International Journal of Project Management*, Vol. 36, No. 2, 331-333, doi: 10.1016/j.ijproman.2017.05.005.

[2] Flyvbjerg, B. (2014). What you should know about megaprojects and why: An overview, *Project Management Journal,* Vol. 45, No. 2, 6-19, doi: 10.1002/pmj.21409.

[3] Ika, L.A., Diallo, A., Thuillier D. (2012). Critical success factors for World Bank projects: An empirical investigation, *International* Journal *of Project Management*, Vol. 30, No. 1, 105-116, doi: 10.1016/j.ijproman. 2011.03.005.

[4] Ika, L.A., Hodgson, D. (2014). Learning from international development projects: Blending critical project studies and critical development studies, *International Journal of Project Management*, Vol. 32, No. 7, 1182-1196, doi: 10.1016/j.ijproman.2014.01.004.

[5] Standish Group, Chaos summary (2009), from *https://www.projectsmart.co.uk/white-papers/chaos-report.pdf*, accessed October 10, 2018.

[6] Brady, T., Davies, A. (2014). Managing structural and dynamic complexity: A tale of two projects, *Project Management Journal*, Vol. 45, No. 4, 21-38, doi: 10.1002/pmj.21434.

[7] Cicmil, S., Cooke-Davies, T., Crawford, L., Richardson, K.A. (2009). *Exploring the complexity of projects: Implications of complexity theory for project management practice*, Project Management Institute, Newtown Square, Pennsylvania, USA.

[8] Geraldi, J., Maylor, H., Williams, T. (2011). Now, let's make it really complex (complicated): A systematic review of the complexities of projects, *International Journal of Operations & Production Management*, Vol. 31, No. 9, 966-990, doi: 10.1108/01443571111165848.

[9] Vidal, L.-A., Marle, F. (2008). Understanding project complexity: Implications on project management, *Kybernetes*, Vol. 37, No. 8, 1094-1110, doi: 10.1108/03684920810884928.

[10] Saynisch, M. (2010). Beyond frontiers of traditional project management: An approach to evolutionary, self-organizational principles and the complexity theory – Results of the research program, *Project Management Journal.*, Vol. 41, No. 2, 21-37, doi: 10.1002/pmj.20159.

[11] Kiridena, S., Sense, A. (2016). Profiling project complexity: Insights from complexity science and project management literature, *Project Management Journal.*, Vol. 47, No. 6, 56-74, doi: 10.1177/875697281604700605.

[12] Williamson O.E., Winter, S.G. (1993). The nature of the firm: origins, evolution, and development, Oxford University Press, UK.

[13] Winter, M., Smith, C., Morris, P., Cicmil, S. (2006) Directions for future research in project management: The main findings of a UK government-funded research network, *International Journal of Project Management*, Vol. 24, No. 8, 638-649, doi: 10.1016/j.ijproman.2006.08.009.

[14] He, Q., Luo, L., Hu, Y., Chan, A.P.C. (2015). Measuring the complexity of mega construction projects in China – A fuzzy analytic network process analysis, *International Journal of Project Management*, Vol. 33, No. 3, 549-563, doi: 10.1016/j.ijproman.2014.07.009.

[15] Geraldi, J., Adlbrecht, G. (2007). On faith, fact, and interaction in projects, *IEEE Engineering Management Review,* Vol. 36, No. 2, 35-49, doi: 10.1109/emr.2008.4534318.

[16] Cooke-Davies, T., Cicmil, S., Crawford, L., Richardson, K. (2007). We're not in Kansas anymore, Toto: Mapping the strange landscape of complexity theory, and its relationship to project mangement, *Project Management Journal*, Vol. 38, No. 2, 50-61, doi: 10.1177/875697280703800206.

[17] Bosch-Rekveldt, M., Jongkind, Y., Mooi, H., Bakker, H., Verbraeck, A. (2011). Grasping project complexity in large engineering projects: The TOE (technical, organizational and environmental) framework, *International Journal of Project Management*, Vol. 29, No. 6, 728-739, doi: 10.1016/j.ijproman.2010.07.008.

[18] Ika, L.A., Donnelly, J. (2017). Success conditions for international development capacity building projects, *International Journal of Project Management*, Vol. 35, No. 1, 44-63, doi: 10.1016/j.ijproman.2016.10.005.

[19] Golini, R., Landoni, P. (2013). *International development projects: Peculiarities and managerial approaches*, Project Management Institute, USA.

[20] Cicmil, S., Williams, T., Thomas, J., Hodgson, D. (2006). Rethinking project management: Researching the actuality of projects, *International Journal of Project Management*, Vol. 24, No. 8, 675-686, doi: 10.1016/ j.ijproman.2006.08.006.

[21] Baccarini, D. (1996). The concept of project complexity – A review, *International Journal of Project Management*, Vol. 14, No. 4, 201-204, doi: 10.1016/0263-7863(95)00093-3.

[22] Williams, T.M. (1999). The need for new paradigms for complex projects, *International Journal of Project Management*, Vol. 17, No. 5, 269-273, doi: 10.1016/s0263-7863(98)00047-7.

[23] Sommer, S.C., Loch, C.H. (2004). Selectionism and learning in projects with complexity and unforeseeable uncertainty, *Management Science*, Vol. 50, No. 10, 1334-1347, doi: 10.1287/mnsc.1040.0274.

[24] Vidal, L.-A., Marle, F., Bocquet, J.-C. (2011). Measuring project complexity using the analytic hierarchy process, *International Journal of Project Management*, Vol. 29, No. 6, 718-727, doi: 10.1016/j.ijproman.2010.07.005.

[25] Bakhshi, J., Ireland, V., Gorod, A. (2016). Clarifying the project complexity construct: Past, present and future, *International Journal of Project Management*, Vol. 34, No. 7, 1199-1213, doi: 10.1016/j.ijproman.2016.06.002.

[26] Lu, Y., Luo, L., Wang, H., Le, Y., Shi, Q. (2015). Measurement model of project complexity for large-scale projects from task and organization perspective, *International Journal of Project Management*, Vol. 33, No. 3, 610-622, doi: 10.1016/j.ijproman.2014.12.005.

[27] Xia, B., Chan, A.P.C. (2012). Measuring complexity for building projects: A Delphi study, *Engineering, Construction and Architectural Management*, Vol. 19, No. 1, 7-24, doi: 10.1108/09699981211192544.
[28] Lessard, D., Sakhrani, V., Miller, R. (2014). House of project complexity – Understanding complexity in large infrastructure projects, *Engineering Project Organization Journal*, Vol. 4, No. 4, 170-192, doi: 10.1080/21573727.2014.907151.
[29] Williams, T. (2003). *Modelling complex projects*, John Wiley & Sons, Chichester, UK.
[30] Flick, U. (2018). *An introduction to qualitative research*, Sixth edition, Sage, London, UK.
[31] Chan, A.P.C., Yung, E.H.K., Lam, P.T.I., Tam, C.M., Cheung, S.O. (2001). Application of Delphi method in selection of procurement systems for construction projects, *Construction Management and Economics*, Vol. 19, No. 7, 699-718, doi: 10.1080/01446190110066128.
[32] Delbecq, A.L., Van de Ven, A.H., Gustafson, D.H. (1975). *Group techniques for program planning: A guide to nominal group and Delphi processes*, Scott Foresman Company, Glenview, Illinois, USA.
[33] Linstone, H.A., Turoff, M. (2002). *The Delphi method: Techniques and applications*, Addison-Wesley, USA.
[34] Liu, S., Zhang, J., Keil, M., Chen, T. (2010). Comparing senior executive and project manager perceptions of IT project risk: A Chinese Delphi study, *Information Systems Journal,* Vol. 20, No. 4, 319-355, doi: 10.1111/j.1365-2575.2009.00333.x.
[35] Hadaya, P., Cassivi, L., Chalabi, C. (2012). IT project management resources and capabilities: A Delphi study, *International Journal of Managing Projects in Business,* Vol. 5, No. 2, 216-229, doi: 10.1108/17538371211214914.
[36] Perera, B.A.K.S., Rameezdeen, R., Chileshe, N., Hosseini, M.R. (2014). Enhancing the effectiveness of risk management practices in Sri Lankan road construction projects: A Delphi approach, *International Journal of Construction Management,* Vol. 14, No. 1, 1-14, doi: 10.1080/15623599.2013.875271.
[37] Lang, T. (1998). An overview of four futures methodologies (Delphi, environmental scanning, issues management and emerging issue analysis), *The Manoa Journal of Fried and Half-Fried Ideas (About the future)*, Hawaii Research Center for Futures Studies, Honolulu, from *http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.455.1045*, accessed October 8, 2018.
[38] Okoli, C., Pawlowski, S.D. (2004). The Delphi method as a research tool: An example, design considerations and applications, *Information & Management*, Vol. 42, No. 1, 15-29, doi: 10.1016/j.im.2003.11.002.
[39] Hallowell, M.R. (2009). Techniques to minimize bias when using the Delphi method to quantify construction safety and health risks, In: Ariaratnam, S.T., Rojas, E.M., (eds.), *Building a Sustainable Future,* American Society of Civil Engineers, Seattle, Washington, USA, 1489-1498, doi: 10.1061/41020(339)151.
[40] Schmidt, R.C. (1997). Managing Delphi surveys using nonparametric statistical techniques, *Decision Sciences*, Vol. 28, No. 3, 763-774, doi: 10.1111/j.1540-5915.1997.tb01330.x.
[41] Von der Gracht, H.A. (2012). Consensus measurement in Delphi studies: Review and implications for future quality assurance, *Technological Forecasting and Social Change*, Vol. 79, No. 8, 1525-1536, doi: 10.1016/j.techfore.2012.04.013.

Appendix A

Complexity measuring in the literature

No.	FACTOR	SUB-FACTOR	ELEMENTS	SUB-ELEMENTS	SOURCE
1.	STUCTURAL COMPLEXITY	TECHNOLOGICAL	Goals	- Number of goals - Goal alignment - Clarity of goals	Lu *et al.* (2015); Brady and Davies (2014); Senescu, Aranda-Mena, And Haymaker (2013); Geraldi *et al.* (2011); Geraldi (2009); Haas (2009); Whitty and Maylor (2009); Maylor *et al.* (2008); Vidal and Marle (2008); Geraldi and Adlbrecht (2007); Remington and Pollack (2007); Ives (2005); Williams (2005); Xia and Lee (2004); Williams (1999); Baccarini (1996); Shenhar and Dvir (1996)
			Scope	- Scope largeness - Quality requirements	
			Tasks	- Number of tasks - Variety of tasks - Dependencies between tasks - Interrelations between technical processes - Conflicting norms and standards	
			Experience	- Newness of technology - Experience with technology	
			Risk	- Technical risks	
		ORGANISATIONAL COMPLEXITY	Size	- Project duration	
			Resources	- Compatibility of different project management methods and tools - Size in budget - Size in Engineering hours - Size of project team - Number of locations included	
			Resources	- Resources and skills availability - Experience with parties involved - Interfaces between different disciplines - Number of financial resources - Contract types	
			Team	- Number of different nationalities - Number of different languages - Cooperation JV partner - Overlapping office hours	
			Trust	- Trust in project team - Trust in the contractor	
			Risk	- Organizational risks	
		ENVIRONMENTAL COMPLEXITY	Stakeholders	- Number of stakeholders - Variety of stakeholders' perspectives - Dependencies on other stakeholders - Political influence - Company (*implementing body*) internal support	
			Location	- Interference with existing site – N/A - Weather conditions – N/A - Remoteness of location - Experience in the country	
			Market conditions	N/A	
			Risk	- Risks from environment	
2.	DYNAMICAL COMPLEXITY			- Changes in all the elements that consist structural complexity	Kiridena et. al (2016)
3.	SOCIO-POLITICAL COMPLEXITY				Beach (2016); Geraldi *et al.* (2011); Maylor *et al.* (2008); Bresnen *et al.* (2005); Cicmil and Marshall (2005); Ives (2005); Shenhar and Dvir (1996); Jones and Deckro (1993)
4.	UNCERTAINTY	STRUC TURAL	Uncertainties in scope		(Shenhar, 2001; Tatikonda and Rosenthal, 2000, Maylor *et al.*, 2008; Mykytyn and Green, 1992) Geraldi and Adlbrecht, 2007; Hobday, 1998;
		DYNAMICAL	Change in elements	- Changes in scope, deviations	Geraldi and Adlbrecht, 2007; Collyer and Warren, 2009; Petit, 2012; Chapman, 2003; Atkinson *et al.*, 2006; Bosch-Rekveldt *et al.*,2011; Maylor *et al.*, 2013; Saun-ders *et al.*, 2015
			External uncertainty	- External contexts - Unclear organizational context - External elements - External political influence	
5.	PACE				Geraldi *et al.* 2011; Dvir, *et al.*, 2006; Shenhar and Dvir, 2007; Williams, 2005
6.	PROJECT SUCCESS	SHORT-TERM	Time		Ika, L. A., Diallo, A., & Thuillier, D. (2012)
			Cost		
			Objectives		
		LONG-TERM	Impact		
			Sustainability		
			Relevance		

