

ISSN 1854-6250

APEM
journal

Advances in Production Engineering & Management

Volume 19 | Number 1 | March 2024





University of Maribor

Published by CPE
apem-journal.org

Advances in Production Engineering & Management

Identification Statement

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

APEM Editorial

Editor-in-Chief

Miran Brezocnik

editor@apem-journal.org, info@apem-journal.org
University of Maribor, Faculty of Mechanical Engineering Smetanova ulica 17, SI – 2000 Maribor, Slovenia, EU

Desk Editor

Martina Meh

desk1@apem-journal.org

Janez Gotlih

desk2@apem-journal.org

Website Technical Editor

Lucija Brezocnik

desk3@apem-journal.org

Editorial Board Members

Eberhard Abele, Technical University of Darmstadt, Germany
Bojan Acko, University of Maribor, Slovenia
Joze Balic, University of Maribor, Slovenia
Agostino Bruzzone, University of Genoa, Italy
Borut Buchmeister, University of Maribor, Slovenia
Ludwig Cardon, Ghent University, Belgium
Nirupam Chakraborti, Indian Institute of Technology, Kharagpur, India
Edward Chlebus, Wroclaw University of Technology, Poland
Igor Drstvensek, University of Maribor, Slovenia
Illes Dudas, University of Miskolc, Hungary
Mirko Ficko, University of Maribor, Slovenia
Vlatka Hlupic, University of Westminster, UK
David Hui, University of New Orleans, USA
Pramod K. Jain, Indian Institute of Technology Roorkee, India
Isak Karabegović, University of Bihać, Bosnia and Herzegovina
Janez Kopac, University of Ljubljana, Slovenia

Changsong Ma, Geely University of China, China
Qingliang Meng, Jiangsu University of Science and Technology, China
Lanndon A. Ocampo, Cebu Technological University, Philippines
Iztok Palcic, University of Maribor, Slovenia
Krstj Pandza, University of Leeds, UK
Andrej Polajnar, University of Maribor, Slovenia
Antonio Pouzada, University of Minho, Portugal
R. Venkata Rao, Sardar Vallabhbhai National Inst. of Technology, India
Rajiv Kumar Sharma, National Institute of Technology, India
Katica Simunovic, J. J. Strossmayer University of Osijek, Croatia
Daizhong Su, Nottingham Trent University, UK
Soemon Takakuwa, Nagoya University, Japan
Nikos Tsourveloudis, Technical University of Crete, Greece
Tomo Udiljak, University of Zagreb, Croatia
Ivica Veza, University of Split, Croatia



Subsidizer: The journal is subsidized by Slovenian Research and Innovation Agency



Creative Commons Licence (CC): Content from published paper in the APEM journal may be used under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Statements and opinions expressed in the articles and communications are those of the individual contributors and not necessarily those of the editors or the publisher. No responsibility is accepted for the accuracy of information contained in the text, illustrations or advertisements. Chair of Production Engineering assumes no responsibility or liability for any damage or injury to persons or property arising from the use of any materials, instructions, methods or ideas contained herein.

Published by CPE, University of Maribor.

Advances in Production Engineering & Management is indexed and abstracted in the **WEB OF SCIENCE** (maintained by **Clarivate**): **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



Advances in Production Engineering & Management

Volume 19 | Number 1 | March 2024 | pp 1–152

Contents

Scope and topics	4
An improved multi-objective firefly algorithm for integrated scheduling approach in manufacturing and assembly considering time-sharing step tariff	5
Xu, E.B.; Zou, F.F.; Shan, P.P.; Wang, Z.Y.; Shi, B.X.	
Research on recovery strategies of supply chain network under disruption propagation using memetic algorithm	21
Li, Z.Y.; Zhao, P.X.; Wang, C.L.; Mi, Y.Z.	
Human-robot collaboration assembly line balancing considering cross-station tasks and the carbon emissions	31
Li, Y.C.; Wang, X.	
Predicting the deep drawing process of TRIP steel grades using multilayer perceptron artificial neural networks	46
Sevšek, L.; Vilkovský, S.; Majerníková, J.; Pepelnjak, T.	
The investigation of production variance in a module-based assembly system: A Markovian Arrival Process approach	65
Zhong, C.W.; Zhang, H.; Zhang, Z.W.; Wu, Z.Q.; Lu, M.L.	
A comparative study of machine learning regression models for production systems condition monitoring	78
Jankovič, D.; Šimic, M.; Herakovič, N.	
Evolutionary game analysis of green innovation in E-commerce closed-loop supply chain WEEE recycling	93
Ma, R.M.; Yao, L.F.; Wang, H.	
FDM process parameter selection by hybrid MCDM approach for flexural and compression strength maximization	108
Begic-Hajdarevic, D.; Klančnik, S.; Muhamedagić, K.; Čekić, A.; Cohodar Husic, M.; Ficko, M.; Gusel, L.	
Integrated optimization of line planning and timetabling on high-speed railway network considering cross-line operation	117
Wang, R.X.; Nie, L.; Fang, W.; Ren, H.Q.; Tan, Y.Y.	
Experimental and numerical investigation of the deep drawing process using a tractrix die – An industrial case study focused on stress and temperature analysis	133
Mandic, V.; Milosavljevic, Dj.; Jurkovic, Z.; Adamovic, D.	
Calendar of events	149
Notes for contributors	151

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

An improved multi-objective firefly algorithm for integrated scheduling approach in manufacturing and assembly considering time-sharing step tariff

Xu, E.B.^{a,b,*}, Zou, F.F.^a, Shan, P.P.^a, Wang, Z.Y.^b, Shi, B.X.^c

^aSchool of Computer and Software, Nanyang institute of technology, Nanyang, P.R. China

^bSchool of Mechanical and Precision Instrument Engineering, Xi'an University of Technology, Xi'an, P.R. China

^cSchool of Naval Architecture, Ocean and Energy Power Engineering, Wuhan University of Technology, Wuhan, P.R. China

ABSTRACT

Today, energy conservation and reduction of consumption are crucial concerns for manufacturing companies. Current research on integrated scheduling of processing and assembly typically focuses only on equipment resources and processing and assembly processes. A new method for energy-saving integrated scheduling in workshops has been proposed, which incorporates the recently introduced time-of-use tiered electricity prices into the scheduling optimization model. This method also introduces an operation strategy of turning equipment on and off during idle periods. A multi-objective mathematical model was developed to minimize energy consumption and assembly delay time in the processing and assembly processes. Due to the complexity of the model, the standard firefly algorithm was improved when used to solve the model. This involved designing a three-layer encoding method and two decoding methods, and providing detailed steps of the algorithm. Using a mixed flow production line as an example, the final scheduling solutions were obtained through model construction and algorithm solving, taking into account the tiered electricity price. The results of the example demonstrate that parallel processing and assembly effectively reduce assembly delay costs, and the implementation of the on/off strategy reduces power consumption during the machining process.

ARTICLE INFO

Keywords:

Energy saving;
Integrated scheduling;
Manufacturing and assembly;
Time-sharing step tariff;
Switch strategy;
Firefly algorithm;
Multi-objective model

*Corresponding author:

baojams@126.com
(Xu, E.B.)

Article history:

Received 11 August 2023
Revised 24 January 2024
Accepted 28 January 2024



Content from this work may be used under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

As the main energy of manufacturing industry, how to save energy has become a problem that manufacturing industry has to consider [1-4]. To solve the above problems, the key is whether through the product manufacturing and assembly process management, to achieve energy-saving cooperative manufacturing of parts manufacturing and assembly. Therefore, it is of great practical significance for energy saving and consumption reduction of manufacturing and assembly manufacturing enterprises to study the energy saving integrated scheduling optimization problem of manufacturing and assembly links [5].

The production process of manufacturing and assembly enterprises includes many links such as process planning before production, workshop scheduling and final assembly molding after workpiece processing. Scholars at home and abroad have studied the workshop scheduling of

single link or the integrated scheduling of multiple links. The research on the integrated scheduling of the two stages of manufacturing and assembly in the production process can reduce the high cost and low efficiency caused by separate scheduling, which has also attracted wide attention of scholars.

Peng and Zheng [6] developed an improved wild horse optimization (IWHO) algorithm to simultaneously optimize three objectives: makespan, maximum machine workload, and total machine workload. Huo and Wang [7] proposed a hybrid dynamic scheduling method with Digital Twin and improved bacterial foraging algorithm (IBFOA) to minimize the maximum completion time and machine load. The results show that the scheduling scheme using the IBFOA can optimize the system performance as a whole and effectively deal with the problem of extended production time caused by disruption. Komaki and Kayvanfar [8] proposed an improved Grey Wolf Optimizer (GWO) to minimize the maximum completion time of two-stage flow shop scheduling with delivery time. Sun *et al.* [9] proposed a dynamic shop scheduling method integrating deep reinforcement learning and convolutional neural network (CNN), and could adaptively select appropriate dispatching rules based on the state features of the production system.

Ren *et al.* [10] tackled Distributed Permutation Flow-shop Scheduling Problems (DPFSPs), aiming to minimize the maximum completion time of workpieces. The authors proposed a NASH Q-Learning algorithm based on Mean Field (MF), employing a two-layer online learning mode within a multi-agent Reinforcement Learning framework, and demonstrated its superior efficiency over similar algorithms. Tian and Zhang [11] focused on the dynamic job-shop scheduling problem (JSP) with the goal of reducing manpower and material costs. They proposed a dynamic job-shop scheduling model using deep learning, specifically employing a long short-term memory network (LSTM) with Dropout technology and adaptive moment estimation (ADAM) to enhance prediction. The optimization was targeted at three objective functions, and the multi-objective problem was solved using an improved multi-objective genetic algorithm (MOGA), with experimental results proving the algorithm's effectiveness. Zhao and Yuan [12] addressed the integrated scheduling of production and maintenance in a parallel machine job-shop environment, considering stochastic machine breakdowns. They developed an integrated model utilizing minimal and preventive maintenance strategies and designed a genetic algorithm to solve the problem. The model and algorithm were verified through an instance, proving their effectiveness. Wang [13] investigated the quality management and control of multi-variety small-batch production (MVSBP) manufacturing logistics, analyzing factors and optimizing elements like site selection, path design, and warehousing. A real case simulation in a porcelain blank processing workshop demonstrated the effectiveness of their approach in aligning with the shift from traditional mass production to MVSBP. Belmahdi *et al.* [14] conducted a comprehensive review of the various scheduling methods and algorithms utilized in Fog Computing, a paradigm that extends the capacities of Cloud and improves performance and QoS for applications. They analyzed, compared, and classified these scheduling approaches according to the algorithm's nature, the optimized QoS, and application types such as critical IoT (CIOT), massive IoT (MIOT), and Industry IoT (IIOT), and provided a comparison of different simulation tools to guide developers and researchers in the field of fog computing. Ren *et al.* [15] developed a two-stage optimization algorithm for the JSP, considering job transport, that combines the improved fast elitist nondominated sorting genetic algorithm II (INSGA-II) with a local search strategy, and an ant colony algorithm based on reinforcement learning (RL-ACA), to minimize makespan, tardiness, and energy consumption, demonstrating superiority over other algorithms in similar problems. Bedhief and Dridi [16] tackled the three-stage hybrid flow shop scheduling problem with two dedicated machines in stage 3, aiming to minimize the maximum completion time (makespan), and proposed an improved genetic algorithm (IGA) that incorporates more than one crossover operator and a 2-opt local search method, demonstrating effectiveness and efficiency in comparison to an existing heuristic approach. Zhao and Yuan [17] developed a multi-objective optimization model for integrating job-shop production scheduling and predictive maintenance, considering constraints like product delivery time and changing machine failure rate, and used an enhanced non-dominated sorting genetic algorithm (NSGA)-II with simulated binary crossover (SBX) to minimize processing cost and time, validating the approach with a case study.

Deng *et al.* [18] research aimed at minimizing the total completion time, including the three-stage integrated scheduling problem of workpiece processing, transportation and assembly, and proposed a hybrid distribution estimation algorithm integrating multiple rules. Aiming at the integrated optimization scheduling problem of mixed-flow assembly line and machining line with parallel machine, Guo and Ryan [19] addresses a real-life uncertainty factor identified in a manufacturer of large vehicles, by modelling unreliable part delivery and quality. Stochastic optimization is applied to find sequencing policies that improve the on-time performance of its mixed-model assembly lines. Wei *et al.* [20] adopts colored Petri nets (CPN) and a minimal spanning tree (MST) to address the type I problems for two-sided assembly line balancing problem (TALBP). Aiming at the integrated optimization problem of manufacturing and assembly workshop, Liang *et al.* [21] establishes a multi-objective mathematical model for integrated scheduling of manufacturing and assembly workshops with the goal of minimizing the completion time of manufacturing and assembly stages, and designs a genetic algorithm to solve the model. The feasibility of the model and the effectiveness of the algorithm are verified by examples.

Current research indicates that the combined scheduling problem of manufacturing and assembly continues to prioritize minimizing completion time and delay, with limited focus on energy conservation. Specifically, there is a lack of studies addressing the impact of time-sharing step tariffs on energy usage. Nevertheless, energy consumption is a crucial factor in scheduling. Therefore, it is essential to investigate integrated scheduling for energy efficiency. In particular, integrating machining and assembly with time-sharing step tariffs holds significant theoretical and practical value.

Therefore, this paper focuses on the processing and assembling manufacturing enterprises, integrating the scheduling issues of manufacturing and assembly links for the processing-assembling mixed-flow production line. It systematically examines the integrated scheduling problems of manufacturing and assembly with a focus on energy conservation, considering time-sharing step tariffs and aiming to reduce assembly delay time and power costs. The paper is structured as follows: Section 2 outlines the energy-saving integrated scheduling problem under time-sharing step tariffs, develops the integrated scheduling mathematical model, and presents the model's objective function and constraint conditions. Section 3 enhances the multi-objective firefly algorithm to address the scheduling model from Section 2. Finally, Section 4 validates the correctness of the model from Section 2 and the effectiveness of the algorithm from Section 3 through an example.

2. Construction of energy-saving integrated scheduling model

In the mixed workshop of production and assembly considering the time-sharing step tariff, the processing equipment completes the processing of the workpiece according to the product requirements, and the assembly equipment completes the product assembly according to the equipment process. In the process of processing, there are many kinds of workpieces, different kinds of workpieces need to be processed by different process routes; In the assembly process, each assembly equipment corresponds to an assembly node. Before assembly, the corresponding workpiece to be assembled needs to be processed. Therefore, in the workshop, the processing sequence of the workpiece to be processed needs to be sorted, and the starting time of the assembly needs to be determined to reduce the waiting time of the assembly process and ensure the timely delivery of the product. To minimize the power cost and delay time of manufacturing and assembly process, the integrated scheduling of manufacturing and assembly process is carried out.

For the integrated optimization scheduling problem of manufacturing and assembly, the following assumptions are made: (1) At the same time, each processing equipment can only process one work piece; (2) The processing route of the workpiece, the processing time of each processing procedure and the assembly time of each assembly procedure are known and determined in advance; (3) Once each processing or assembly process starts to complete the processing task, it cannot be interrupted; (4) The preparation time of tool change and clamping is included in the processing or assembly time of each process; (5) Each processing procedure of the workpiece must be completed before the subsequent processing procedure is completed; (6) The workpiece needed at the beginning of the assembly process has been processed.

To establish a mathematical model for integrated scheduling optimization of machining and assembly, the following symbols are defined:

i	Workpiece number, $i \in [1, \dots, I]$;
j	identification number, $j \in [1, \dots, J]$;
k	Electricity price period, $k \in [1, \dots, K]$;
v	Electricity price ladder, $v \in [1, \dots, V]$;
a	Assembly node number, $a \in [1, \dots, A]$;
w_i	Number of processes of workpiece i , $w \in [1, \dots, w_i]$;
u_j	Number of locations of machine j , $u \in [1, \dots, u_j]$;
ST	Planned start time;
ET	Planned finish time;
O_{iwju}	Boolean variable, the w -way program for the job i is processed at the u position on machine j , then the value is 1, otherwise 0;
SM_{iwju}	The starting time of w -channel process of workpiece i at u position on machine j ;
EM_{iwju}	The completion time of w -channel procedure of workpiece i at u position on machine j ;
SA_{ai}	Assembly node a Requirement time for artifacts i ;
EA_{ai}	Workpiece i completion time at assembly node a ;
RT_j	The time consumed when machine j shuts down once;
W_j^{Te}	The energy consumption of machine j shutdown once;
x_{ju}	Boolean variable, machine j takes the switch policy after finishing the machining position u , then the value is 1, otherwise it is 0;
AT_{ai}	Boolean variable, whether the assembly node a needs workpiece i , if so, the value is 1, otherwise the value is 0;
$h_{ai'l}$	Boolean variable, if assembly node a requires that workpiece i after work piece i' , then the value is 1, otherwise it is 0;
t^{idle}_{jk}	The idle time of machine j in period k ;
p^{idle}_j	The standby power of machine j ;
p^{asse}_a	Average power of assembly node a ;
t^{asse}_{ai}	The installation time of workpiece i on assembly node a ;
t^{mach}_{iwj}	The processing time of process w of workpiece i on machine j ;
p^{mach}_{iwj}	The power requirement of process w of workpiece i when machining on machine j ;
p^{pub}	Power of public equipment;
PS_k	Start time of period k ;
PE_k	Finish time of period k ;
PCS_{kv}	Electricity consumption standard for v -th price of period k ;
y_{kv}	Boolean variable, whether the power consumption of period k meets the standard of v , if so, the value is 1, otherwise, it is 0;
P_{kv}	The v -th price of period k ;
Z_{iwjk}	Boolean variable, whether machine j processes w operations of workpiece i in K period. If so, the value is 1; otherwise it is 0;
az_{ak}	Boolean variable, if assembly node a is assembled in period k , the value is 1, otherwise it is 0;
at_{ak}	Assembly time of assembly node a in time k ;
t_{iwjk}	The processing time of w -channel sequence of machine j processing workpiece i in period k ;
R_{jk}	The number of times that machine j adopts switch strategy in period k .