Appendix B

Complexity factors in international development projects: Mean values

FACTOR	MEAN	Type of complexity
Number of goals in ID projects affects project complexity	3.86	TECH
Goal alignment in ID projects influences project complexity	3.43	TECH
Clarity of goals in ID projects influence project complexity	4.57	TECH
Scope largeness in ID projects influences project complexity	3.27	TECH
Uncertainties in scope in ID projects influence project complexity	4.06	TECH
Quality requirements in ID projects influence project complexity	2.73	TECH
Number of tasks in ID projects influences project complexity	3.86	TECH
Variety of tasks in ID projects influences project complexity	3.45	TECH
Dependencies between tasks in ID projects influence project complexity	4.43	TECH
Uncertainty in technical methods to be applied in ID projects influences project complexity	3.73	TECH
Conflicting norms and standards in ID projects influence project complexity	3.57	TECH
Newness of technology (world-wide) in ID projects influences project complexity	3.86	TECH
Experience with technology in ID projects influences project complexity	3.71	TECH
Technical risks in ID projects influence project complexity	3.27	TECH
Duration of ID projects influences project complexity	2.82	ORG
Compatibility of different project management methods and tools in ID projects influences project complexity	2.73	ORG
Size in CAPEX influences project complexity	2.55	ORG
Size of the project team in ID projects influences project complexity	3.64	ORG
Number of locations for implementation of ID projects influences project complexity	3.91	ORG
Resource and skills availability in ID projects influence project complexity	4.00	ORG
Experience with parties involved in ID projects influences project complexity	3.64	ORG
Interfaces between different disciplines in ID projects influence project complexity	4.29	ORG
Number of different nationalities in ID projects influences project complexity	3.14	ORG
Number of different languages in ID projects influence project complexity	2.73	ORG
Trust in project team (JV partner) in ID projects influences project complexity	4.00	ORG
Organisational risks in ID projects influence project complexity	3.27	ORG
Number of stakeholders (internal and external) in ID projects influences project complexity	4.09	ENV
Variety of stakeholders' perspectives in ID projects influences project complexity	4.45	ENV
Dependencies on other stakeholders in ID projects influence project complexity	4.27	ENV
Political influence in ID projects influences project complexity	4.14	ENV
Organisational internal support in ID projects influences project complexity	3.55	ENV
Required local content in ID projects influences project complexity	3.00	ENV
Experience in the country of implementation of ID project influences project complexity	3.71	ENV
Stability of project environment (exchange rates, material pricing) in ID project influences project complexity	3.71	ENV
Risks from environment in ID project influence project complexity	4.27	ENV
Overlap of the project phases influences project complexity	3.43	ORG
Interdependence among different stakeholders	3.71	ENV
Diversity of stakeholder expectations	3.86	ENV
Lack of clarity or consensuses on project benefits among project stakeholders	4.14	ENV
Variation (1st type of uncertainty)	3.57	UNC
Foreseen uncertainty (2nd type of uncertainty)	3.71	UNC
Unforeseen uncertainty (3rd type of uncertainty)	4.14	UNC
Chaos (4th type of uncertainty)	3.86	UNC

APEM journal

Advances in Production Engineering & Management
Volume 14 | Number 2 | June 2019 | pp 239–250
https://doi.org/10.14743/apem2019.2.325

ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

The investment strategy and capacity portfolio optimization in the supply chain with spillover effect based on artificial fish swarm algorithm

Zheng, Z.L.[a], **Bao, X.**[b,*]

[a]College of Economic and Management, China Agricultural University, Beijing, P.R. China
[b]College of Business Administration, Zhejiang University of Finance and Economics, Hangzhou, P.R. China

ABSTRACT

Spillover effect can lead to the free-riding behavior when joint investment takes place in the supply chain. This study examined the investment strategies of two competitive retailers who considered whether to invest a shared contract manufacturer (CM) or not. The supply chain members' operational decisions in four scenarios were analyzed through a Cournot competition model, and the paths of the retailers' investment strategies were examined. The CM's capacity portfolio optimization was NP-hard in nature, and was modelled by an investment portfolio problem. Results show that both retailers jointly invest the CM only when the difference of production costs is not high, and the intentions of joint investment will decrease when the coefficient of spillover and the degree of substitutability between products increase. The CM always benefits as long as one retailer invests, and allocates more investment on the capacity with highest revenue when he emphasizes more on the profit. For optimizing the CM's capacity portfolio problem, an artificial fish swam algorithm with uniform mutation (AFSA_UM) is developed and it shows better convergent performance and higher robustness than the basic AFSA.

ARTICLE INFO

Keywords:
Supply chain;
Investment strategy;
Capacity portfolio;
Spillover;
Artificial fish swam algorithm

**Corresponding author:*
goldbxing@outlook.com
(Bao, X.)

Article history:
Received 22 February 2019
Revised 25 April 2019
Accepted 2 May 2019

1. Introduction

Joint investment has been frequently witnessed in the supply chain, where several members (e.g., retailers in the downstream) might invest to a shared partner (e.g., a manufacturer in the upstream). Investment could vary in different kinds, such as financial injection, advertisement subsidy, investment on the most up-to-date equipment, and cooperation on the R&D of new product etc. [1, 2]. Joint investment could benefit all members in the supply chain, such as lowering the production costs, increasing the market share by providing lower retailing prices to the customers, thus performing better than the price-based coordination [3].

Despite the merits of joint investment, there are two questions for both the investor and the investee in the supply chain. For the investor, she could be surprised to find that her effort in some activities could benefit her competitors in the same industries. For example, in 2012, Intel and Samsung jointly provided $4.4 billion to ASML (the largest manufacturer of wafer steppers and scanners for microchip fabrication in the world) through a "Customer Co-Investment Program" to accelerate a 450 mm technology and a next-generation EUV development project. However, the ASML claimed that the results of development projects and capacity investment would be available to every semiconductor company with no restrictions. For the ASML (the investee), he has to consider the loss brought by the obsolete risk of production capacity. For example, the capacity to produce 450 mm chips might be outdated quickly due to the change of market demand. When offered several kinds of capacity with different production rate and obsolete loss, the AMSL has to carefully determine the proportion of each capacity by balancing the revenue and loss with the consideration of budget constraint.

This study is motivated by the ASML case. In this study, Intel and Samsung are referred as the retailers, who have the incentives to invest the ASML or free ride the competitor's investment. The ASML is referred as the shared contract manufacturer (CM) who is at the upstream of the supply chain and optimizes the capacity portfolio after receiving the retailers' investment. There are two questions worthy of investigation: (1) What the retailers' investment decisions will be when spillover effect exists? Jointly invest, invests alone or free rides? (2) How could the CM optimize the capacity portfolio when given multiple production capacity with different combinations of production output (equivalently, revenue) and obsolete risk (equivalently, loss)? With these two questions in mind, this study considers a supply chain consisted of two competitive retailers and a shared CM. In this supply chain, both retailers compete to sell the products with a certain degree of substitutability in the market, either retailer could free ride the other one's investment in reducing the production cost due to the spillover effect. Once the CM receives the investment, he should appropriately spilt the retailers' total amount of investment on multiple production capacities for maximizing his revenue while minimizing the loss.

The rest of this study is organized as follows. In section 2, the relevant literatures are reviewed. In section 3, the supply chain members' decisions in four scenarios are firstly analysed under the Cournot model, and then the paths of retailers' investment decisions are investigated. The CM's capacity portfolio problem is described through a classic investment portfolio model, and an artificial fish swam algorithm (AFSA) improved by the uniform mutation is developed to optimize the problem. In section 4, numerical studies are carried out to investigate the managerial insights. In section 5, the conclusions and future researches are given.

2. State of the art

In recent decades, many researchers have carried out voluminous studies on the investment in the field of operation management (OM) [4]. The literatures related to this study can be categorized into the following two streams:

The first stream of studies mainly focused on the influence of spillover effect on the players' operational decisions when they compete in the same market. Due to the spillover effect, the player who did not invest could have the opportunity to free ride the outcome of the other one's investment on some activities (e.g., R&D and cost reduction). Actually, the spillover effect was so prevailing in the R&D investment that many researchers paid their attentions in this field in the past decade. Those who were interested in this topic, we refer the studies given by [5-8]. In recent decade, researchers found that the spillover effect also existed in the operational decisions in the supply chain field. For example, the spillovers in process knowledge increased the likelihood of observing decentralized channel structures under some conditions (e.g., spillovers were involuntary, firms' innovative activities were non-overlapping, and firms benefited directly from the results of competitors' innovations) [9]. In a supply chain where an original equipment manufacturer (OEM) and a CM competed in the finished goods market, the technology spillovers could strengthen the OEM's incentive to strategically outsource the production to the CM [10]. Under a wholesale price contract, both firms in the supply chain could achieve win-win via cartelization in R&D only if their contribution levels were Pareto matched (e.g., when each firm's contribution level was comparable to its partner's even when spillover existed) [11]. However, the degree of competition might change the players' operational decisions. For example, the manufacturer's improvement effort usually declined in market competition, market uncertainty or spillover effect, although its expected profit typically increased in spillover effect [12]. In the situation when two rival firms' operations and technology managers were given bonuses for cost reduction, the prisoner's dilemma occurred in case spillovers were less than 50%, or when spillovers were higher and process improvement capability was relatively high [13]. In the situation when two competing firms invested in a shared supplier, the spillover actually discouraged firms' investment despite that it supposedly intensified competition [14]. However, in the work of [15], after investigating the effect of learning, spillover and competition in affecting the optimal strategies for two firms to invest in their shared suppliers, the investment strategies of two firms were characterized by a region of preemption and a region of war of attrition.

The second stream focused on the optimization of capacity portfolio problem, which was a branch of capacity management and concerned with specifying the amounts or locations for the given multiple capacity types [16]. For some capacity-intensive industries (e.g., semiconductor manufacturing), capacity portfolio planning could significantly affect the capacity effectiveness and final profit via forecasting various demands of products [17]. From the point of investment's view, balancing the production system's revenue (profit) and risk (loss) was the critical consideration of the decision maker who aimed to optimize the capacity portfolio under different settings of environments [18]. However, it was difficult to optimize the capacity portfolio, because it was NP hard in nature to determine the sizes or proportions among different types of capacity with budget constraints. Some researchers modelled the capacity portfolio problems by various kinds of mixed-integer problems. For example, in [19], the capacity portfolio problem was formulated a two-stage stochastic mixed-integer model when the production system faced multiple demands with associated probability, then a robust stochastic programming approach was proposed to solve problem. Similarly, a mixed integer linear programming was proposed in [18] to seek an optimal capacity allocation plan and capacity expansion policy under single-stage, multi-generation, and multi-site structures. Some researchers modelled the capacity portfolio problem as a Markov decision process, and solved the optimal investment levels on two types of capacity (mainly the dedicated and flexible capacity) through efficient dynamical programming [20] or heuristic algorithm [21], or robust optimization [22, 23]. When there were multiple capacities (such as dedicated, flexible and reconfigurable capacity) to be invested, the capacity portfolio problem turned to be a mixed integer programming model [24], and hence, heuristic method could be adopted to obtain the optimal solution.

Most published studies on the spillover effect in the supply chain field focused on the investment decisions, supplier's reliability, or increasing the quality of product from the perspective of the investors (e.g., retailers in this study). Few studies paid attention to the spillover effect could alter the investors' decisions on investing or free riding, and hence, the path of investors' investment strategies were lack of investigating. On the other hand, investee (e.g., the shared CM in this study) might optimize the capacity portfolio when given more than two types of capacity associated with different combination of revenue and loss. In this study, we adopt the game model to analytically show the investors' (e.g., the retailers) investment decisions and the path of changing investment decisions when considering the competitors' advantage or disadvantage. We model the investee's (e.g., the CM) capacity portfolio problem by the classic investment portfolio problem, which is solved through an artificial fish swam algorithm (AFSA) improved by the uniform mutation.

3. Methodology

3.1 Operational decisions in four scenarios

In this study, we consider a supply chain with two competitive retailers and a shared CM. Each retailer is intent to lower the production cost by investing or free riding the competitor's investment due to the spillover effect. To describe the retailers' competitive behaviour, we follow the assumption in [25] where the Retailer *i*'s retailing price, p_i could be given by

$$p_i(q_i, q_j) = a - q_i - bq_j, \quad i = 1, or\ 2;\ j = 3 - i \tag{1}$$

where a denotes the potential demand of two products in the market, q_i and q_j are the ordering quantities, b is the degree of substitutability between two products. b ranges between 0 and 1. The larger value of b is, the more substitutable of the products are to the customers, thus the high market competition is.

c_i denotes the initial production cost when the Retailer *i* does not invest the CM. For simplicity but without loss of generality, we assume $c_1 = c$ and $c_2 = c_1 + \delta$. δ is the difference of production costs between two products. $\delta > 0$ denotes that the Retailer 1 enjoys a cost advantage over Retailer 2 in production cost, vice versa.

We assume that the spillover effect takes place in production process. Because both retailers share the same CM, one retailer could have the opportunity of enjoying a lower production cost by free riding. C_i denotes the production cost after the retailers provide the investment, which is given by

$$C_i = c_i - x_i - \alpha x_j \tag{2}$$

where x_i and x_j are the levels of cost reduction that the Retailer *i* and *j* want to achieve by investing the CM, respectively. α is the coefficient of spillover. α ranges between 0 and 1. The larger value of α is, the higher spillover effect is, thus leading to Retailer *i*'s higher free-riding behavior. By the Eq. 2, C_i is the result of both retailers' joint investment decisions. Furthermore, we assume the Retailer *i*'s investment is quadratic in the level of production cost reduction [11, 26]. Due to the spillover effect, each retailer has an incentive to free ride the other one's investment. Thus, $\rho_i x_i^2$ is the Retailer *i*'s investment on the CM for achieving x_i. ρ_i is a parameter related to the investment and $\rho_i \geq 1$. The Retailer *i* does not invest any when $x_i = 0$.

When the Retailer *i* invests the CM, her profit is given by

$$\pi_i(q_i, x_i) = (p_i - w_i)q_i - \rho_i x_i^2 \tag{3}$$

where w_i is the wholesale price that CM sets for Retailer *i*. The CM's profit is shown as follow:

$$\pi_M(w_1, w_2) = \sum_{i=1}^{2} (w_i - C_i)q_i \tag{4}$$

We assume that the retailers compete under Cournot competition model, thus the retailer's ordering quantities are determined simultaneously. The retailing prices could be determined through the backward induction. Note that, we assume that all members in the supply chain have complete knowledge of the game participants. Consider the question of whether to invest the CM or not, each retailer has two options: Yes or No (denoted as Y or N). Therefore, there are four scenarios: the YY scenario where both retailers invest the CM; the YN and NY scenarios where only retailer invests the CM and the other free rides; and the NN scenario where no one invests the CM.

We start with the derivation of the retailers' and CM's equilibrium decisions in the YY scenario. As both retailers invest the CM, the Retailer *i*'s production cost is given by the Eq. 2. The Retailer *i* aims to maximize her profit in Eq. 3. Take first and second order derivatives of π_i with respect to q_i, we could have π_i is convex in q_i. Therefore, the Retailer *i*'s optimal ordering quantity is given by $q_i(q_j) = \frac{1}{2}(a - bq_j - w_i)$, which can be reformulated in below:

$$q_i(w_i, w_j) = \frac{a(2-b) - 2w_i + bw_j}{4 - b^2} \tag{5}$$

By inserting the Eq. 5 into Eq. 3, we have the Retailer *i*'s profit, $\pi_i(x_i) = q_i^2 - \rho_i x_i^2$.