2.1 Objective functions

The processing period is divided into machine tool processing energy consumption and idle standby energy consumption. The assembly period has assembly energy consumption, and then the public energy consumption such as lighting is considered.

(1) Machine tool processing energy consumption

Processing energy consumption of machine j in period k :

$$W_{jk}^{mach} = \sum_{i \in I} \sum_{w \in W_i} t_{iwjk} p_{iwj}^{mach} \tag{1}$$

where

$$t_{iwjk} = \begin{cases} t_{iwj}^{mach}, & \sum_{k \in K} z_{iwjk} = 1 \\ EM_{iwju} - PS_k, & SM_{iwju} \leq PS_k \leq EM_{iwju} \\ PE_k - SM_{iwju}, & SM_{iwju} \leq PE_k \leq EM_{iwju} \end{cases} \tag{2}$$

(2) Standby power consumption

The standby energy consumption of machine j period k :

$$W_{jk}^{idle} = t_{jk}^{idle} p_j^{idle} + R_{jk} W_j^{re} \tag{3}$$

The machine standby time t_{jk}^{idle} :

$$t_{jk}^{idle} = PS_k - PE_k - \sum_{i \in I} \sum_{w \in W_i} t_{iwjk} - \sum_{i \in I} \sum_{w \in W_i} \sum_{u \in U_j} [(S_{irwjr,u+1} - E_{iwju}) x_{ju} z_{iwjk}] \tag{4}$$

(3) Assembly energy consumption

Assembly energy consumption of assembly node a at period k :

$$W_{ak}^{asse} = at_{ak} \times p_a \tag{5}$$

among them:

$$at_{ak} = \begin{cases} t_{ai}^{asse}, & \sum_{k=1}^K az_{ak} = 1 \\ EA_{ai} - PS_k, & SA_{ai} \leq PS_k \leq EA_{ai} \\ PE_k - SA_{ai}, & SA_{ai} \leq PE_k \leq EA_{ai} \end{cases} \tag{6}$$

(4) Public energy consumption of public equipment including lighting equipment

Public energy consumption for period k :

$$W_k^{pub} = (PS_k - PE_k) p^{pub} \tag{7}$$

In summary, the total energy consumption of workshop in period k is:

$$W_k = \sum_{j \in J} (W_{jk}^{mach} + W_{jk}^{idle}) + W_k^{pub} + \sum_{a \in A} W_{ak}^{asse} \tag{8}$$

Energy consumption minimization objective function:

$$\min \left\{ \sum_{k=1}^K \left[P_{k1} W_k + \sum_{v=2}^V (y_{kv} (W_k - PCS_{kv}) (P_{kv} - P_{k,v-1})) \right] \right\} \tag{9}$$

At the same time, the delay time of manufacturing and assembly caused by workpiece processing should be considered. Therefore, the processing time and assembly time of workpiece should be comprehensively considered to establish the second optimization objective:

$$\min \left\{ \begin{aligned} & \sum_{a \in A} \sum_{i \in I} [SA_{ai} - \max(AT_{ai} \times EM_{iwju})] \\ & + \sum_{a=1}^{A-1} [\min(SA_{a+1,i}) - \max(EA_{a,ir})] + \sum_{i \in I} \sum_{w \in W_i} (EM_{iwju} - SM_{i,w-1,jru'}) \end{aligned} \right\} \quad (10)$$

In summary, the objective function is established as follows:

$$\min \left\{ \begin{aligned} & \sum_{k=1}^K \left[P_{k1} W_k + \sum_{v=2}^V (y_{kv} (W_k - PCS_{kv}) (P_{kv} - P_{k,v-1})) \right] \\ & \sum_{a \in A} \sum_{i \in I} [SA_{ai} - \max(AT_{ai} \times EM_{iwju})] + \\ & \sum_{a=1}^{A-1} [\min(SA_{a+1,i}) - \max(EA_{a,ir})] \\ & + \sum_{i \in I} \sum_{w \in W_i} (EM_{iwju} - SM_{i,w-1,jru'}) \end{aligned} \right\} \quad (11)$$

2.2 Constraint conditions

Consider the following constraints.

Uniqueness constraint of machines:

$$SM_{iwrj,u+1} \geq EM_{iwju} + RT_j \times x_{ju} \quad (12)$$

Uniqueness constraint of workpiece:

$$SM_{i,w+1,jru'} \geq EM_{iwju} + RT_j \times x_{ju} \quad (13)$$

Ensure the continuity of the machining process, once the process begins to process, uninterrupted before completion:

$$SM_{iwju} + t_{iwj}^{mach} = EM_{iwju} \quad (14)$$

$$\sum_{k=1}^K t_{iwjk} = t_{iwj}^{mach} \quad (15)$$

Start processing and complete processing tasks within the specified period of time:

$$SM_{iwju} \geq ST \quad (16)$$

$$EM_{iwju} \leq ET \quad (17)$$

Constraints on the number of switching strategies adopted by processing equipment:

$$\sum_{k=1}^K R_{jk} = \sum_{u=1}^{u_j} x_{ju} \quad (18)$$

One process for the same workpiece can only select one machine:

$$\sum_{i \in I} \sum_{w \in W_i} o_{iwju} = 1 \quad (19)$$

When the processing equipment adopts the switch strategy, the free time meets the minimum time requirement of the processing equipment shutdown once:

$$(SM_{iwrj,u+1} - EM_{iwju})x_{ju} \geq RT_j \quad (20)$$

When processing equipment adopts switch strategy, standby energy consumption of idle time is greater than that of shutdown once:

$$(SM_{iwrj,u+1} - EM_{iwju})x_{ju}p_j^{idle} \geq W_j^{re} \quad (21)$$

Start time and finish time of assembly within specified time period:

$$SA_{ai} \geq ST \quad (22)$$

$$EA_{ai} \leq ET \quad (23)$$

Workpiece i completes machining before assembly node needs workpiece i :

$$SA_{ai} \geq \max(AT_{ai} \times EM_{iwju}) \quad (24)$$

The start time of the next assembly node $a + 1$ is less than the end time of the current assembly node a :

$$SA_{a+1,i} \geq EA_{a,i} \quad (25)$$

Ensure the order of assembly requirements on assembly nodes:

$$(SA_{a,i} - EA_{ai})h_{ai} \geq 0 \quad (26)$$

Continuity of assembly, that is, once the assembly process begins, it cannot be interrupted before the assembly process is completed:

$$SA_{ai} + t_{ai}^{asse} = EA_{ai} \quad (27)$$

3. Improved multi-objective firefly algorithm

We strive to reduce the energy expenses and production time during the manufacturing and assembly process, which is a common multi-objective optimization challenge. Currently, there are numerous algorithms available for addressing multi-objective scheduling issues, including Genetic algorithm [22, 23], particle swarm optimization [24], migratory bird optimization [25], and firefly algorithm [26]. The firefly algorithm, as a novel swarm intelligence optimization approach, offers benefits such as a straightforward model, fewer adjustable parameters, simple parallel processing, and rapid convergence, and has been successfully applied in various domains.

Considering the complexity of the problem, the standard multi-objective firefly algorithm cannot completely solve the energy-saving integrated scheduling mathematical model established in Section 2, and several important parts of the algorithm need to be redesigned and improved.

3.1 Encoded mode

In this paper, the combined scheduling optimization issue of manufacturing and assembly under the time-based step tariff can be broken down into four sub-problems. The initial sub-problem involves establishing the processing sequence for the workpiece. The second sub-problem is to determine the start time for each process of the workpiece. The third sub-problem is to ensure the completion of workpiece processing and the start time for assembly at the assembly node. The fourth sub-problem occurs during the processing stage, where the machine decides whether to implement a switch strategy after completing the current processing task, based on specific criteria.

According to the four sub-problems, the three-layer coding is designed as follows: The first layer is based on the workpiece sequence coding, including two parts of the workpiece manufacturing and assembly, the assembly process is regarded as a workpiece, considering the precedence constraints of the assembly workpiece; The second layer coding is the starting time of the workpiece, which corresponds to the first layer coding, and also includes the manufacturing and assembly of the workpiece, respectively corresponding to the processing starting time or assembly starting time of the workpiece in the first layer coding. The third layer code is the code whether to adopt the switch strategy, which is whether the current machine adopts the switch restart operation after the workpiece is processed.

It is presumed that there are three pieces of work to be completed, requiring three processing steps and two assembly steps to finish assembly after processing. The viable design solution is depicted in Figure 1. In the first layer, 1, 2, and 3 represent the three workpieces, while 4 and 5 represent the assembly process. The processing sequence is indicated in the first layer's coding sequence. The second layer shows the corresponding start times, and the third layer indicates whether a shutdown restart operation is necessary.

the first layer:	1	2	3	1	3	2	1	3	2	4	5
the second layer:	0	4	4	3	9	8	9	14	18	25	30
the third layer:	0	0	0	1	0	0	0	1	0	0	0

Fig. 1 Schematic diagram of three-layer coding

3.2 Decoding method

In accordance with the specifications of this issue, a single decoding method is not entirely suitable for this problem, which differs from the typical job-shop scheduling problem by taking energy consumption into account. As a result, this study proposes two decoding methods: one based on workpiece sequence and the other based on start time.

(1) Decoding based on workpiece sequence

Decoding based on workpiece sequence is one of the common decoding methods. Decoding for the first layer of coding, the sequence of the workpiece corresponds to the processing sequence of the workpiece. The starting time of the workpiece is determined by the idle time of the workpiece, the idle time of the processing equipment and the processing time of the workpiece, that is, the workpiece and the machine can start processing when they are idle at the same time. After decoding by this decoding method, the maximum completion time of the workpiece can be obtained, and then whether the workpiece can complete the processing task within the specified time can be judged. If the maximum completion time of the workpiece is within the specified time, this solution meets the requirements and can enter the next operation. Otherwise, this solution does not meet the requirements and cannot carry out the next operation. It needs to be iterated again in order to enter the next step.

(2) Decoding based on start time

The decoding based on the start time is to determine the processing sequence of the workpiece according to the coding based on the workpiece sequence in the first layer of the coding, determine the start time of the workpiece according to the coding based on the start time in the second layer, and determine whether the machine performs the shutdown restart operation after finishing the current workpiece according to the coding based on the switch strategy in the third layer. From the decoding operation, it can be seen that in the final feasible solution, the decoding operation based on the start time is adopted.

3.3 Key link

(1) Location update

The location update formula when the firefly i is attracted by the brighter firefly j and moves to j is:

$$x_i = x_i + \beta(d)(x_j - x_i) + \alpha(rand - \frac{1}{2}) \tag{28}$$

In Eq. 28, x_i and x_j are the spatial positions of individual i and individual j of firefly; α is a step factor, which is a constant between 0-1; 'rand' is a random number with uniform distribution between 0-1; $\alpha(rand - \frac{1}{2})$ is a random disturbance term to avoid falling into local optimum too early in the population iteration process.

In the multi-objective firefly algorithm, in order to maintain the diversity of Pareto solution set, Yang [27] added a random moving method, that is, a new method of position updating, as shown in Eq. 29:

$$x_i = g^* + \alpha \left(\text{rand} - \frac{1}{2} \right) \quad (29)$$

where g^* is the optimal solution in the current population solution space obtained by weighting multiple objective function values according to the random weight (the sum of random weights is 1).

Therefore, in the multi-objective firefly algorithm, the firefly individual has two ways of position updating. When the firefly individual x_i exists in the current population, the position updating operation is carried out according to the Eq. 28. When the firefly individual x_i does not exist in the current solution space, the position is updated according to Eq. 29.

For the discrete domain space encoded by the first layer, when the individual firefly is dominated, the Precedence preserving order-based crossover (POX), which is commonly used in workshop scheduling, is used for crossover operation. This operator can well retain the excellent characteristics of the parent generation, inherit the processing order of the parent generation, and ensure the feasibility of the offspring [28]. If individual fireflies are not dominated, swap (SWAP) operation is adopted.

For the continuous domain space of the second layer coding, the position update operation in the multi-objective firefly algorithm is adopted. If the firefly individual is dominated, the position update is carried out according to Eq. 28. If the firefly individuals are not dominated, they are randomly moved according to Eq. 29 to increase the diversity of Pareto solution set.

(2) Judgement and repair of feasible solution

There are two types of infeasible solutions. The first is the infeasible solution caused by the sequence of the assembly process does not meet the constraints, and the second is the infeasible solution caused by the conflict of the starting time of the workpiece.

The initial solution is not possible due to an infeasible assembly sequence resulting from the interchange operation of the first layer code in the firefly individual. This is caused by constraints in the assembly sequence between the workpiece assembly processes. Therefore, for this kind of infeasible solution, the following feasible solution judgment and repair operation should be implemented.

- Step 1: Find the correct assembly sequence constraint;
- Step 2: Find the assembly process in the first layer of code by traversal and record the location;
- Step 3: Compare the assembly process and assembly sequence constraints in the code, if so, do step 4, otherwise, insert the correct assembly process in turn according to the assembly process location in the code, and return step 2;
- Step 4: Output the feasible solution of the first layer encoding.

The second layer coding has an impractical solution due to the requirement that the workpiece's starting time must align with the unique constraints of the processing equipment and the workpiece when they are both idle. In such cases of infeasible solutions, the following determination and repair operations are implemented.

- Step 1: Determine whether the second layer code is planned after the start time, if so, the next step; otherwise, the coding that is not within the specified time is set as the planned start time before the next step;
- Step 2: The first layer coding is traversed one by one, and the next idle time of the workpiece and the processing equipment is determined according to the current processing status of the workpiece and the processing equipment and the processing time of the workpiece. And determine the feasibility of the second layer encoding corresponding coding, if feasible, the next step, otherwise, with the workpiece and processing equipment at

the next time the idle time to replace the corresponding coding, and then to the next step;

Step 3: Determine whether the second layer code is before the completion time of the plan, if so, go to step 5, otherwise, the code that is not within the specified time range is set as the completion time of the plan, and continue the next step;

Step 4: Inversely traverse the first layer encoding to determine the last idle time of the workpiece and the processing device according to the processing time of the workpiece and the current processing state of the workpiece and the processing device. Determine whether the second layer encoding is feasible, if feasible, take the next step, if not, replace the corresponding encoding with the last free time of the workpiece and processing equipment, and then go to step 1;

Step 5: Output the feasible solution of the second layer encoding.

3.4 Algorithm procedure

The algorithm procedure is designed as follows:

Step 1: Setting initial parameters. m : the number of fireflies, also known as population size, γ : light absorption coefficient, β_0 : maximum attraction, α : step size factor, T : population iteration number, and initializing the location of fireflies;

Step 2: Calculate and sort the dominance level and crowding distance for each firefly individual;

Step 3: The individual firefly x_j is selected to update the position of the first layer coding. If the individual x_j is not dominated, SWAP switching operation is performed on the first layer coding. If the individual x_j is dominated, POX crossover operation is performed on the first layer coding;

Step 4: Judge and repair the feasibility of the first layer coding;

Step 5: The maximum completion time of the work piece is obtained by decoding based on the work piece sequence. If the completion time is within the specified time, the next step is continued, otherwise step 3 is returned;

Step 6: Select the firefly individual x_j to update the location for the second layer code: If the individual x_j is not dominated, then move randomly according to Eq. 29; If the individual x_j is dominated, then update the location according to Eq. 28;

Step 7: The feasibility of the second layer coding is judged and repaired, and then the third layer coding is calculated according to the switch strategy requirements, and a firefly individual is obtained;

Step 8: If all firefly individuals have a location update operation, continue the next step, otherwise, return step 3;

Step 9: Implement elite strategy, merge father and son generation to select elite, keep good individual;

Step 10: Determines whether the number of iterations satisfies the population iteration number T , and if so, continues the next step, otherwise returns step 2;

Step 11: Calculate and output Pareto front solution set.

The procedure of the algorithm is shown in the Fig. 2.