Now, it is the CM's turn to optimize his profit. The CM aims to maximize his profit by considering the Retailer *i*'s optimal ordering quantity $q_i(w_i, w_j)$. The CM's profit is given by $max \pi_M = \sum_{i=1}^{2}(w_i - c_i - x_i - \alpha x_j) q_i(w_i, w_j)$. By taking the first order derivative of π_M with respect to w_i, we have the optimal wholesale price $w_i(x_i, x_j)$ for the Retailer *i*, which is given by $w_i(x_i, x_j) = \frac{a + c_i - x_i - \alpha x_j}{2}$. Insert $w_i(x_i, x_j)$ into Eq. 5, and we have the Retailer *i*'s profit, which is $\pi_i(x_i) = \left[\frac{a(2-b) - 2c_i + bc_j + x_i(2 - b\alpha) - x_i}{2(4-b^2)}\right]^2 - \rho_i x_i^2$.

By summarizing the results above, we have the proposition 1.

Proposition 1: In the YY scenario,

(1) To make sure that the Retailer i's ordering quantity is positive, the difference of production costs should be bounded, i.e., $\delta_1 \leq \delta \leq \delta_2$, *where* $\delta_1 = (a-c)(1-2/b)$ and $\delta_2 = (a-c)(1-b/2)$.

(2) $\frac{\partial q_i(x_i)}{\partial x_j} \leq 0$ *when* $\alpha \leq b/2$, *and* $\frac{\partial q_i(x_i)}{\partial x_j} \geq 0$ *when* $\alpha > b/2$.

(3) the Retailer i's profit is strictly concave in x_i. *Therefore, her optimal decision is* $x_i(x_j) = \frac{(2-\alpha b)[a(2-b)-2c_i+bc_j-x_j(b-2\alpha)]}{2\rho_i(4-b^2)^2-(2-\alpha b)^2}$ *for a given* x_j.

By Proposition 1, we could backwardly derive the retailers and CM's optimal decisions in the YY Scenario, which are given in Table 1.

The procedures of deriving the operational decisions in YN, NY and NN scenarios are same with those in YY scenario. Therefore, we omit the derivations for saving the pages. The retailers' and CM's operational decisions in four scenarios are summarized in Table 1.

Table 1 The Supply Chain's operational decisions in four scenarios

	YY Scenario	NY Scenario	YN Scenario	NN Scenario
x_1	$\frac{D_1(A_1B_2 - A_2D_1D_2)}{B_1B_2-(D_1D_2)^2}$	0	$\frac{A_1D_1}{B_1}$	0
x_2	$\frac{D_1(A_2B_1 - A_1D_1D_2)}{B_1B_2-(D_1D_2)^2}$	$\frac{A_2D_2}{B_2}$	0	0
p_1		$w_1 + q_1$		
p_2		$w_2 + q_2$		
q_1	$\frac{2(4-b^2)^2\rho_1(A_1B_2 - A_2D_1D_2)}{B_1B_2-(D_1D_2)^2}$	$\frac{(A_1B_2 - A_2D_1D_2)}{2(4-b^2)B_2}$	$2\rho_1(4-b^2)\frac{A_1}{B_1}$	$\frac{A_1}{2(4-b^2)}$
q_2	$\frac{2(4-b^2)^2\rho_2(A_2B_1 - A_1D_1D_2)}{B_1B_2-(D_1D_2)^2}$	$2\rho_2(4-b^2)\frac{A_2}{B_2}$	$\frac{(A_2B_1 - A_1D_1D_2)}{2(4-b^2)B_1}$	$\frac{A_2}{2(4-b^2)}$
w_1	$\frac{1}{2}\left\{a+c_1-\frac{D_1[A_2(B_1-\alpha D_1D_2)+A_1(\alpha B_1 - D_1D_2)]}{B_1B_2-(D_1D_2)^2}\right\}$	$\frac{1}{2}(a+c_1-\alpha x_2)$	$\frac{1}{2}(a+c_1-x_1)$	$\frac{1}{2}(a+c_1)$
w_2	$\frac{1}{2}\left\{a+c_2-\frac{D_1[A_1(B_2-\alpha D_1D_2)+A_2(\alpha B_1 - D_1D_2)]}{B_1B_2-(D_1D_2)^2}\right\}$	$\frac{1}{2}(a+c_2-x_2)$	$\frac{1}{2}(a+c_2-\alpha x_1)$	$\frac{1}{2}(a+c_2)$
π_1	$\frac{\rho_1B_1(A_1B_2 - A_2D_1D_2)^2}{[B_1B_2-(D_1D_2)^2]^2}$	$\frac{(A_1B_2 - A_2D_1D_2)^2}{[2(4-b^2)B_2]^2}$	$\frac{\rho_1A_1^2}{B_1}$	$\frac{A_1^2}{B_1^2}$
π_2	$\frac{\rho_2B_2(A_2B_1 - A_1D_1D_2)^2}{[B_1B_2-(D_1D_2)^2]^2}$	$\frac{\rho_2A_2^2}{B_2}$	$\frac{(A_2B_1 - A_1D_1D_2)^2}{[2(4-b^2)B_1]^2}$	$\frac{A_2^2}{B_1^2}$
π_M	$\sum_{i=1}^{2}(w_i - c_i + x_i + \alpha x_{3-i})q_i$			

3.2 Paths of the retailers' investment strategies

In this section, we start with the analysis of the path that the Retailer 1 changes her decision from NN to YN scenario. Apparently, Retailer 1 decides to invest the CM only when her profit increases in YN scenario compared with that in NN Scenario, i.e., $\pi_1^{YN} - \pi_1^{NN} > 0$ should hold. From Table 1, $\pi_1^{YN} - \pi_1^{NN} > 0$ holds only when $\delta > \delta_2$. Further, we have $q_1^{YN} - q_1^{NN} > 0$, which means the Retailer 1 will always order more products when she invests the CM.

We investigate the CM's benefit when Retailer 1 invests. From Table 1, we have $w_i^{YN} - w_I^{NN} < 0$, which means that the CM will lower the wholesale prices to both retailers as long as Retailer 1 invests. Let $\Delta P_{Mi}^{YN-NN} = (w_i^{YN} - C_i) - (w_i^{NN} - c_i)$ be the CM's difference of marginal profits when he wholesales the product to the Retailer *i* in YN and NN scenarios. We could prove that $\Delta P_{Mi}^{YN-NN} > 0$. Obviously, the CM's marginal profits will always increase when the Retailer 1 invests the CM. Further, the difference of the CM's profits, $\pi_M^{YN} - \pi_M^{NN}$ is strictly positive. Therefore, the CM's profit always increases as long as one retailer invests.

Corollary 1 summarizes the results above.

Corollary 1: *Compared with the operational decisions in the NN* scenario*, in the YN scenario where the Retailer 1 invests the CM,*

(1) The CM offers lower wholesale prices for both retailers when only the Retailer 1 provides the investment. He could increase the marginal and total profits concurrently.
(2) The Retailer 1 always orders more in YN scenario than she does in NN scenario, and she enjoys a higher profit only when $\delta > \delta_1$; otherwise, she will not invest the CM.

Corollary 1 shows the path of the Retailer 1's investment strategy from NN to YN scenario. We find that, in YN scenario, the Retailer 1 could benefit in two folds: (i) she orders (also sells) more products, enjoys higher market share, thus deterring the competition from the Retailer 2. (ii) She could have higher profit by investment.

We could derive the path of Retailer 2 when she change her investment decision from NN to NY scenario in the same way. For saving pages, we omit the derivation and give the results in Corollary 2.

Corollary 2: *Compared with the equilibrium decisions in the NN scenario, in the NY scenario where only the Retailer 2 invests the CM,*

(1) The CM offers lower wholesale prices for both retailers. He could increase the marginal and total profit concurrently.
(2) The Retailer 2 will always orders more in YN scenario than she does in NN scenario, and she invests the CM when $\delta < \delta_2$; otherwise, she will not invest the CM.

Corollary 2 shows that the Retailer 2 changes her investment strategy from NN scenario to NY scenario under two situations: (i) the Retailer 1's cost advantage is not high, and (ii) the Retailer 1 is obvious cost disadvantageous. Both situations should satisfy $\delta < \delta_2$.

Since both retailers compete in the market, no one would like her investment to be free rode by the other one. Therefore, joint investment would be the ideal equilibrium state for both retailers. To investigate the dynamic path by which one retailer quits free riding and enters the equilibrium state of YY scenario, we compare the equilibrium decisions in NY (YN) and YY scenarios.

To make sure that the Retailer 1 stays in YY scenario, her profit in YY scenario should be more than that in NY scenario (i.e., $\pi_1^{YY} - \pi_1^{NY} > 0$ should hold). From Table 1, $\pi_1^{YY} - \pi_1^{NY} > 0$ holds when $\delta > \frac{(a-c)(2-b)(B_2-D_1D_2)}{bB_2+2D_1D_2}$. $q_1^{YY} - q_1^{NY} > 0$ always holds.

Corollary 3 summarizes the results above.

Corollary 3: *The Retailer 1 enters into the equilibrium state of YY scenario from NY scenario only when $\delta > \delta_3$, where $\delta_3 = \frac{(a-c)(2-b)(B_2-D_1D_2)}{bB_2+2D_1D_2}$. In this case, the Retailer 1 will not only increase the profit, but also increase the ordering quantities.*

Corollary 3 indicates that the Retailer 1 will jointly invest the CM with Retailer 2 only when her cost advantage is above a threshold over the Retailer 2's production cost (i.e., $\delta > \delta_3$).

Similarly, we could analyse the path of the Retailer 2 to jointly invest the CM with Retailer 1. For saving the pages, we only presents the results in Corollary 4.

Corollary 4: *The Retailer 2 will quit the equilibrium state of YN scenario and enter into the equilibrium state of YY scenario only when $\delta < \delta_4$, where $\delta_4 = \frac{(a-c)(2-b)(B_1-D_1D_2)}{2B_2+bD_1D_2}$. When compared with the YN scenario, the Retailer 2 will increase her profit, but increase her ordering quantities only when $\alpha > b/2$.*

Corollary 4 indicates that the Retailer 2 will jointly invest the CM with Retailer 1 only when the Retailer 1's cost advantage should not be above a threshold over her production cost (i.e., $\delta < \delta_4$).

3.3 The CM's capacity portfolio optimization based on AFSA_UM

Shown by Corollary 1 to 4, the CM always benefits as long as one retailer invests. However, the CM still has to carefully split the retailers' investment on different types of capacity before launching the production. The reasons is that each capacity varies in the production output and obsolete risk, thus the potential revenue and loss is different to each capacity.

For simplicity without loss of generality, we assume that the total amount of retailers' investment to be one, and the CM allocates the investment on n types of capacity $(S_j, j = 1, \dots, n)$ with different proportion (y_j). For the capacity S_j, the potential revenue and loss are r_j and u_j. The total revenue of CM is given by $V = \sum_{j=1}^{n} y_j r_j$, and the maximal risk of CM's investment allocation is given by $U = max_{1 \le j \le n}\, y_j u_j$. Therefore, the CM's objective in allocating the investment turns into a classic portfolio optimization problem, which is maximizing the total revenue while minimizing the maximal risk. According to [27], the CM's capacity portfolio optimization problem could be modelled by

$$maxF = \beta V + (1 - \beta)U, \qquad s.t. \sum_{j=1}^{n} y_j = 1, y_j \ge 0, j = 1, \dots, n \tag{6}$$

where β is the CM's attitude on revenue, and $0 < \beta < 1$. By introducing β, the multi-objective capacity portfolio optimization problem in Eq. 6 turns into single-objective problem.

Unfortunately, the optimization of CM's capacity portfolio given in Eq. 6 is NP hard in nature. It is rather time-consuming to optimize the problem when the numbers of decision variables and constraints increase. In this study, we develop an improved AFSA (named AFSA_MU) to optimize the CM's capacity portfolio.

The AFSA and its variants are widely used in the optimizations in the OM field [28-30]. The basic principle AFSA relies on the phenomenon that a fish can discover the more nutritious area by searching or following other fish, the area with more fish is generally most nutritious. The AFSA imitates the fish behaviours such as preying, swarming, moving and following with local search of an artificial fish (AF) for reaching the global optimum. The AF's behaviours are defined as follows:

Preying behaviour. An AF is generated with $Y(y_j, j = 1, \dots, n)$ being its current position, where y_j is the proportion of investment on the capacity S_j and y_j is the decision variable to be optimized. The AF randomly searches its neighbour's position Y_v in its vision.

$$Y_v = Y + Visual \cdot rand() \tag{7}$$

where $rand()$ is uniform distributed in 0 and 1, $Visual$ represents the AF's visual distance. The AF compares the value functions of $F(Y)$ and $F(Y_v)$. If $F(Y_v) > F(Y)$ in the maximum problem, then the AF goes forward a step in this direction and arrives at the position Y_{next}, which is given by

$$Y_{next} = Y + rand() \cdot step \cdot \frac{Y - Y_v}{\|Y - Y_v\|} \tag{8}$$

where $step$ is the step length. If $F(Y_v) < F(Y)$, the AF continues an inspecting tour in its visual range, which is described by Eq. 7.

Swarming behaviour. The AF will assemble in groups naturally, which is a kind of living habits. Let Y_c be the center position and N_f be the number of its companions in the current neighbourhood. n is the total fish number. Δ measures the crowdedness of the AF's neighbourhood. If $F(Y_c) > F(Y)$ and $\frac{N_f}{n} < \Delta$, which means that the companion center has more food and is not very crowed, the AF goes forward a step to the companion center, which is given by

$$Y_{next} = Y + rand() \cdot step \cdot \frac{Y - Y_c}{\|Y - Y_c\|} \tag{9}$$

Otherwise, the AF executes the preying step.

Following or moving behaviour. If the position Y_v is better than Y and the surrounding of Y_v is not crowded, the AF goes forward a step to the position Y_v, which is given in Eq. 8. If the AF cannot find a better position, it continues to search in its vision, which is given in Eq. 7.

The basic AFSA is an iterative algorithm, which gradually converges into global optimum. However, the AFSA could easily fall into the local optimum when the $step$ is fixed. For a larger value of $step$, the AF could quickly assemble in groups in the first iterations, but it could oscillate around the global optimum. However, for a smaller value of $step$, it will take a longer time for the AF to converge and could easily fall into the local optimum. To improve the performance of AFSA, we introduce the principle of uniform mutation to adaptively change the length of visual range and step. Therefore, the improved AFSA is called AFSA_UM in this study.

Let $kesi$ ($kesi < step$) be the threshold to trig the uniform mutation process. If the variation between the current and the previous positions is less than $kesi$, then keep the current position and add a uniform random number on the other fish. The uniform mutation allows us to use larger values of visual range and step at the beginning iterations and switch to the small values at the ending iterations. Therefore, it could greatly increase the convergent speed and precision in finding the global optimum.

To adaptively change the lengths of AF's visual range and step, we first calculate three kinds of vision ranges: (1) $Visual\_avg$ denotes the average distances between the AF and all its neighbour fish, (2) $Visual\_best$ is the distance between the AF and the best neighbour fish, and (3) $Visual\_nrst$ represents the distance between the AF and its nearest neighbour. The AF could use $Visual\_avg$ as the visual range in moving and swarming steps, and use $Visual\_best$ and $Visual\_nrst$ as the visual range in preying step. The adaptive step length could be given by $step\_adp = \gamma \cdot visual$, where γ is a constant between 0 and 1.

The procedure of AFSA_UM to optimize the CM's capacity portfolio is described below.

Step 1: Initialization. Set the fish number (*fishnum*), maximal generation ($maxgen$), the maximal try number (*try_num*), randomly generate an artificial fish swam with the population size *fishnum* ($Y_1, \ldots, Y_{fishnum}$).
Step 2: The CM's capacity portfolio problem is served as the value function.
Step 3: Calculate the visual ranges (e.g., $Visual\_avg$, $Visual\_best$, $Visual\_nrst$) and the adaptive step ($step\_adp$).
Step 4: For Y_j, execute the swarming and following behaviours with $Visual\_avg$, and execute the preying behavior with $Visual\_best$ and $Visual\_nrst$. If a better solution Y_v is found, replace Y_j with Y_v.
Step 5: Update the optimal solution on the bulletin board. If the optimal solution reach the convergent precision, then terminate the iteration; otherwise, go to the next step.
Step 6: Uniform mutation. Calculate the variation between the current and previous optimal positions. If the variation is less than *kesi* (i.e., 0.01), then keep the better fish and add uniform random numbers on the other fish with worse positions.
Step 7: Termination criteria. $gen = gen + 1$. If gen > $maxgen$, or the optimal solution meet the convergent precision, then terminate the iteration.

Before using the AFSA_UM to solve the CM's optimal capacity portfolio, we use the $min\, f(x,y) = x^2 + y^2 - 10(cos2x + cos2y)$ and $5.12 \leq x, y \leq 5.12$ to test the robustness of the AFSA_UM in avoiding falling into the local optimum. The parameters settings are: *fishnum*=100, *step*=0.1, Δ=0.5, *kesi*=0.01, γ=0.5. The convergent precision to terminate is 1.0e-5.

To show the robustness of AFSA_UM, we test the performance of three algorithms: genetic algorithm (GA), the basic AFSA, and the AFSA_UM. We run each algorithm on the testing function for 20 times. Table 2 gives the results. Among three algorithms, AFSA_UM performs the best in convergent precision and speed.

Table 2 Precision and iterations of GA, AFSA and AFSA_UM on the testing function

	Precision_best	Precision _worst	Precision _avg	Iterations_avg	Success rate
GA	3.913e-5	9.16e-4	4.89e-5	97.5	70.3%
AFSA	3.213e-5	8.16e-4	4.34e-5	91.3	74%
AFSA_UM	2.53e-11	1.09e-11	6.73e-11	30.6	94%

4. Results and discussion

4.1 Influences of δ on the retailers' investment decisions

From Corollary 1 to 4, we have shown that the value of δ determines each retailer's investment decision. Recall the thresholds of δ in Corollary 1 to 4, we could have $\delta_1 < 0$ and $\delta_4 > \delta_2 > 0$. However, the sign of δ_3 depends on the values of α and b. Table 3 shows each retailer's investment decision under different values of δ.

As $(\delta_2 - \delta_1)$ denotes the region that both retailers jointly invest the CM, we investigate the influences of α and b on the value of $(\delta_2 - \delta_1)$ given in Table 3. From Corollary 1 and 2, we could obtain that $(\delta_2 - \delta_1)$ decreases with α and b. It means that, when the free-riding behaviour and the market competition are high, both retailers will lower their intentions to jointly invest the CM.

Table 3 Each retailers' investment decisions under the value of δ

In the situation when $a > b/2$
Both retailers jointly invest the CM when $\delta_1 < \delta < \delta_2$
Retailer 1 free rides and Retailer 2 invests the CM when $\delta < \delta_3 < 0$
Retailer 1 invests the CM and Retailer 2 free rides when $0 < \delta_4 < \delta < \delta_3$
Both retailers do not invest the CM when $0 < \delta_2 < \delta < \delta_4$ and $\delta_3 < \delta < \delta_1 < 0$
In the situation when $a \le b/2$
Retailer 1 free rides, Retailer 2 invests the CM when $\delta > \max(\delta_3, \delta_4)$

4.2 Influence of β on the CM's capacity portfolio optimization

To investigate the CM's attitude on revenue (β) on the solution of capacity portfolio, we use the numerical studies where the AFSA and AFSA_UM are utilized. Table 4 gives the parameters of five types of capacity with different combinations of revenue and risk.

The parameters for the AFSA and AFSA_UM are: *Visual* = 2.5, *step* = 0.5, *firshnum* = 100, *maxgen* = 100, *try_num* = 100, Δ = 0.6, *kesi* = 0.01, *γ* = 0.6. The CM's attitude factor on revenue, *β* = {0.1, 0.2, 0.3, 0.4, 0.5}. Both algorithms are run for 20 times, and y_j is the average of the proportion of total investment on the capacity S_j. Table 5 gives the numerical results.

Table 4 The parameters of the capacity portfolio to be optimized

	S_1	S_2	S_3	S_4	S_5
r_i	0.28	0.23	0.21	0.05	0.25
u_i	0.025	0.055	0.015	0.00	0.026

Table 5 The results of CM's capacity portfolio optimization by AFSA and AFSA_UM

		y_1	y_2	y_3	y_4	y_5	R	U	F
$\beta = 0.1$	AFSA	0.401	0.003	0.437	0.053	0.106	0.2339	0.0100	0.0324
	AFSA_UM	0.327	0.18	0.493	0	0	0.2365	0.0099	0.0326
$\beta = 0.2$	AFSA	0.397	0.043	0.3	0.007	0.253	0.2477	0.0099	0.0575
	AFSA_UM	0.365	0.229	0.126	0	0.28	0.2450	0.0126	0.0591
$\beta = 0.1$	AFSA	0.539	0.056	0.206	0.009	0.19	0.2550	0.0135	0.0859
	AFSA_UM	0.505	0.18	0	0	0.315	0.2616	0.0126	0.0873
$\beta = 0.4$	AFSA	0.639	0.056	0.206	0	0.099	0.2598	0.0160	0.1135
	AFSA_UM	0.627	0	0.183	0	0.19	0.2615	0.0157	0.1140
$\beta = 0.5$	AFSA	0.754	0	0.006	0	0.24	0.2724	0.0189	0.1456
	AFSA_UM	0.773	0	0	0	0.227	0.2800	0.0210	0.1525

From Table 5, we have two findings:

- The AFSA_UM always outperforms the AFSA in two folds: (i) it provides a higher total revenue while lowering the maximal loss, and (ii) it provides higher value of objective function than the AFSA.
- The CM prefers to increase the proportion of S_1 when β increases. It means that the CM tends to increase the installment of the capacity with the highest revenue when he increase his attitude on the revenue.

4.3 Managerial findings of the study

In this section, we summarize the managerial findings of the study.

- From Table 3, we find that the retailer with obvious cost advantage over her competitor will invest the CM, because (i) she could increase her profit by enjoying a lower production cost, and (ii) she could dominate the market share by providing a lower price to the customers, thus deterring the competitor. Her competitor, who is disadvantageous in cost, will choose to free ride. Because, the competitive inferiority in production cost is so obvious that she gives up investing the CM. The competitor benefits from free riding in two folds: (i) she could enjoy a lower production cost, and (ii) order more products when $\alpha > b/2$. However, the competitor's profit is not surely to increase.
- In the situation when the absolute value of δ (i.e., $|\delta|$) is low (see Table 3), both retailers will jointly invest the CM. The reason is that, when the cost advantage (or disadvantage) is not obvious, the retailer who does not invest the CM will have a lower profit and a lower market share, and hence, be in a state of competitive inferiority. However, both retailers' intentions of jointly investment will decrease in α and b. The reason is that, high substitutable products and high spillover effect will increase both retailers' intention to free ride when the difference of production costs is not obvious.
- In the situation when $|\delta|$ is large, the retailer with obviously cost advantageous will continue to invest the CM, because she could dominate the market with even lower production cost and her profit will increase by investment. However, the retailer with obviously cost disadvantage will free ride, because the difference of production costs is so high that she will be in competition inferiority even she provides the investment.
- In the situation when $|\delta|$ is medium, both retailers will not invest the CM. The reasons are in two folds: (i) if the retailer with cost advantage invests the CM, then the competitor could benefit more through free riding, such as lowering the retailing price, increasing the profit and market share. (ii) Oppositely, if the retailer with cost disadvantage invests the CM, then her competitor could further lower the production cost, lower the retailing price, increase the profit and dominate the market. Therefore, both retailers will choose not to invest the CM as their investment decisions.
- The CM always benefits as long as one retailer provides the investment. The reasons are in two folds: (i) he will lower the wholesale prices for both retailers to incentivize the retailers to order more products, thus increasing the production quantity; and (ii) he could increase the marginal profit of wholesaling the product, and hence, increase the total profit. The numerical study shows that the CM prefers to install more capacity with the highest revenue in his capacity portfolio when his attitude on the revenue increases. AFSA_UM outperforms the AFSA in optimizing the CM's capacity portfolio problem in two folds: (i) it converges with higher speed and precision, and (ii) it provides higher revenue and lower risk, and higher value of objective function than the AFSA. Because the AFSA_UM has the ability to adaptively change the AF's searching step by introducing the uniform mutation.

5. Conclusion

In this study, the operational decisions of a supply chain with two competitive retailers and a shared contract manufacturer (CM) was firstly examined in this study. Then the paths of the retailers' investment strategies were investigated in detail. The CM's capacity portfolio problem was model as a classic investment portfolio problem which was solved by the artificial fish swarm algorithm modified by uniform mutation (AFSA_UM). The following conclusions were obtained in this study:

- For the CM, he always benefits from the retailers' investment, because both his marginal profit and total profit will increase. The CM will install a higher proportion of capacity with higher revenue when he emphasizes more on the revenue. The CM's capacity portfolio problem is optimized by the AFSA_UM which introduces the principle of uniform mutation to boost the convergent speed and precision. The numerical results show that the AFSA_UM is much more robust than the basic AFSA, and it could provide the CM a better capacity portfolio with higher revenue, lower maximal risk, and higher value of objective function.
- For the retailers, their investment decisions are significantly influenced by the absolute difference of production costs (i.e., $|\delta|$). Specifically, (i) both retailers jointly invest the CM when the value of $|\delta|$ is low. The retailers will lower their intentions to jointly invest when the free riding behaviour and the substitutability of the products are high. (ii) No retailer would invest the CM when the value of $|\delta|$ is medium. (iii) Only the retailer with obvious cost advantage will invest the CM when the value of $|\delta|$ is high.

Three research directions can follow from this study. First, no market uncertainty is considered in this study. It would be interesting to investigate how the downside risk of demand will influence the competitive retailers' investment decisions. Second, the information asymmetry phenomenon is not considered yet in this study. In the industrial practice, no retailer could have the compete knowledge of her competitors. However, the introduction of this phenomenon will greatly complicate the mathematic model and the procedure of analyses. Third, the optimization process of CM's capacity portfolio is independent of the retailers' investment decisions in this study. However, the value function of optimizing the CM's capacity portfolio will be complicated when the retailers' decisions are incorporated.

References

[1] Bendoly, E., Bharadwaj, A., Bharadwaj, S. (2012). Complementary drivers of new product development performance: Cross-functional coordination, information system capability, and intelligence quality, *Production and Operations Management*, Vol. 21, No. 4, 653-667, doi: 10.1111/j.1937-5956.2011.01299.x.

[2] Cajigas-Romero, M., Haro, M.D.C., Ramirez, E. (2018). Working capital investment strategy, *Dyna*, Vol. 93, No. 4, 359, doi: 10.6036/8869.

[3] Vizinger, T., Žerovnik, J. (2018). Coordination of a retail supply chain distribution flow, *Tehnički Vjesnik – Technical Gazette,* Vol. 25, No. 5, 1298-1305, doi: 10.17559/TV-20161219120040.

[4] Harabi, N. (2002). The impact of vertical R&D cooperation on firm innovation: An empirical investigation, *Economics of Innovation and New Technology*, Vol. 11, No. 2, 93-108, doi: 10.1080/10438590210900.

[5] Cellini, R., Lambertini, L. (2009). Dynamic R&D with spillovers: Competition vs cooperation, *Journal of Economic Dynamics and Control*, Vol. 33, No. 3, 568-582, doi: 10.1016/j.jedc.2008.08.006.

[6] Shibata, T. (2014). Market structure and R&D investment spillovers, *Economic Modelling*, Vol. 43, 321-329, doi: 10.1016/j.econmod.2014.08.014.

[7] Xing, M. (2017). The optimal risk choice of cooperative and noncooperative R&D in duopoly with spillovers, *Bulletin of Economic Research*, Vol. 69, No. 4, 173-185, doi: 10.1111/boer.12109.

[8] Wang, J., Ou, S., Li, S. (2017). Pricing strategy for end-of-life electrical and electronic products with remanufacturing under joint recovery mode, *Journal of Engineering Science and Technology Review*, Vol. 10, No. 5, 76-83, doi: 10.25103/jestr.105.09.

[9] Gupta, S. (2008). Research note – Channel structure with knowledge spillovers, *Marketing Science*, Vol. 27, No. 2, 247-261, doi: 10.1287/mksc.1070.0285.

[10] Chen, Y.-J., Chen, Y. (2014). Strategic outsourcing under technology spillovers, *Naval Research Logistics*, Vol. 61, No. 7, 501-514, doi: 10.1002/nav.21599.

[11] Ge, Z., Hu, Q., Xia, Y. (2014). Firms' R&D cooperation behavior in a supply chain, *Production and Operations Management*, Vol. 23, No. 4, 599-609, doi: 10.1111/poms.12037.

[12] Wang, Y., Xiao, Y., Yang, N. (2014). Improving reliability of a shared supplier with competition and spillovers, *European Journal of Operational Research*, Vol. 236, No. 2, 499-510, doi: 10.1016/j.ejor.2014.01.015.

[13] Veldman, J., Gaalman, G.J.C. (2015). Competitive investments in cost reducing process improvement: The role of managerial incentives and spillover learning. *International Journal of Production Economics*, Vol. 170, Part B, 701-709, doi: 10.1016/j.ijpe.2015.08.016.