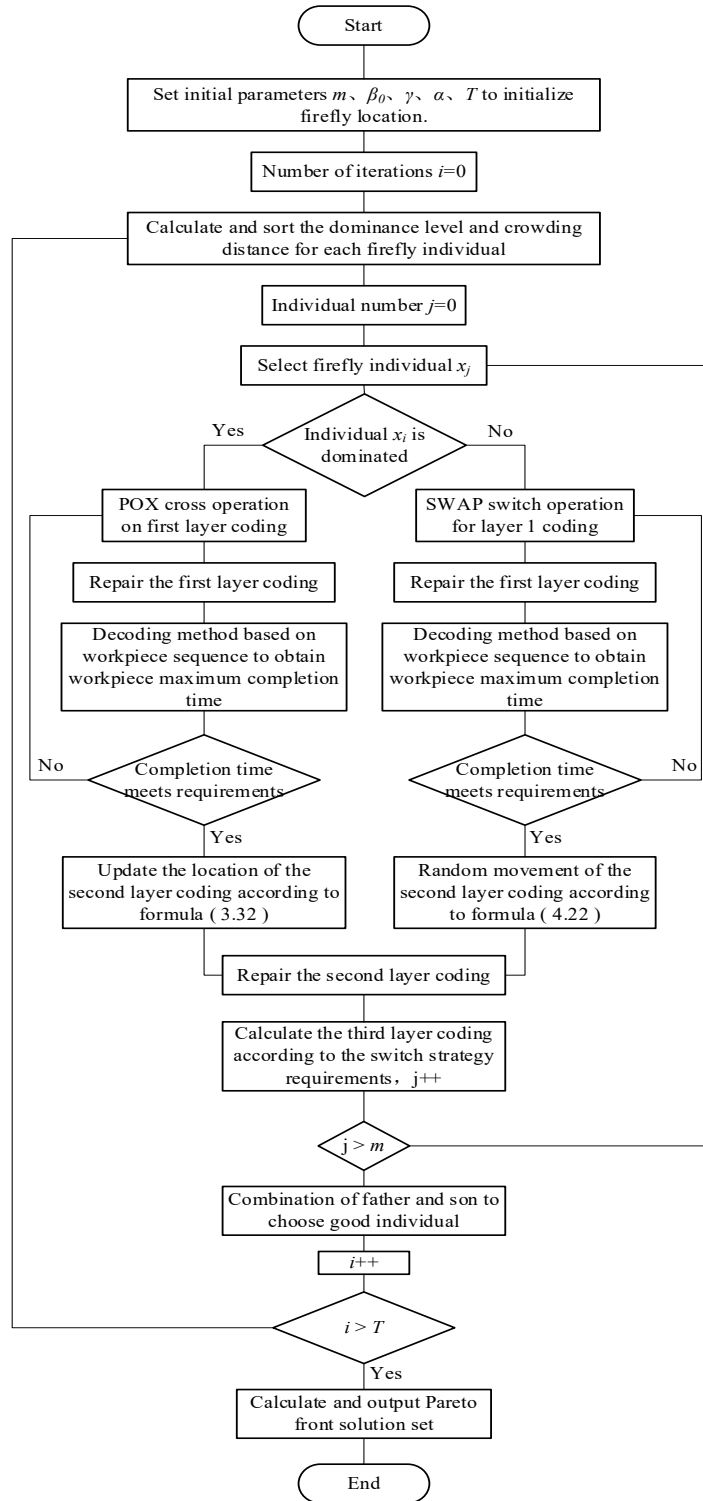


Fig. 2 Flowchart of integrated scheduling algorithm for machining and assembly

4. Results and discussion of case study

4.1 Case study results

The processing-assembly manufacturing enterprise has a mixed-flow production line for processing and assembling, taking into account the time-sharing step price. After creating standard parts in other processing workshops, the personalized core parts are produced and the products are assembled on the mixed-flow production line. For an existing product A, the product structure, as depicted in Figure 3, consists of 8 parts and 6 assembly tasks.

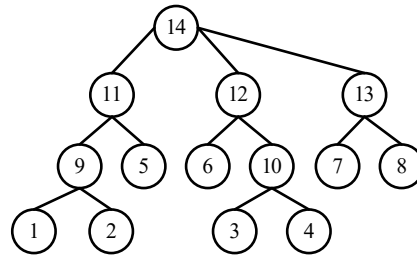


Fig. 3 Product structure diagram

Each workpiece undergoes five processing steps, requiring passage through five processing machines. The processing equipment and time required for each processing step of the workpiece are detailed in Table 1. Six assembly tasks are carried out using three assembly machines. The assembly equipment and time required for each assembly task are outlined in Table 2. Table 3 provides information on the standby power, processing power, shutdown restart power, and time for both the processing and assembly equipment, with the public equipment having a power rating of 5. Assembly tasks must commence after the corresponding processing steps are completed, and assembly and processing tasks can be conducted concurrently.

Table 1 Corresponding processing equipment and processing time for workpiece processing

Workpiece serial number	(Equipment sequence, processing time /min)				
1	(3,73)	(5,68)	(2,64)	(1,64)	(4,79)
2	(3,76)	(2,78)	(5,64)	(1,63)	(4,77)
3	(2,74)	(1,66)	(4,61)	(3,66)	(5,81)
4	(1,75)	(4,62)	(2,69)	(3,61)	(5,79)
5	(2,73)	(1,65)	(5,64)	(3,63)	(4,79)
6	(1,79)	(4,65)	(2,68)	(5,68)	(3,84)
7	(1,76)	(3,67)	(5,69)	(2,72)	(4,75)
8	(2,78)	(1,68)	(3,73)	(4,65)	(5,76)

Table 2 Assembly equipment and assembly time corresponding to the assembly process

Assembly task	9	10	11	12	13	14
Equipment sequence	6	6	7	7	7	8
Length of assembly /min	119	109	119	128	111	115

Table 3 Some parameters of processing and assembly equipment

	Equipment 1	Equipment 2	Equipment 3	Equipment 4
Processing power (kW)	25	28	27	31
Standby power (kW)	4	5	4	3
Shutdown restart time (min)	20	25	30	38
Shutdown restart energy consumption (kWh)	6	9	6	5
	Equipment 5	Equipment 6	Equipment 7	Equipment 8
Processing power (kW)	30	40	45	42
Standby power (kW)	4	5	6	5
Shutdown restart time(min)	40	45	48	55
Shutdown restart energy consumption (kWh)	6	10	14	10

Referring to the current time-sharing electricity price and residential ladder electricity price in Shaanxi Province, assuming that there are two electricity prices in each period, this paper constructs the time-sharing ladder electricity price model. The first stage is normal electricity consumption, and the second stage is high standard electricity consumption. It is concluded that the time-sharing electricity price is shown in Table 4 below. The cut-off point of the step price is 60 kWh per hour, starting at 7 a.m. and requiring completion within one day.

Table 4 Time-sharing tier electricity price period and price list

	Time periods	First-grade electricity price	Second-grade electricity price
Peak hours	8:00-11:00 18:00-23:00	0.9831	1.3131
Normal period	7:00-8:00 11:00-18:00	0.6712	1.0012
Low valley period	23:00-7:00	0.3594	0.6894

According to the model established in section 2 of this paper, the MATLAB algorithm is written, and the program is run on the MATLAB R2014 b version. The population size is 100, the light intensity absorption coefficient γ is 0.5, the maximum attraction β_0 is 1, the step factor α is 0.5, and the number of iterations is 500. The Pareto solution set is obtained, as shown in Figure 4. Statistics were conducted on the target values corresponding to the solutions in the Pareto set, and the results are shown in Table 5.

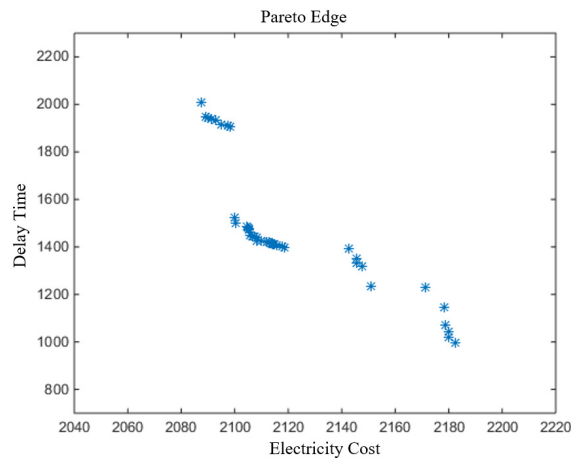


Fig. 4 Pareto Frontier

Table 5 Pareto solution set statistical results corresponding to each target value

Electricity cost (Yuan)			Delay time (min)		
Maximum value	Minimum value	Average value	Maximum value	Minimum value	Average value
2182.4	2087.6	2121.2	2008	997	1472.5

The Gantt chart for the single best solution is depicted in Figures 5 and 6 within the Pareto solution set.