[14] Qi, A., Ahn, H.-S., Sinha, A. (2015). Investing in a shared supplier in a competitive market: Stochastic capacity case, *Production and Operations Management*, Vol. 24, No. 10, 1537-1551, doi: 10.1111/poms.12348.
[15] Agrawal, A., Kim, Y., Kwon, H.D., Muthulingam, S. (2016). Investment in shared suppliers: Effect of learning, spillover, and competition, *Production and Operations Management*, Vol. 25, No. 4, 736-750, doi: 10.1111/poms.12503.
[16] Van Mieghem, J.A. (2003). Commissioned paper: Capacity management, investment, and hedging: Review and recent developments, *Manufacturing & Service Operations Management*, Vol. 5, No. 4, 269-302, doi: 10.1287/msom.5.4.269.24882.
[17] Chien, C.-F., Dou, R., Fu, W. (2017). Strategic capacity planning for smart production: Decision modeling under demand uncertainty, *Applied Soft Computing*, Vol. 68, 900-909, doi: 10.1016/j.asoc.2017.06.001.
[18] Chen, T.-L., Chen, Y.-Y., Lu, H.-C. (2013). A capacity allocation and expansion model for TFT-LCD multi-site manufacturing, *Journal of Intelligent Manufacturing*, Vol. 24, No. 4, 847-872, doi: 10.1007/s10845-012-0634-9.
[19] Barahona, F., Bermon, S., Günlük, O., Hood, S. (2005). Robust capacity planning in semiconductor manufacturing, *Naval Research Logistics*, Vol. 52, No. 5, 459-468, doi: 10.1002/nav.20086.
[20] Wu, C.-H., Chuang, Y.-T. (2010). An innovative approach for strategic capacity portfolio planning under uncertainties, *European Journal of Operational Research*, Vol. 207, No. 2, 1002-1013, doi: 10.1016/j.ejor.2010.05.015.
[21] Wu, C.-H., Chuang, Y.-T. (2012). An efficient algorithm for stochastic capacity portfolio planning problems, *Journal of Intelligent Manufacturing*, Vol. 23, No. 6, 2161-2170, doi: 10.1007/s10845-011-0562-0.
[22] Jakubovskis, A. (2017). Flexible production resources and capacity utilization rates: A robust optimization perspective, *International Journal of Production Economics*, Vol. 189, 77-85, doi: 10.1016/j.ijpe.2017.03.011.
[23] Jian, M., Wang, Y.L. (2018). Decision-making strategies in supply chain management with a waste-averse and stockout-averse manufacturer, *Advances in Production Engineering & Management*, Vol. 13, No. 3, 345-357, doi: 10.14743/apem2018.3.295.
[24] Niroomand, I., Kuzgunkaya, O., Bulgak, A.A. (2012). Impact of reconfiguration characteristics for capacity investment strategies in manufacturing systems, *International Journal of Production Economics*, Vol. 139, No. 1, 288-301, doi: 10.1016/j.ijpe.2012.05.012.
[25] Feng, Q., Lu, L.X. (2012). The strategic perils of low cost outsourcing, *Management Science*, Vol. 58, No. 6, 1196-1210, doi: 10.1287/mnsc.1110.1481.
[26] Song, J., Li, F., Wu, D.D., Liang, L., Dolgui, A. (2017). Supply chain coordination through integration of innovation effort and advertising support, *Applied Mathematical Modelling*, Vol. 49, 108-123, doi: 10.1016/j.apm.2017.04.041.
[27] Cai, Y. (2010). Artificial fish school algorithm applied in a combinatorial optimization problem, *International Journal of Intelligent Systems and Applications*, Vol. 2, No. 1, 37-43, doi: 10.5815/ijisa.2010.01.06.
[28] He, L.X., He, S.H. (2015). Solving water resource scheduling problem through an improved artificial fish swarm algorithm, *International Journal of Simulation Modelling*, Vol. 14, No. 1, 170-181, doi: 10.2507/ijsimm14(1)CO5.
[29] Omega, R.S., Noel, V.M., Masbad, J.G., Ocampo, L.A. (2016). Modelling supply risks in interdependent manufacturing systems: A case study, *Advances in Production Engineering & Management*, Vol. 11, No. 2, 115-125, doi: 10.14743/apem2016.2.214.
[30] Lin, C., Yang, B. (2018). Simulation of flow line scheduling of production enterprises based on improved artificial fish swarm algorithm, *International Journal of Simulation Modelling*, Vol. 17, No. 3, 512-520, doi: 10.2507/ijsimm17(3)CO11.

APEM journal

Advances in Production Engineering & Management ISSN 1854-6250
Volume 14 | Number 2 | June 2019 | pp 251–263 Journal home: apem-journal.org
https://doi.org/10.14743/apem2019.2.326 Original scientific paper

Experimental investigation and multi-objective optimization of micro-wire electrical discharge machining of a titanium alloy using Jaya algorithm

Singh, M.[a], **Ramkumar, J.**[a], **Rao, R.V.**[b,*], **Balic, J.**[c]

[a]Department of Mechanical Engineering, Indian Institute of Technology Kanpur, Kanpur, India
[b]Department of Mechanical Engineering, Sardar Vallabhbhai National Institute of Technology Surat, Surat, India
[c]Faculty of Mechanical Engineering, Laboratory for Intelligent Manufacturing Systems, University of Maribor, Maribor, Slovenia

A B S T R A C T

Micro-wire electrical discharge machining (Micro-WEDM) process exhibits superior precision and greater relative accuracy for the efficient machining of difficult-to-machine materials. The micro-slit cutting operation using WEDM process has been experimentally investigated for the objective of analysing the average kerf-loss and responses pertaining to the economic viability of the process viz. average cutting rate and volumetric material removal rate (MRR_v). The experiments are performed using a Tungsten wire of diameter 70 µm on titanium grade 5 alloy (Ti-6Al-4V). Three different controllable process variables (input parameters) associated with the Resistance-Capacitance (RC) based power generator namely discharge energy, wire feed-rate and wire travelling speed are varied to demonstrate their impacts on typical responses such as average kerf-loss, average cutting rate and MRR_v. The experimental analysis revealed a close relationship that cutting rate bears with discharge energy, wire feed-rate and efficient flushing of molten liquid as well as fine debris particles. An advanced multi-objective optimization technique popularly known as Multi Objective-Jaya (MO-Jaya) algorithm has been adopted for the simultaneous optimization of average kerf-loss, average cutting rate and volumetric material removal rate. The best set of input parameters have been selected to suggest the most optimum responses for micro wire-cutting operations.

A R T I C L E I N F O

Keywords:
Micro-wire electrical discharge machining (Micro-WEDM);
Multi-objective optimization;
Titanium alloy;
Kerf-loss;
Cutting rate;
Volumetric material removal rate;
Feed-rate;
Jaya algorithm;
Multi objective-Jaya algorithm (MO-Jaya)

**Corresponding author:*
ravipudirao@gmail.com
(Rao, R.V.)

Article history:
Received 2 November 2018
Revised 5 May 2019
Accepted 27 May 2019

1. Introduction

Despite several advancements in the applicability of micro-WEDM process, the kerf-loss and its detrimental influence on the relative accuracy and precision have remained an area to be addressed. Due to the occurrence of secondary discharges between the wire and workpiece surface, the feature size is intended to be larger than the wire diameter. Therefore, the width of the feature perpendicular to the motion of wire-travel is commonly termed as kerf-loss. It consists of the diameter of the wire and the radial overcut on either side of the wire. The radial overcut introduces the geometric inaccuracies in the intended feature as it adds the dimension of the feature in addition to the wire diameter. Moreover, the lateral vibration of the wire which is extremely difficult to detect and prevent owing to the complex phenomenon happening in the narrow gap results in the increased kerf-loss [1]. Minimization of kerf-loss demands either reduction in the diameter of the wire-electrode or lessening in the radial overcut generated during the machining. Owing to the thermal nature of the process there is always a limitation on the minimum diameter of the wire that can be applied productively. Reducing the wire diameter below a certain limit would result in poor tensile strength and consequently ruptures due to thermal softening. Therefore, the reduction of radial/diametric overcut is the only possible way to diminish the kerf-loss. However, a minimum amount of overcutting is always desirable and bound to happen as this narrow pass allows the effective evacuation of fine machining products such as molten liquid and fine debris. Several factors those govern the kerf-loss in the WEDM process include secondary-sparks, bridging of the gap, wire-vibration, low dielectric strength of the fluid, etc. Evacuation of machining products such as fine debris is utmost important for the improved efficiency of the process [2]. Moreover, the electrical and thermal properties of both tool and workpiece material determine the responses in EDM [3] as well as in the WEDM process. However, the non-contact nature of the process makes it a prominent contender for burr-free machining of delicate parts irrespective of its hardness [4]. Fig. 1 shows the schematic diagram of a typical wire slit cutting operation and inclusion of kerf-loss.

Varieties of input parameters affect the material removal rate as well as the kerf-loss in the WEDM process, but the open-circuit voltage and pulse-duration are the noticeable parameters which largely affect the kerf-loss and machining rate [5]. It has also been established that the kerf-loss in the WEDM process is primarily due to the lateral vibration of the wire electrode perhaps caused by the poor wire-tension. Incorporating the appropriate parameters influencing the wire vibration the kerf-loss was reduced with a minimum value of 30.8 μm using a wire of 30 μm diameter [1]. Wire speed has a less significant effect on the outputs as compared to other parameters [5, 6]. The uniformity of sparks across the thickness of the workpiece is essential in minimizing the wire vibration [7]. Majority of the work carried out in the direction of minimization of kerf-loss and maximization of machining rate in the WEDM process is related to the transistor-based circuit where most of the input parameters are arbitrarily decided by the user. However, in the RC based circuit, the capacitance of the capacitor largely determines the discharge-frequency, pulse-duration, pulse-interval and energy per-spark.Therefore, it is essential to establish an experimental analysis pertaining to RC based circuit where the capacitance is the deciding factor. Moreover, the simultaneous optimization of conflicting output responses is paramount importance for the economic viability and improved dimensional accuracy of the process.

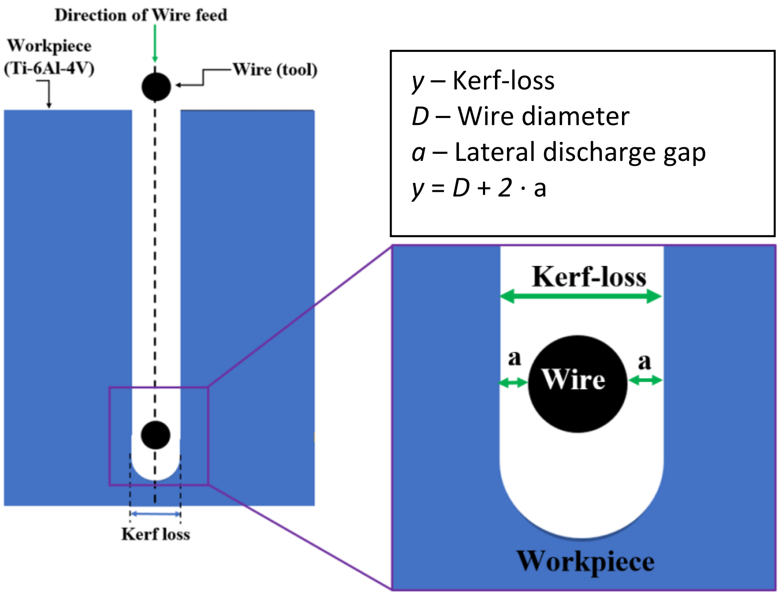


Fig. 1 Schematic illustration of slit cutting operation and the concept of kerf-loss in WEDM process

The rest of the paper is organized into two major sections. The first section consists of the experimentation of micro-slit cutting operation and the influence of various input parameters on the selected output parameters. Finally, adopting an advanced multi-attribute decision-making technique, multi-objective optimization of the micro-WEDM process is conducted, and the results are shown. In order to perform multi-objective optimization of micro-WEDM process, an advanced technique called multi objective-Jaya algorithm is used. The priori approach of the Jaya algorithm first optimizes the individual responses and then convert them into a single objective function to determine the global optimized input parameters and responses.

2. Materials, methods, and design of experiments

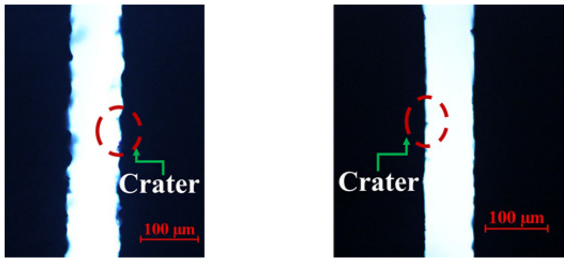
The following nomenclature/abbreviations are used in the paper: WEDM: Wire electrical discharge machining; EDM: Electrical discharge machining, MRR_v; Volumetric material removal rate; RC: Resistance capacitance; MO: Multi objective, SEM: Scanning electron microscopic; μm: Micrometer; μJ: Micro Joule; C: Capacitance of the capacitor; V_0: Open circuit voltage; R^2: R square value; Adj. R^2: Adjusted R square value.

The experiments are conducted using the integrated multipurpose machine tool DT-110 (Make Mikrotools Pte. Ltd. Singapore) equipped with an RC based power generator. The micro-WEDM unit available with the machine tool is mounted at the vertical stage of the setup. The micro-WEDM unit consists of a wire-spool, wire guide cum tensioner, wire cutting envelope and a separate spool for collecting the eroded wire. The process parameters which can be varied during the experiments are discharge energy, wire feed-rate and wire travelling speed, whereas the constant parameters consist of charging resistance (1 kOhm), initial spark-gap (20 μm), dielectric fluid (EDM-Oil), wire (Tungsten, 70 μm diameter), wire-tension (15 %). Titanium grade 5 alloy (Ti-6Al-4V) having a thickness of 800 μm has been used as a workpiece material in the study. The discharge energy is a combined term which includes capacitance (C) of the capacitor used and open circuit voltage (V_0) according to the following relationship [8].

$$E = \frac{1}{2} C V_0^2 \tag{1}$$

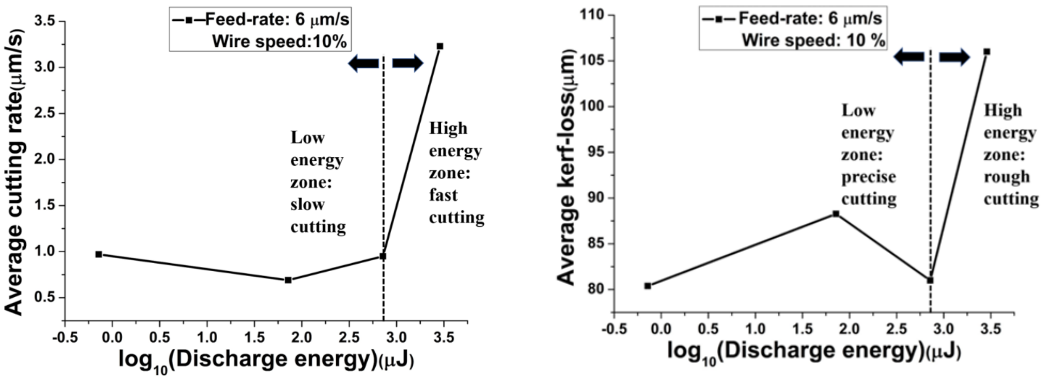
V_0 is kept constant at 120 V, as the Eq. 1 is pertinent for the condition of maximum energy delivered to the gap [8]. However, the maximum energy condition is hardly fulfilled due to several losses in the process. Therefore, it is the discharge/breakdown voltage which determines the discharge energy. In other cases, the discharge energy is largely influenced by the capacitance value as the discharge voltage invariably remains constant with the constant set of electrodes material, dielectric and inter-electrode-gap.

Exploratory experiments are performed at four distinct levels of discharge energy ranging from 0.72 μJ to 2880 μJ. It has been observed from preliminary experiments that the average cutting rate is exceptionally high at 2880 μJ discharge energy. However, the average kerf-loss and quality of kerf in terms of sharp cutting are observed to be very meagre, and therefore it is recommended to use this energy-level merely for cutting purposes where kerf quality is not a prime interest. The kerf quality of Ti-6Al-4V workpiece machined at two different conditions of discharge energy (2880 μJ and 0.72 μJ) are clearly visualized in Fig. 2(a) and 2(b), respectively. The variation in average cutting rate and average kerf-loss with discharge energy are shown in Fig. 3(a) and 3(b), respectively. A logarithmic scale is used for discharge energy to circumvent the asymmetry of data. It is clearly evident that there is a sudden increase in both average cutting rate as well as average kerf-loss at discharge energy 2880 μJ as compared to other energy levels. Accordingly, the machining operation can be divided into two zones. High energy zone can be used for rough cutting operation whereas the low energy zone is appropriate for the precise cutting operation. Hence, the highest level of discharge energy (2880 μJ) is not considered for detailed experimental investigations. Similarly, the range of feed-rate and wire speed are also determined through preliminary experiments considering the short-circuits and wire-vibration separately.



(a) (b)

Fig. 2 Optical microscopic image of slit machined at discharge energy (a) 2880 µJ, (b) 0.72 µJ (Wire feed-rate: 6 µm/s, wire speed: 10 %)



(a) (b)

Fig. 3 Influence of discharge energy on (a) average cutting rate, and (b) average kerf-loss

Table 1 Input parameters and their levels

S. No.	Parameter	Unit	Levels		
1	Discharge energy	µJ	0.72	72	720
2	Feed-rate	µm/s	2	4	6
3	Wire speed	%	10	15	20

Based on the preliminary experiments the ranges of input parameters for detailed experimental investigations are decided. Table 1 displays the three input parameters and their levels used for the slit-cutting operation whereas, Table 2 shows full randomized experimental runs (27) and measured values of responses after conducting the experiments. Std. and Run in Table 2 represent the standard and randomized order of experiments, respectively.

Full factorial design of experiments is performed for analyzing the influence of different process parameters on typical responses. The total number of experiments are $(m)^n = (3)^3 = 27$, where n is the number of factors and m is the number of levels for each factor.

The average kerf-loss is the mean value of the readings of the kerf-width measured at three different locations (inlet, centre and outlet) of the slit. Though the average cutting rate, as well as MRR_v, represent the productivity aspect as well as the rate of machining, yet the MRR_v is related to the total volume of material removed during the operation. Whereas, average cutting rate merely considers the length of the slit machined in the given time and hence neglects the kerf-loss.

$$Average\ cutting\ rate\ (\mu m/s) = \frac{L}{T} \tag{2}$$

where L is the length of the slit machined and T is the total time of machining.

The MRR_v is given as follows [5]:

$$MRR_v\ (\mu m^3/s) = K \cdot t \cdot f \tag{3}$$

where K is the average kerf-loss (µm), t is the thickness of the workpiece (µm), and f is the average cutting rate (µm/s) given by Eq. 2. The workpiece thickness for all experiments is kept constant at 800 µm.

Table: 2 Experimental runs and responses according to full factorial design

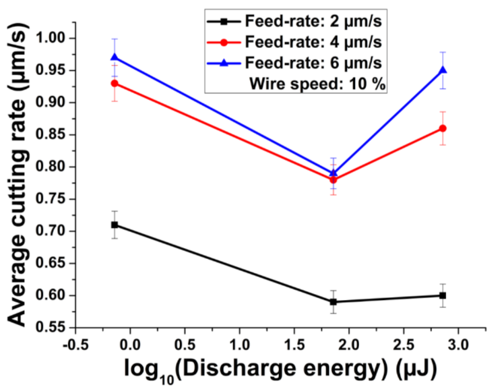
		Input parameters			Responses		
Std.	Run	*Discharge energy* (μJ)	*Feed-rate* (μm/s)	*Wire speed* (%)	*Average kerf-loss* (μm)	*Average cutting rate* (μm/s)	MRR_V $\times 10^3$ ($\mu m^3/s$)
4	1	0.72	4	10	80.35	0.93	59.78
21	2	720	2	20	82.19	0.59	38.79
15	3	720	4	15	80.99	0.71	46
8	4	72	6	10	88.27	0.69	48.73
24	5	720	4	20	83.43	0.56	37.38
13	6	0.72	4	15	79.68	0.89	56.73
17	7	72	6	15	87.93	0.73	51.35
11	8	72	2	15	88.89	0.56	39.82
22	9	0.72	4	20	80.72	0.79	51.02
18	10	720	6	15	80.41	0.81	52.11
12	11	720	2	15	81.17	0.59	38.31
25	12	0.72	6	20	78.99	0.94	59.4
16	13	0.72	6	15	80.74	0.98	63.3
14	14	72	4	15	86.48	0.9	62.27
9	15	720	6	10	79.89	0.95	60.72
3	16	720	2	10	81.20	0.6	38.98
20	17	72	2	20	86.32	0.55	37.98
26	18	72	6	20	87.95	0.93	65.44
2	19	72	2	10	88.13	0.59	41.60
7	20	0.72	6	10	80.38	0.97	62.38
5	21	72	4	10	87.09	0.78	54.35
19	22	0.72	2	20	81.23	0.66	42.89
1	23	0.72	2	10	81.04	0.71	46.03
23	24	72	4	20	88.19	0.83	58.59
27	25	720	6	20	80.92	0.87	56.32
6	26	720	4	10	80.30	0.86	55.25
10	27	0.72	2	15	80.72	0.72	46.50

3. Results and discussion

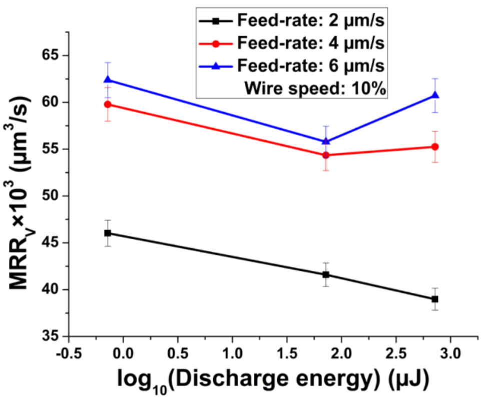
The typical responses of interest include average cutting rate, MRR_V and average kerf-loss. The responses are analyzed with the variation in input-parameters used in the study.

3.1. Effect of discharge energy

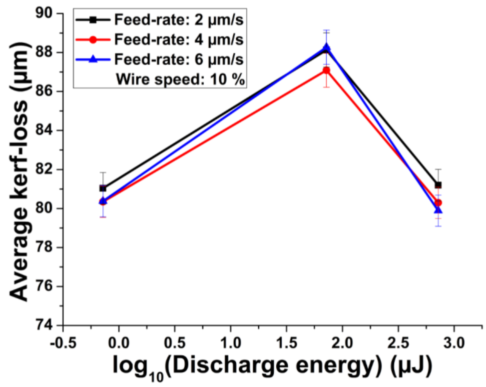
Fig. 4(a) shows the variation in average cutting rate with discharge energy. It is clearly evident from the graph that the average cutting rate is invariably constant at lowest (0.72 μJ) and highest (720 μJ) discharge energy levels used in the study. Though there is a considerable increase in the discharge energy, the feed-rate and flushing conditions are such that the substantial improvement in machining rate is not achieved. It can be attributed to the fact that an increase in discharge energy increases the volume of material removal per-spark, so the quantity of molten material which is to be removed is large. However, the feed-rate, as well as flushing conditions, are not adequate to remove the whole liquid metal and therefore the expected increase in machining rate is not achieved. The combination of sparks frequency and discharge energy suggest that at lower discharge energy the sparks frequency is high, but the volume of material per pulse is small. Whereas at the high discharge energy level the volume of molten material per pulse is large but low sparks frequency allows adequate pulse-off duration for the evacuation of molten material pool and debris. However, at the intermediate value of discharge energy, the combination of sparks frequency and molten volume of material is critical to remove it efficiently. This can be established by the fact that the average cutting rate at 72 μJ is lower than that at 0.72 μJ. The inability to remove the entire molten material in the frontal gap prevents the wire to proceed forward and therefore more interaction time is available for secondary discharges to happen. This fact can be clearly verified by the increased kerf-loss at this energy level as shown in Fig. 4(c). The same phenomenon is responsible for MRR_v though it is slightly higher owing to larger kerf-loss at 72 μJ, Fig. 4(b). Therefore flushing is an utmost important factor to consider in the micro-EDM as well as micro-WEDM process [9]. Moreover, an exact trend is difficult to establish in RC based power generator due to variation in charging of capacitor [10].The average kerf-loss at the lowest and highest energy levels are not considerably different. The constant value of feed-rate enables the servo system to move with constant speed and, therefore, the time available for secondary sparks to occur is invariably the same. Though the volume of the crater formed with higher discharge energy is large yet the frequency of sparks is low as the higher discharge energy is achievable with increased capacitance value. Higher the value of capacitance greater will be the discharge energy and lower is the frequency of sparks.



(a)



(b)



(c)

Fig. 4 Influence of discharge energy on: (a) average cutting rate, (b) MRR_v and (c) average kerf-loss

3.2 Effect of wire feed-rate

The average cutting rate increases continuously with an increase in wire feed-rate in the given range. Wire feed-rate in the micro-WEDM/EDM process comes into action only when there is a short circuit between wire and workpiece or there is a gap enlargement due to the formation of the crater. In both the situations, the wire is either retracted away from the workpiece or moves forward to maintain an equilibrium gap with speed specified by the user termed as wire feed-rate. Therefore, the non-productive time associated with wire retraction in case of short circuit and forward motion to adjust the gap reduces with the increased wire feed-rate. Both the average cutting rate as well as MRR_v increases with the increase in wire feed-rate in the given range of 2-6 μm/s, Fig. 5(a) and 5(b). However, beyond a certain value further increase in wire feed-rate, continuous short circuits are observed and machining ceases. Therefore, the judicious decision must be taken while selecting the appropriate wire feed-rate. However, the preliminary experiments have revealed the use of wire feed-rate as high as 10-15 μm/s with discharge energy of 2880 μJ as the large volume of molten material creates a large size crater.

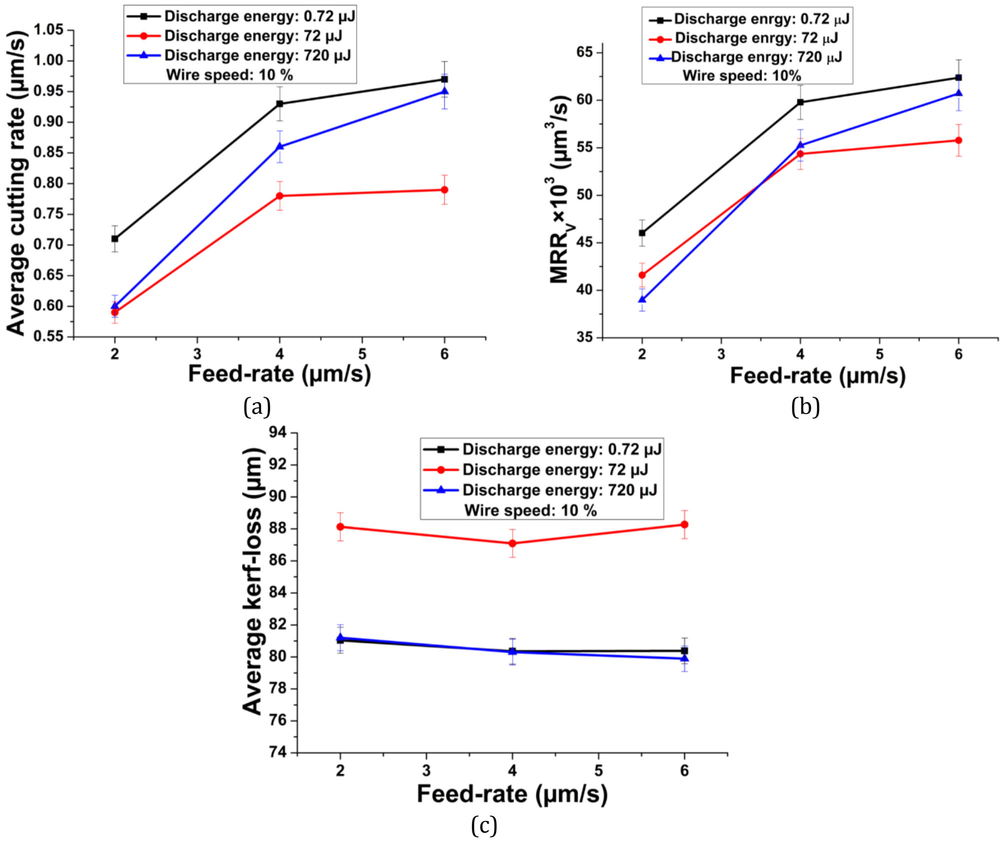


Fig. 5 Influence of wire feed-rate on (a) average cutting rate, (b) MRR_v and (c) average kerf-loss

The average kerf-loss is not much influenced by the wire feed-rate though it is decreasing slightly at 720 µJ and 0.72 µJ discharge energy levels. The occurrence of secondary sparks decreases as the wire travels with higher feed-rate and thus the kerf-loss reduces slightly. At 72 µJ discharge energy the average kerf-loss is observed to be decreasing with an increase in wire feed-rate from 2 µm/s to 4 µm/s and then it increases slightly at 6 µm/s, Fig. 5(c), due to frequent short circuits as the wire is not able to move forward and more secondary sparks are taking place in the lateral gap.