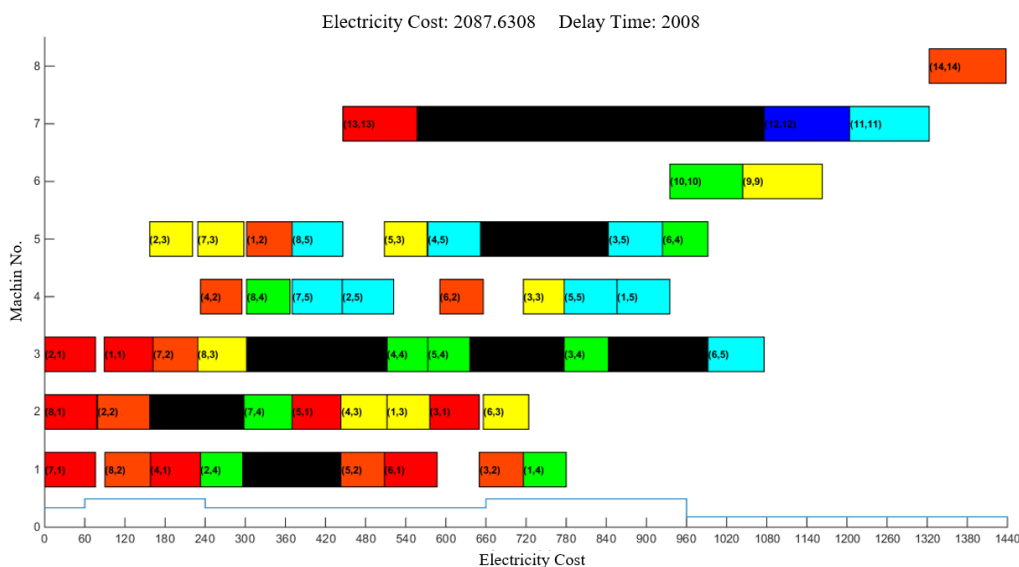


Fig. 5 Pareto solution to the concentrated Gantt chart for optimal power cost

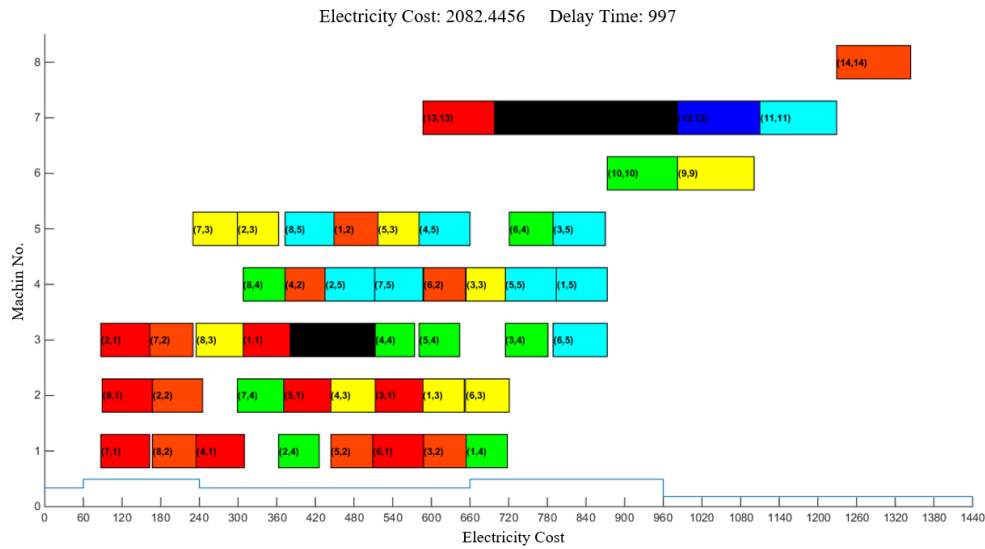


Fig. 6 Pareto solution to the concentrated Gantt chart for optimal delay time

It can be seen from Table 5 and Fig. 4 that the single pursuit of the optimal power cost or the optimal delay time of processing and assembly will lead to another goal too high. When the optimal power cost is 2087, the delay time reaches the high value of 2008, and when the minimum delay time is 997, the power cost reaches the high value of 2182. Therefore, multi-objective integrated scheduling is necessary. From the results of the algorithm, a set of scheduling schemes is obtained for this instance. The scheme set is non-dominated, and the decision makers can select the most suitable implementation scheme according to certain subjective preferences.

As shown in the Gantt chart, the implementation of the switch strategy reduces the power cost in the processing process, and the parallel processing and assembly effectively reduces the waiting time of the components, and completes the assembly task in time in the idle time of the equipment. So as to improve production efficiency and achieve the purpose of energy saving and punctual production.

4.2. Discussion

In the 21st century, manufacturing enterprises cannot overlook the issue of energy-efficient production. The scheduling problem, which takes energy consumption into account, is more intricate and involves a wider range of influencing factors compared to traditional production planning and scheduling. This study focuses on the integrated scheduling problem of machining and assembly in manufacturing enterprises. While this is a conventional issue in production and manufacturing, the study incorporates the time-sharing ladder price and switch strategy to establish a new optimization goal and a comprehensive scheduling model for machining and assembly with the aim of conserving energy. The firefly algorithm is chosen to solve the model, and by adjusting certain aspects, the results are quickly obtained. The example results demonstrate that this method, which considers the parallelism of processing and assembly processes, can effectively utilize equipment resources and reduce power loss during idle equipment, ultimately achieving punctual production and energy savings.

5. Conclusion

This paper considers time-of-use tiered electricity pricing and focuses on integrated scheduling of processing and assembly for mixed flow production lines in processing assembly firms. Establish a mathematical model for this problem and introduce shutdown and restart procedures when the device is idle with the optimization goal of minimizing power consumption costs and delay time during the machining and assembly process. A multi-objective firefly technique based on three-layer encoding was created for the ease of solving the model. Lastly, examples were used to confirm the accuracy of the solving method and the efficacy of the constructed mathematical

model. Empirical evidence indicates that the approach put forward in this work yields certain advantages over conventional scheduling techniques for mixed flow production lines that combine processing and assembly. In addition, during peak and off peak hours, the on/off strategy adopted based on the time of use tiered electricity price reduces standby time of equipment, lowers power costs, and achieves the goal of on-time production and energy conservation for enterprises. On the one hand, it reduces waiting delays for components during processing and assembly, and improves production efficiency.

Owing to space constraints, this article ignores the comparison of various techniques and instead concentrates on solving model solving problems. To confirm their superiority, the solution algorithms will be compared and examined in further studies.

Acknowledgments

This research is supported by Doctoral Research Initiation Fund Project of Nanyang Institute of Technology.

Reference

- [1] Meng, L.L., Zhang, C.Y., Xiao, H.J., Zhan, X.L., Luo, M. (2019). Mathematical modeling of energy-efficient flexible job shop scheduling problem with controllable processing times, *Computer Integrated Manufacturing Systems*, Vol. 25, No. 5, 1062-1074, doi: [10.13196/j.cims.2019.05.004](https://doi.org/10.13196/j.cims.2019.05.004).
- [2] Mazlan, N.A.S., Nawawi, M.N., Saputra, J., Muhamad, S.B., Abdullah, R. (2022). Classification of attributes on green manufacturing practices: A systematic review, *International Journal of Sustainable Development and Planning*, Vol. 17, No. 6, 1839-1847, doi: [10.18280/ijstdp.170618](https://doi.org/10.18280/ijstdp.170618).
- [3] Toró, G. (2023). Production of electricity at the European Union level vs. Romania, *Journal of Green Economy and Low-Carbon Development*, Vol. 2, No. 1, 11-18, doi: [10.56578/jgelcd020102](https://doi.org/10.56578/jgelcd020102).
- [4] Kilci, E.N. (2022). Incentives for sustainability: relationship between renewable energy use and carbon emissions for Germany and Finland, *Opportunities and Challenges in Sustainability*, Vol. 1, No. 1, 29-37, doi: [10.56578/ocs010104](https://doi.org/10.56578/ocs010104).
- [5] Li, C.B., Kou, Y., Lei, Y.F., Xiao, Q.G., Li, L.L. (2020). Flexible job shop rescheduling optimization method for energy-saving based on dynamic events, *Computer Integrated Manufacturing Systems*, Vol. 26, No. 2, 288-299, doi: [10.13196/j.cims.2020.02.002](https://doi.org/10.13196/j.cims.2020.02.002).
- [6] Peng, F., Zheng, L. (2023). An improved multi-objective wild horse optimization for the dual resource-constrained flexible job shop scheduling problem: A comparative analysis with NSGA-II and a real case study, *Advances in Production Engineering & Management*, Vol. 18, No. 3, 271-287, doi: [10.14743/apem2023.3.472](https://doi.org/10.14743/apem2023.3.472).
- [7] Huo, L., Wang, J.Y. (2022). Flexible job shop scheduling based on digital twin and improved bacterial foraging, *International Journal of Simulation Modelling*, Vol. 21, No. 3, 525-536, doi: [10.2507/IJSIMM21-3-CO14](https://doi.org/10.2507/IJSIMM21-3-CO14).
- [8] Komaki, G.M., Kayvanfar, V. (2015). Grey Wolf Optimizer algorithm for the two-stage assembly flow shop scheduling problem with release time, *Journal of Computational Science*, Vol. 8, 109-120, doi: [10.1016/j.jocs.2015.03.011](https://doi.org/10.1016/j.jocs.2015.03.011).
- [9] Sun, Z.Y., Han, W.M., Gao, L.L. (2023). Real-time scheduling for dynamic workshops with random new job insertions by using deep reinforcement learning, *Advances in Production Engineering & Management*, Vol. 18, No. 2, 137-151, doi: [10.14743/apem2023.2.462](https://doi.org/10.14743/apem2023.2.462).
- [10] Ren, J.F., Ye, C.M., Li, Y. (2021). A new solution to distributed permutation flow shop scheduling problem based on NASH Q-Learning, *Advances in Production Engineering & Management*, Vol. 16, No. 3, 269-284, doi: [10.14743/apem2021.3.399](https://doi.org/10.14743/apem2021.3.399).
- [11] Tian, W., Zhang, H.P. (2021). A dynamic job-shop scheduling model based on deep learning, *Advances in Production Engineering & Management*, Vol. 16, No. 1, 23-36, doi: [10.14743/apem2021.1.382](https://doi.org/10.14743/apem2021.1.382).
- [12] Zhao, Z.Y., Yuan, Q.L. (2022). Integrated scheduling of the production and maintenance of parallel machine job-shop considering stochastic machine breakdowns, *Journal of Engineering Management and Systems Engineering*, Vol. 1, No. 1, 15-22, doi: [10.56578/jemse010103](https://doi.org/10.56578/jemse010103).
- [13] Wang, D. (2023). Quality management and control for the whole-process logistics service of multi-variety small-batch production and manufacturing, *Journal Européen des Systèmes Automatisés*, Vol. 56, No. 1, 69-76, doi: [10.18280/jesa.560110](https://doi.org/10.18280/jesa.560110).
- [14] Belmahdi, R., Mechta, D., Harous, S. (2021). A survey on various methods and algorithms of scheduling in Fog Computing, *Ingénierie des Systèmes d'Information*, Vol. 26, No. 2, 211-224, doi: [10.18280/isi.260208](https://doi.org/10.18280/isi.260208).
- [15] Ren, J.F., Ye, C.M., Li, Y. (2020). A two-stage optimization algorithm for multi-objective job-shop scheduling problem considering job transport, *Journal Européen des Systèmes Automatisés*, Vol. 53, No. 6, 915-924, doi: [10.18280/jesa.530617](https://doi.org/10.18280/jesa.530617).
- [16] Bedhief, A.O., Dridi, N. (2020). A genetic algorithm for three-stage hybrid flow shop scheduling problem with dedicated machines, *Journal Européen des Systèmes Automatisés*, Vol. 53, No. 3, 357-368, doi: [10.18280/jesa.530306](https://doi.org/10.18280/jesa.530306).