3.3 Effect of wire travelling speed

The wire travelling speed refers to the speed of wire circulation from the wire supplying wheel to the wire collecting wheel. The prevalent phenomenon of tool wear in EDM as well as WEDM process necessitates the wire to travel continuously so that a fresh or uneroded wire can be presented for cutting operation. If the wire is stationary, then the high temperature in the vicinity of the plasma channel would reduce its tensile strength and eventually the wire breaks. However, higher wire travelling speed results in the consumption of more wire. Therefore, the wire speed is decided while considering the wire breakage and less consumption simultaneously. The Fig. 6(a) and 6(b) reveal a less significant influence of wire speed on both average cutting rate and MRR_v. This insignificant variation of output parameters with wire speed is also supported by the study conducted in [5]. A slight decrease in the responses with an increase in wire speed from 15-20 % is attributed to the vibration of wire resulting from the vibration induced in the wire guides as the speed increases. The average kerf-loss is not considerably affected by the wire speed as the number of secondary sparks are not reduced significantly with the variation in wire speed. However, at 72 µJ discharge energy, it is observed that the increased wire speed improves the flushing of molten material in the frontal gap and therefore the interaction of the wire in the lateral gap reduces. This can be established with a decrease in kerf-loss at 72 µJ discharge energy while increasing the wire speed from 15-20 %, Fig. 6(c).

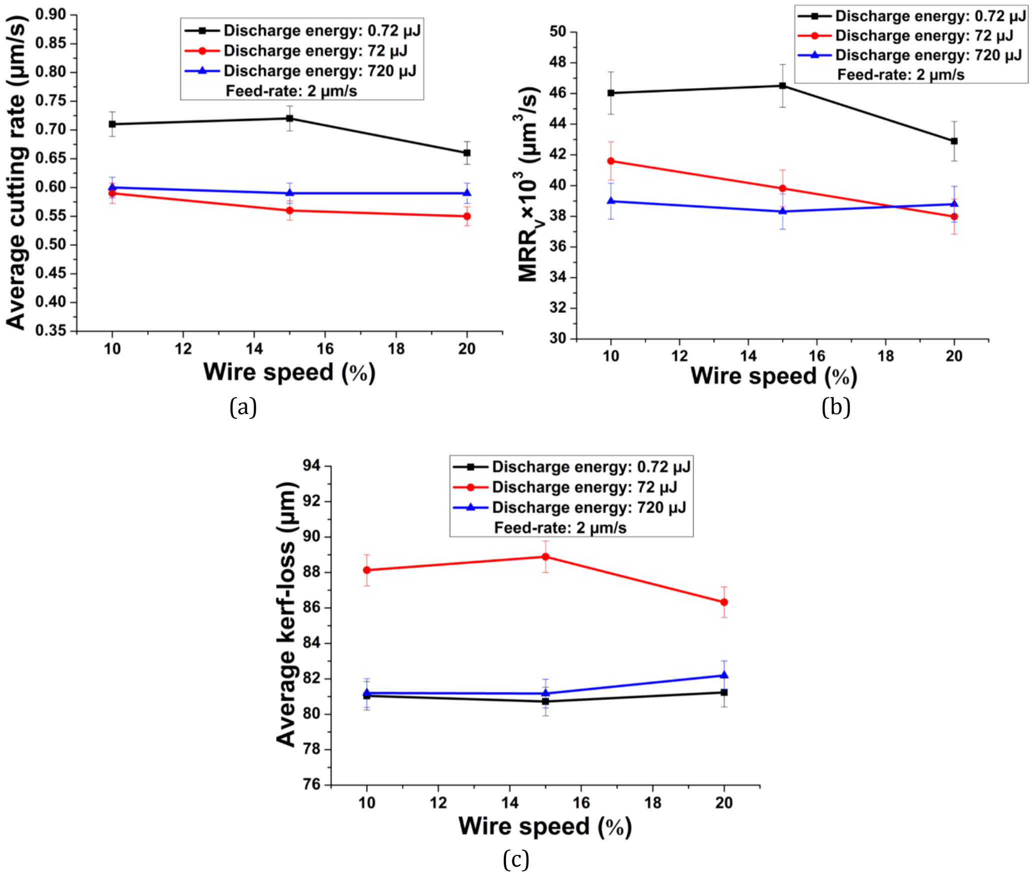


Fig. 6 Influence of wire speed on (a) average cutting rate, (b) MRR_v and (c) average kerf-loss

3.4 Multi-objective optimization of the micro-WEDM process with multi objective-Jaya (MO-Jaya) algorithm

Multi-objective optimization problem deals with the assertion of the best set of input parameters in the given range to satisfy the responses with conflicting desirability simultaneously [11]. The values of process parameters used in experimental investigations are rarely optimum, and thus multi-objective optimization is required to come with the optimum set of input parameters [12]. The micro slit-cutting operation presented in current work consists of three responses with contradictory natures, i.e. two of them (average cutting rate and MRR_v) demand maximization whereas one of the responses (average kerf-loss) is to be minimized. The problem of synchronized optimization of various responses with incompatible desirability is categorized as a multi-objective optimization problem.

Multi objective-Jaya (MO-Jaya) algorithm

Jaya algorithm is an advanced optimization technique. Unlike other similar techniques such as genetic algorithm, particle swarm algorithm and artificial bee colony algorithm, the Jaya algorithm doesn't have algorithm specific parameters which sometimes affect the final results if a fine tuning between the parameters specific to the algorithm is not achieved [13-14].

The multi-objective optimization consists of the determination of global optimum input parameters which can simultaneously satisfy all the conflicting responses. The optimum values of each response while solving them individually are used to convert them into a single/combined objective function. The single objective function is either maximized or minimized to evaluate the optimum global values of input parameters. This method is termed as a priori approach of the Jaya algorithm.

The present experimental analysis of micro-WEDM process consists of three input parameters and three number of output responses which are expressed in the mathematical form in Eqs. 4, 5 and 6. The prediction models given by Eqs. 4, 5 and 6 represent the output responses such as average cutting rate, MRR_v and average kerf-loss respectively in terms of input parameters. These prediction models can predict the output at distinct levels of input parameters in the given range. Among three output responses, average kerf-loss is to be minimized whereas the other two namely average cutting rate and MRR_v are to be maximized. Performing regression analysis, the quadratic prediction models for each response in terms of input parameters are obtained as follows.

Average cutting rate:

$$f_1 = 0.520896284 - 0.001555172 \cdot A + 0.149297749 \cdot B - 0.00829 \cdot C + 2.22741 \cdot 10^{-6} \cdot A^2 - 0.014722222 \cdot B^2 + 4.44444 \cdot 10^{-5} \cdot C^2 + 1.36956 \cdot 10^{-5} \cdot A \cdot B + 0.001916667 \cdot \mathrm{B} \cdot \mathrm{C} - 1.781 \times 10^{-5} \cdot \mathrm{C} \cdot \mathrm{A} \tag{4}$$

Volumetric material removal rate:

$$f_2 = 33734.23458 - 34.2746937 \cdot \mathrm{A} + 9953.068531 \cdot B - 684.5006238 \cdot C + 0.051817546 \cdot A^2 - 1001.25 \cdot \mathrm{B}^2 + 6.333333333 \cdot C^2 + 0.592846241 \cdot \mathrm{A} \cdot \mathrm{B} + 135.6666667 \cdot \mathrm{B} \cdot \mathrm{C} - 1.023688223 \cdot C \cdot A \tag{5}$$

Average kerf-loss:

$$f_3 = 82.82242011 + 0.109846399 \cdot \mathrm{A} - 0.291825702 \cdot B - 0.210612268 \cdot C - 0.000155738 \cdot A^2 + 0.026527778 \cdot B^2 + 0.005711111 \cdot C^2 - 0.00024381 \cdot \mathrm{A} \cdot \mathrm{B} - 0.000416667 \cdot \mathrm{B} \cdot \mathrm{C} + 0.000293298 \cdot \mathrm{C} \cdot \mathrm{A} \tag{6}$$

where A, B and C are the actual values of discharge energy (μJ), wire feed-rate (μm/s) and wire speed (%), respectively. The accuracy of the quadratic fit models is adjudged by the values of R^2 and adj. R^2 for each case. The R^2 and adj. R^2 values are 0.78 and 0.66 for f_1, 0.76 and 0.64 for f_2 and 0.96 and 0.94 for f_3.

The flowchart (Fig. 7) shows the various steps involved in the Jaya algorithm. The same steps are executed for individual optimization of each output response and multi-objective optimization when a combined objective function is used.

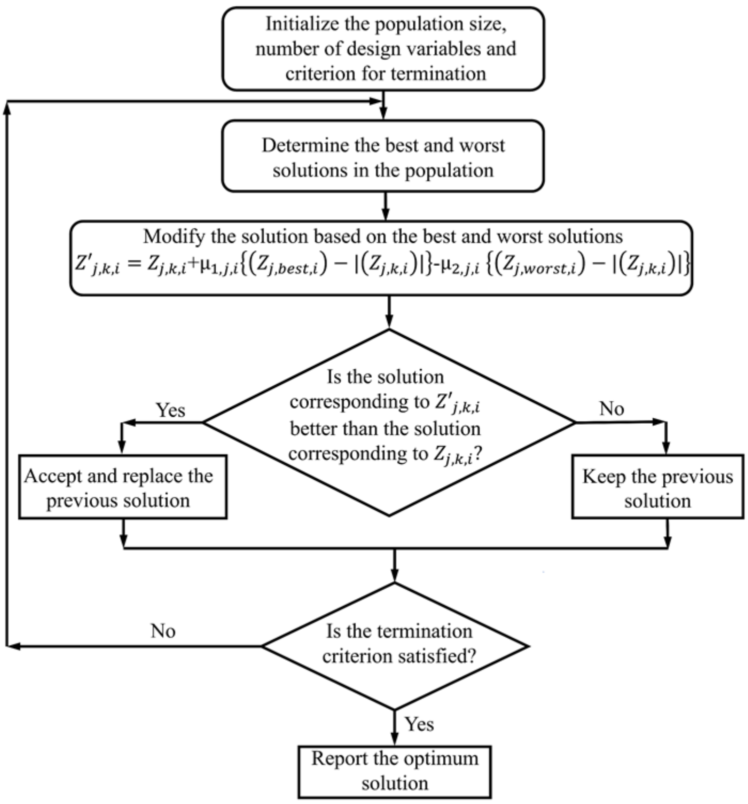


Fig. 7 Flowchart showing the steps in the Jaya algorithm

Individual optimization

First of all, the three output responses given by the regression Eqs. 4, 5 and 6 are optimized separately without taking into consideration the other two outputs. For this step, the regression equation of the respective output responses is assumed to be an objective function.

- (i)The input variables pertaining to Jaya algorithm such as the number of design variables (equal to the number of input parameters, i.e. three), population size (number of candidate solutions) and the termination criterion (i.e. the number of iterations) are decided by the decision maker.
- (ii)The design variables, i.e. the input parameters are assigned the random values in the given range of the respective variables. The range for different variables is: A (0.72-720 μJ), B (2-6 μm/s), C (10-20 %).
- (iii)For each candidate solution, the value of the objective function is calculated. The best and worst value of the objective function is determined. For a maximization problem, the best value corresponds to the maximum value of objective function across all the candidate solutions whereas the worst value corresponds to the minimum value of the objective function. The vice-versa is true for a minimization problem. The corresponding values of input parameters are known as the best and worst values for the respective input parameter.
- (iv)The values of the design variables for each candidate solution is modified as per the following equation [15]:

$$Z'_{j,k,i} = Z_{j,k,i} + \mu_{1,j,i}\left\{\left(Z_{j,\,best,i}\right) - |\left(Z_{j,k,i}\right)|\right\} - \mu_{2,j,i}\left\{\left(Z_{j,worst,i}\right) - |\left(Z_{j,k,i}\right)|\right\} \tag{7}$$

  where, $Z'_{j,k,i}$ is the new value of $Z_{j,k,i}$, $Z_{j,\,best,i}$ and $Z_{j,worst,i}$ are the value of objective function corresponding to the best candidate and worst candidate respectively. $\mu_{1,j,i}$ and $\mu_{2,j,i}$ are two random numbers between 0 and 1.
- (v) Step (iii) is repeated.
- (vi) Compare the values of the objective functions of different candidate solutions in step (v) with the corresponding values of step (iii). If the value of the objective function in step (v) is better than that in step (iii), this new value is accepted along with the corresponding input parameters. Else, the previous solution is retained.
- (vii) This completes one iteration. The algorithm continues to execute until the last iteration at which the values of all candidate solutions become equal.

The value of the objective function and input parameters are accepted as the optimum value of output and input parameters respectively. A similar procedure is repeated for the other two output responses. The optimum values of output responses and the corresponding optimum input parameters while solving them distinctly are given in Table 3. Max_f_1, Max_f_2 are the optimum values of outputs f_1 and f_2, respectively, while solving them individually for maximization whereas Min_f_3 is the optimum value of f_3 while solving it for minimization using Jaya algorithm.

Table 3 Optimized results obtained by using Jaya algorithm when each response is solved individually

Responses	The optimum value of the response	Optimum discharge energy (μJ): A	Optimum feed-rate (μm/s): B	Optimum wire speed (%): C
Max_f_1 (μm/s)	1.2990	0.72	6	20
$Max\_f_2 \times 10^3$ (μm³/s)	62.4941	0.72	6	20
Min_f_3 (μm)	78.1048	720	6	10

Multi-objective optimization

The optimum input parameters given in Table 3 are pertinent only when one of the outputs is to be looked-for while neglecting the other two. However, the success of any process is determined by the combined effects of all output parameters. Therefore, the multi-objective optimization is performed to come up with a set of best input parameters that can satisfy all the responses simultaneously. This can be achieved by converting the outputs with conflicting desirability into a single objective function and then optimizing that combined objective function.

In the present study the combined objective function is formulated by the following equation:

$$Z_{Max} = \frac{f_1}{Max\_f_1} + \frac{f_2}{Max\_f_2} - \frac{f_3}{Min\_f_3} \tag{7}$$

Here, f_1, f_2, f_3 are the prediction models given by Eqs. 4, 5 and 6, respectively, whereas Max_f_1, Max_f_2, Min_f_3 are the optimized values of f_1, f_2 *and* f_3 given in Table 3. Negative sign in the third term indicates the conversion of minimization problem (Min_f_3) into a maximization problem (Z_{Max}).

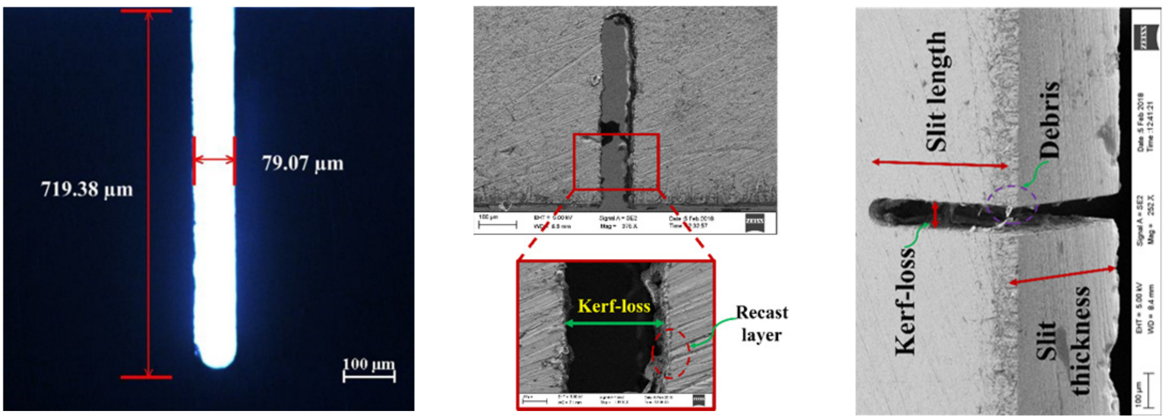
The Jaya algorithm is repeated while considering Z_{Max} to be the objective function. The Jaya algorithm used for this optimization is named as multi objective-Jaya algorithm. The algorithm is executed using a rigorous MATLAB code written in version R2016a (9.0).

The global optimum input parameters satisfying all the output responses simultaneously are given in Table 4. Putting these optimum input values into the regression Eqs. 4, 5 and 6, the corresponding optimum values of responses are also calculated.

MO-Jaya algorithm predicts minimum value of discharge energy (0.72 μJ) and maximum values of feed-rate (6 μm/s) and wire speed (20 %) as an optimum input parameters for simultatneous optimization of three output response. The global optimum input parameters correspond to experiment number 12 of Table 2. Fig. 8(a) shows the optical microscopic image of the micro-slit machined at the global optimum input parameters. Whereas Fig. 8(b) shows the SEM image of a micro-slit machined. The magnified view of the machined surface reveals the formation of recast layer due to the re-solidification of molten liquid. Therefore, the appropriate flushing of molten liquid and debris particles is a paramount necessity for superior dimensional accuracy as well as the better surface characteristics of the process. Fig. 8(c) depicts the uniformity of slit width along the workpiece length and thickness respectively.

Table 4 Results of MO-Jaya algorithm for multi-objective optimization

	Global optimum input parameters			Corresponding optimum values of responses		
Parameters	*Discharge energy* (μJ): A	*Feed-rate* (μm/s): B	*Wire speed* (%): C	*Kerf-loss* (μm)	*Average cutting rate* (μm/s)	*MRR_v* $\times 10^3$ (μm^3/s)
Values	0.72	6	20	80.1790	1.2990	62.4941



(a) (b) (c)

Fig. 8 (a) Optical microscopic image of optimized slit according to MO-Jaya method (Experiment No. 12, Table 2) (b) SEM image of a slit machined (c) SEM image of a slit machined showing kerf-loss along the length and thickness (720 μJ discharge energy and 6 μm/s feed-rate)

4. Conclusions

The paper discusses the variation in average cutting rate, MRR_v and average kerf-loss with input parameters such as discharge energy, wire feed-rate and wire speed. Further, an exertion has been made to obtain the global optimum parameters for simultaneous optimization of various conflicting responses. Following conclusions have been drawn from the work presented in the paper:

- On the basis of preliminary experiments for average cutting rate and average kerf-loss, it is observed that for rough cutting operation where the kerf quality is not utmost important the highest energy level (2880 μJ) can be used. However, for the precise cutting operation, it is recommended to use lower discharge energy levels (0.72-720 μJ).
- Unlike theoretical prediction, the average cutting rate and MRR_v have a close relation with discharge energy, wire feed-rate and efficient flushing of molten material as well as debris particles. The adequate flushing condition is essential for the requisite improvement in machining rate.
- A maximum average cutting rate of 0.98 μm/s and MRR_v of 65.44 ×10^3 μm^3/s are reported from the experimental investigations. Whereas the minimum value of average kerf-loss using a tungsten wire (70 μm diameter) is found to be 78.99 μm which implies that the diametral overcut is 12.84 % of the wire diameter. Therefore, the minimum radial side gap of 4.49 μm on either side of the wire is attained.
- Multi objective-Jaya algorithm has been applied for the simultaneous optimization of output responses. The algorithm results suggest minimum discharge energy (0.72 μJ), highest wire feed-rate (6 μm/s) and wire speed (20 %) as the global optimum input parameters.
- Unlike the other advanced optimization algorithms, the MO-Jaya algorithm doesn't have algorithm-specific parameters and this is an advantageous feature. The presence of algorithm-specific parameters in other optimization algorithms may adversely affect the final results if a fine tuning between the algorithm-specific parameters is not achieved. However, the proposed MO-Jaya algorithm needs to be run many times so as to get a Pareto front to get non-dominated solutions if different weights are to be assigned to the objectives.
- Multi-objective optimization with some other advanced evolutionary computation and swarm intelligence optimization methods will be taken up in near future.
- The future prospects of the micro-WEDM process would include thin wall micromachining of Ti-6Al-4V.

References

[1] Di, S., Chu, X., Wei, D., Wang, Z., Chi, G., Liu, Y. (2009). Analysis of kerf width in micro-WEDM, *International Journal of Machine Tools and Manufacture,* Vol. 49, No. 10, 788-792, doi: 10.1016/j.ijmachtools.2009.04.006.

[2] Narasimhan, J., Yu, Z., Rajurkar, K.P. (2005). Tool wear compensation and path generation in micro and macro EDM, *Journal of Manufacturing Processes,* Vol. 7, No. 1, 75-82, doi: 10.1016/S1526-6125(05)70084-0.

[3] D'Urso, G., Ravasio, C. (2017). Material-technology index to evaluate micro-EDM drilling process, *Journal of Manufacturing Processes,* Vol. 26, 13-21, doi: 10.1016/j.jmapro.2017.01.003.

[4] Joshi, S.N., Pande, S.S. (2010). Thermo-physical modeling of die-sinking EDM process, *Journal of Manufacturing Processes,* Vol. 12, No. 1, 45-56, doi: 10.1016/j.jmapro.2010.02.001.

[5] Tosun, N., Cogun, C., Tosun, G. (2004). A study on kerf and material removal rate in wire electrical discharge machining based on Taguchi method, *Journal of Materials Processing Technology,* Vol. 152, No. 3, 316-322, doi: 10.1016/j.jmatprotec.2004.04.373.

[6] Scott, D., Boyina, S., Rajurkar, K.P. (1991). Analysis and optimization of parameter combinations in wire electrical discharge machining, *International Journal of Production Research,* Vol. 29, No. 11, 2189-2207, doi: 10.1080/00207549108948078.

[7] Okada, A., Uno, Y., Nakazawa, M., Yamauchi, T. (2010). Evaluations of spark distribution and wire vibration in wire EDM by high-speed observation, *CIRP Annals,* Vol. 59, No. 1, 231-234, doi: 10.1016/j.cirp.2010.03.073.

[8] Masuzawa, T., Tönshoff, H.K. (1997). Three-dimensional micromachining by machine tools, *CIRP Annals,* Vol. 46, No. 2, 621-628, doi: 10.1016/S0007-8506(07)60882-8.

[9] Karthikeyan, G., Ramkumar, J., Dhamodaran, S., Aravindan, S. (2010). Micro electric discharge milling process performance: An experimental investigation, *International Journal of Machine Tools and Manufacture,* Vol. 50, No. 8, 718-727, doi: 10.1016/j.ijmachtools.2010.04.007.

[10] Jahan, M.P., Wong, Y.S., Rahman, M. (2009). A study on the quality micro-hole machining of tungsten carbide by micro-EDM process using transistor and RC-type pulse generator, *Journal of Materials Processing Technology,* Vol. 209, No. 4, 1706-1716, doi: 10.1016/j.jmatprotec.2008.04.029.
[11] Konak, A., Coit, D.W., Smith, A.E. (2006). Multi-objective optimization using genetic algorithms: A tutorial, *Reliability Engineering & System Safety,* Vol. 91, No. 9, 992-1007, doi: 10.1016/j.ress.2005.11.018.
[12] Rao, R.V., Rai, D.P., Balic, J. (2017). A multi-objective algorithm for optimization of modern machining processes, *Engineering Applications of Artificial Intelligence,* Vol. 61, 103-125, doi: 10.1016/j.engappai.2017.03.001.
[13] Rao, R.V. (2019). *Jaya: An advanced optimization algorithm and its engineering applications,* Springer International Publishing AG, Cham, Switzerland, doi: 10.1007/978-3-319-78922-4.
[14] Rao, R.V., Rai, D.P., Ramkumar, J., Balic, J. (2016). A new multi-objective Jaya algorithm for optimization of modern machining processes, *Advances in Production Engineering & Management,* Vol. 11, No. 4, 271-286, doi: 10.14743/apem2016.4.226.
[15] Rao, R.V. (2016). Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems, *International Journal of Industrial Engineering Computations,* Vol. 7, 19-34, doi: 10.5267/j.ijiec.2015.8.004.

Calendar of events

- 17th Annual Industrial Simulation Conference, June 5-7, 2019, Lisbon, Portugal.
- International Conference on Design and Implementation of Intelligent Manufacturing Systems, June 10-11, 2019, Tokyo, Japan.
- 6th International Conference and Exhibition on Automobile & Mechanical Engineering, July 8-9, 2019, Zurich, Switzerland.
- 35th World Congress on Materials Science and Nanotechnology, July 22-23, 2019, Melbourne, Australia.
- 23rd International Conference on Advanced Materials & Nanotechnology, August 19-20, 2019, Tokyo, Japan.
- AI Manufacturing 2019: Machine Learning and Artificial Intelligence, The Fourth Industrial Revolution, August 28-29, 2019, Westin O'Hare, Rosemont, Illinois, USA.
- 9th IFAC Conference on Manufacturing Modeling, Management, and Control, August 28-30, 2019, Berlin, Germany.
- European Summit on Robots, Artificial Intelligence and Autonomous Vehicles, October 7-8, 2019, Paris, France.
- 25th World Congress on Advanced Materials, October 21-22, 2019, Singapore.
- 30th DAAAM International Symposium, October 23-26, 2019, Zadar, Croatia.
- 32nd Materials Science and Engineering Conference: Advancement & Innovations, October 24-25, 2019, Helsinki, Finland.
- 4th European Conference on Design and Production Engineering, October 28-29, 2019, Rome, Italy.
- 5th International Conference on 3D Printing Technology and Innovations, March 16-17, 2020, Berlin, Germany.

Notes for contributors

General

Articles submitted to the *APEM journal* should be original and unpublished contributions and should not be under consideration for any other publication at the same time. Manuscript should be written in English. Responsibility for the contents of the paper rests upon the authors and not upon the editors or the publisher. Authors of submitted papers automatically accept a copyright transfer to *Chair of Production Engineering*, *University of Maribor*. For most up-to-date information on publishing procedure please see the *APEM journal* homepage *apem-journal.org*.

Submission of papers

A submission must include the corresponding author's complete name, affiliation, address, phone and fax numbers, and e-mail address. All papers for consideration by *Advances in Production Engineering & Management* should be submitted by e-mail to the journal Editor-in-Chief:

Miran Brezocnik, Editor-in-Chief
UNIVERSITY OF MARIBOR
Faculty of Mechanical Engineering
Chair of Production Engineering
Smetanova ulica 17, SI – 2000 Maribor
Slovenia, European Union
E-mail: editor@apem-journal.org

Manuscript preparation

Manuscript should be prepared in *Microsoft Word 2010* (or higher version) word processor. *Word .docx* format is required. Papers on A4 format, single-spaced, typed in one column, using body text font size of 11 pt, should not exceed 12 pages, including abstract, keywords, body text, figures, tables, acknowledgements (if any), references, and appendices (if any). The title of the paper, authors' names, affiliations and headings of the body text should be in *Calibri* font. Body text, figures and tables captions have to be written in *Cambria* font. Mathematical equations and expressions must be set in *Microsoft Word Equation Editor* and written in *Cambria Math* font. For detail instructions on manuscript preparation please see instruction for authors in the *APEM journal* homepage *apem-journal.org*.

The review process

Every manuscript submitted for possible publication in the *APEM journal* is first briefly reviewed by the editor for general suitability for the journal. Notification of successful submission is sent. After initial screening, and checking by a special plagiarism detection tool, the manuscript is passed on to at least two referees. A double-blind peer review process ensures the content's validity and relevance. Optionally, authors are invited to suggest up to three well-respected experts in the field discussed in the article who might act as reviewers. The review process can take up to eight weeks on average. Based on the comments of the referees, the editor will take a decision about the paper. The following decisions can be made: accepting the paper, reconsidering the paper after changes, or rejecting the paper. Accepted papers may not be offered elsewhere for publication. The editor may, in some circumstances, vary this process at his discretion.

Proofs

Proofs will be sent to the corresponding author and should be returned within 3 days of receipt. Corrections should be restricted to typesetting errors and minor changes.

Offprints

An e-offprint, i.e., a PDF version of the published article, will be sent by e-mail to the corresponding author. Additionally, one complete copy of the journal will be sent free of charge to the corresponding author of the published article.

APEM journal

Advances in Production Engineering & Management

Chair of Production Engineering (CPE)
University of Maribor
APEM homepage: apem-journal.org

Volume 14 | Number 2 | June 2019 | pp 139-266

Contents

ISSN 1854-6250

9 771854 625008

apem-journal.org