- [17] Zhao, Z.Y., Yuan, Q.L. (2022). Integrated multi-objective optimization of predictive maintenance and production scheduling: perspective from lead time constraints, *Journal of Intelligent Management Decision*, Vol. 1, No. 1, 67-77, doi: [10.56578/jimd010108](https://doi.org/10.56578/jimd010108).
- [18] Deng, C., Qian, B., Hu, R., Wang, L. (2019). Hybrid EDA for three-phase heterogeneous parallel machine assembly integrated scheduling problem, *Information and Control*, No. 5, 552-558, doi: [10.13976/j.cnki.xk.2019.8565](https://doi.org/10.13976/j.cnki.xk.2019.8565).
- [19] Guo, G., Ryan, S.M. (2022). Sequencing mixed-model assembly lines with risk-averse stochastic mixed-integer programming, *International Journal of Production Research*, Vol. 60, No. 12, 3774-3791, doi: [10.1080/00207543.2021.1931978](https://doi.org/10.1080/00207543.2021.1931978).
- [20] Wei, N.-C., Liu, S.-F., Chen, C.-H., Xu, Y.-X., Shih, Y.-Y. (2023). An integrated method for solving the two-sided assembly line balancing problems, *Journal of Advanced Manufacturing Systems*, Vol. 22, No. 1, 181-203, doi: [10.1142/S0219686723500105](https://doi.org/10.1142/S0219686723500105).
- [21] Liang, Y.J., Yang, M.S., Gao, X.Q., Ba, L., Lei, F.D. (2016). Multi-objective optimizing model for solving mixed model shop of manufacturing and assembly, *Computer Engineering and Applications*, Vol. 52, No. 10, 247-253, doi: [10.3778/j.issn.1002-8331.1406-0463](https://doi.org/10.3778/j.issn.1002-8331.1406-0463).
- [22] Boudjemline, A., Chaudhry, I.A., Rafique, A.F., Elbadawi, I.A., Aichouni, M., Boujelbene, M. (2022). Multi-objective flexible job shop scheduling using genetic algorithms, *Tehnički Vjesnik – Technical Gazette*, Vol. 29, No. 5, 1706-1713, doi: [10.17559/TV-20211022164333](https://doi.org/10.17559/TV-20211022164333).
- [23] Sun, H. (2023). Optimizing manufacturing scheduling with genetic algorithm and LSTM neural networks, *International Journal of Simulation Modelling*, Vol. 22, No. 3, 508-519, doi: [10.2507/IJSIMM22-3-C013](https://doi.org/10.2507/IJSIMM22-3-C013).
- [23] Rahman, N.Z.A., Aziz, N.A.A., Abas, F.S., Adnan, R., Hamzah, N.A.A., Khan, C.T., Rosli, N.R. (2022). BLE beacons positioning algorithm using particle swarm optimization for indoor navigation system, *Journal of System and Management Sciences*, Vol. 12, No. 6, 31-49, doi: [10.33168/JSMS.2022.0603](https://doi.org/10.33168/JSMS.2022.0603).
- [24] Ren, C.L., Yang, X.D., Zhang, C.Y., Meng, L., Hong, H., Yu, J. (2019). Modeling and optimization for energy-efficient hybrid flow-shop scheduling problem, *Computer Integrated Manufacturing Systems*, Vol. 25, No. 8, 1965-1980, doi: [10.13196/j.cims.2019.08.011](https://doi.org/10.13196/j.cims.2019.08.011).
- [25] Zhang, B., Pan, Q.-K., Gao, L., Zhang, X.-L., Peng, K.-K. (2019). A multi-objective migrating birds optimization algorithm for the hybrid flowshop rescheduling problem, *Soft Computing*, Vol. 23, 8101-8129, doi: [10.1007/s00500-018-3447-8](https://doi.org/10.1007/s00500-018-3447-8).
- [26] Zhang, Z.C., Liu, S.Y. (2019). Firefly algorithm based on topology improvement and crossover strategy, *Computer Engineering and Applications*, Vol. 55, No. 7, 1-8, doi: [10.3778/j.issn.1002-8331.1812-0263](https://doi.org/10.3778/j.issn.1002-8331.1812-0263).
- [27] Yang, X.-S. (2013). Multiobjective firefly algorithm for continuous optimization, *Engineering with Computers*, Vol. 29, 175-184, doi: [10.1007/s00366-012-0254-1](https://doi.org/10.1007/s00366-012-0254-1).
- [28] Ba, L., Li, Y., Yang, M.S., Liu, Y. (2017). Investigating integrated process planning and scheduling problem with flexible batch splitting considered, *Mechanical Science and Technology for Aerospace Engineering*, Vol. 36, No. 3, 426-435, doi: [10.13433/j.cnki.1003-8728.2017.0317](https://doi.org/10.13433/j.cnki.1003-8728.2017.0317).

Research on recovery strategies of supply chain network under disruption propagation using memetic algorithm

Li, Z.Y.^a, Zhao, P.X.^{b,*}, Wang, C.L.^b, Mi, Y.Z.^b

^aSchool of Public Administration and Policy, Shandong University of Finance and Economics, P.R. China

^bSchool of Management, Shandong University, P.R. China

ABSTRACT

In the context of the economic globalization, there is an increased disruption risk in the supply chain network due to the outsourcing, complexity and uncertainty. At the same time, the disruption may propagate across the entire supply chain network because of the interdependence. With the resource constraints, appropriate recovery strategies which can minimize the impact of disruption propagation and effectively improve the supply chain network resilience have attracted a great deal of attention. In this paper, we first construct the disruption propagation model considering the recovery strategy based on the characteristics of the competitiveness, time delay and underload cascading failure in the supply chain network. This model uses the memetic algorithm to determine the set of recovery nodes among all disruption nodes, which can minimize the impact of disruption propagation. And then, the simulation analysis is conducted on the synthetic network and the real-world supply chain network. We compare the proposed recovery strategy with other strategies (according to the genetic algorithm, according to the descending order of the load of failure node, according to the ascending order of the load of failure node, according to the descending order of the node degree, according to the ascending order of the node degree) and provide decision-making reference against supply chain disruptions.

ARTICLE INFO

Keywords:
Supply chain network;
Disruption propagation;
Recovery strategy;
Memetic algorithm

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

Article history:
Received 13 September 2023
Revised 24 January 2024
Accepted 8 February 2024



Content from this work may be used under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

In recent years, public health events, global transportation network congestion and large-scale natural disasters have led to frequent disruptions in the supply chain network. Moreover, with the deepening of division of labour and cooperation, supply chains are becoming more global and intertwined. The disruption events may spread in the supply chain network and even cause ripple effects [1, 2]. In 2020, coronavirus-driven supply chain disruptions affected 94 % of the Fortune 1000 companies [3]. On March 23 2021, the Suez Canal was blocked by vast container ship – the Ever Given, which affected over 400 vessels and held up about USD15 billion to USD17 billion [4]. Many researchers have studied to construct agile [5, 6], sustainable [7], and resilient [3] supply chains to withstand disruption risks.

Many industries have increasingly begun to implement supply chain management, including manufacturing, service and so on [8]. Disruption events may cascade through the supply chain resulting in disruption propagation [9] and have a powerful impact on most economic sectors [10]. Other terms related to supply chain disruption propagation in literature include cascading failure [11], ripple effect [2], risk diffusion [12] and so on. Constructing disruption propagation model can dynamically analyse the propagation process in the supply chain network and find the

critical firms more accurately in the disruption events. Some studies use the method of complex network [13] to analyse the disruption propagation of supply chain network based on the cascading failure model [9, 14], which commonly used in power, transportation and infrastructure disruption. However, the disruption propagation process in the supply chain network is different from that in infrastructure network because of the characteristics of competition [13, 15], time delay [16], underload failure [17] and adaptivity [13]. Wang and Zhang [17] innovatively used underload failure instead of overload failure to analyse the propagation process of supply network disruption, and constructed synthetic network to conduct a numerical simulation. Zhao *et al.* [13] created a real-world network and competition supply chain network provided by Mergent, and simulated how disruptions propagate in the supply chain network through cascading failures.

What's more, scope of the disruption and its performance impact rely on the speed and scale of recovery strategies [18]. Some studies search for the recovery firms according to characteristics of complex networks, such as the betweenness centrality [9] and degree centrality of nodes [16]. Wang and Xiao [19] developed a resilience method to cascading failures in cluster supply chain network using social resilience of ant colony. They compared the generated random number with the recovery probability to determine whether the node restores timely. Fu *et al.* [9] divided the recovery process into 3 kinds of situations and compared the effect of dynamic recovery strategies which were in descending and ascending orders of node degrees and betweenness centrality. Jing and Tang [16] designed the recovery probability is related to the nodes' degree. However, the above recovery strategies are qualitative methods, and there are few studies using optimization methods to quantitatively recover the critical firms in the supply chain network.

In this paper, we construct a disruption propagation model based on the characteristics of the competitiveness, time delay and underload cascading failure in the supply chain network. Then, considering the restrictions from recovery resources, we use memetic algorithm to quantitatively search for the critical enterprises to determine the recovery strategies, so as to more effectively reduce the impact of supply chain disruptions. The remaining sections of this paper is organized as follows. In Section 2, we construct a supply chain network disruption propagation model considering the recovery strategies. The numerical simulation is conducted on the synthetic network and the real-world supply chain in Section 3. The paper delivers a brief conclusion in Section 4.

2. Simulation model

This section constructs a supply chain network disruption propagation model considering the recovery strategies and the whole process is shown in Fig. 1. This study analyses the above process from three parts: main metrics of supply chain network, recovery strategies, and disruption propagation process. The specific process can be described as follows.

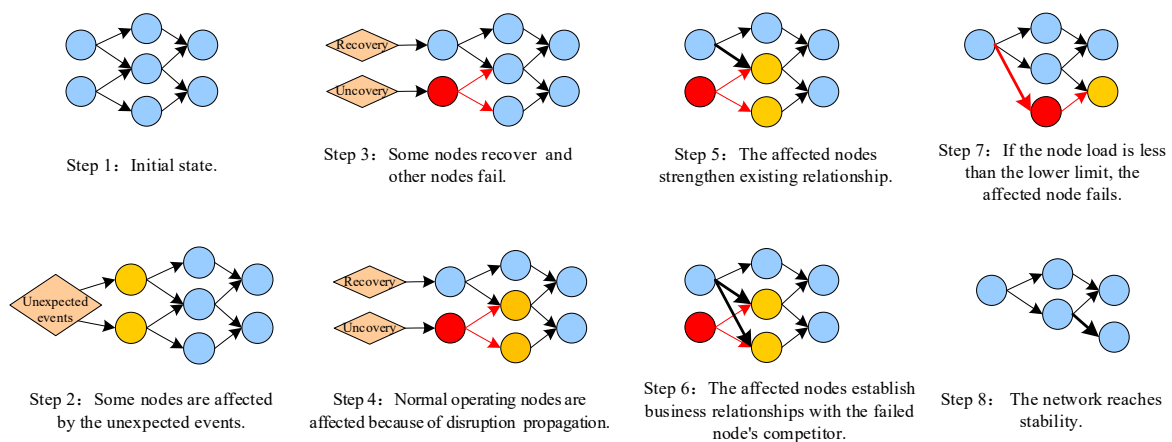


Fig. 1 Process of disruption propagation considering the recovery strategies (red nodes denote failed nodes, yellow nodes denote affected nodes, and blue nodes denote normal operating nodes)

2.1 Main metrics of supply chain network

This section mainly determines the initial node load, capacity and resilience measurement of the supply chain network. The node load represents the operation scale of supply chain members. At present, the initial load of a node is generally estimated by the node degree [20], the total number of shortest paths [21, 22], eigenvector centrality [15] and node degree multiplied by the neighbour node degree[17]. Actually, the business scale of a firm is not only related to the number of neighbour firms, but also to the importance of neighbour firms. Eigenvector centrality of the node load is not only related to the number of neighbour nodes, but also to the importance of neighbour nodes. Therefore, we use eigenvector centrality to measure the initial load of the node i , that is $L_i^0 = c \sum_{j=1}^n a_{ij}L_j^0$, where c is the proportional constant, and a_{ij} is the value of row i and column j in the adjacency matrix A . If node i is connected to node j in the graph, 1; otherwise, 0.

There are upper and lower limits to a firm’s capacity, which are determined by firm’s competitiveness. According to the reference [17], the upper limit of node i load capacity $C_{i(max)}$ is defined as $C_{i(max)} = \alpha L_i^0$ and α is the upper limit parameter. The lower limit of node i load capacity $C_{i(min)}$ is defined as $C_{i(min)} = \beta L_i^0$, and β is a lower limit parameter.

Supply chain network resilience (RN) is a network attribute, which refers to its ability to resist disruption [23]. There are currently a variety of measurement metrics for the resilience, including the size of the network (total number of nodes [24], size of the largest functional sub-network [25], density [24, 26]), network availability (supply availability rate [27], the proportion of suppliers [24]), network diameter (average shortest-path length [28], average supply-path length [27]) and centrality [24, 29] (betweenness centrality, freeman centralization and eigenvector centrality).

In real-world supply chain networks, the normal operation of downstream members may depend on the operation of suppliers. Therefore, this section chooses the size of the largest functional sub-network (LFSN) as a metric, which differs from the largest connected component (LCC) in that there must be at least one supply node in LFSN according to the reference [27]. That is, $RN = N^t$, where N^t is the total number of nodes in LFSN at time t .

2.2 Recovery strategies

The external environment and core enterprises often take strategies to help the affected enterprises against unexpected events. However, the resources of the external environment are limited, which can only help some enterprises to recover initial normal operation and some enterprises are unrecovered. The unrecovered enterprises will still affect the upstream and downstream enterprises in the supply chain until the network reaches a stable state. Recovery resource C_R refers to the sum of available recovery resources. Therefore, constraint $\sum_{i=1}^N L_i^0 \leq C_R$ need to be satisfied, where N is the number of restored nodes. As shown in Fig. 2, different recovery strategies lead to the different network resilience when it reaches stability.

In order to maximize the resilience of the supply chain network, this section analyses the disruption propagation and takes RN after the network reaches stability as the objective. We obtain the optimal recovery strategy through the memetic algorithm under the capacity constraint of nodes and the recovery resource constraint. Compared with other algorithms, memetic algorithm [30, 31] can not only retain the advantages of genetic algorithm, but also improve the efficiency of local search. The specific process is as follows.

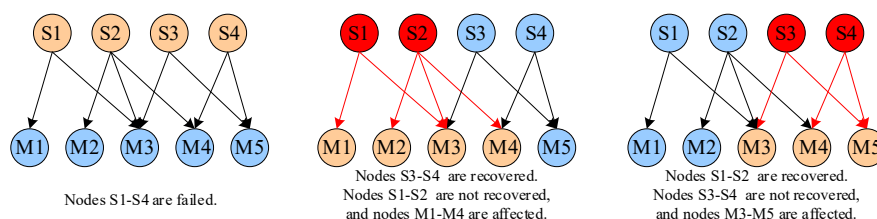


Fig. 2 Supply chain network resilience under different recovery strategies

(2) *Cascading failure.* After a node fails, the effect of neighbour nodes is related to the closeness of it. If node i fails at time t , its upstream and downstream neighbor nodes will be affected and the load of neighbor node j reduces to $L_j^{t+1} = L_j^t - \Delta L_{ij}^{t+1}$, where $\Delta L_{ij}^{t+1} = \min(\delta_{ij} L_i^t, L_j^t)$. δ_{ij} is the strength of relationship between the nodes i and j , where $\delta_{ij} = \frac{W_{ij}}{\sum_{j \in \Gamma_i(out)} W_{ij}}$, $\Gamma_i(out)$ refers to the set of downstream neighbour nodes of node i , $W_{ij} = (L_i^0 L_j^0)^\tau$ and τ is a constant. The effect on the upstream nodes is the same as that on the downstream nodes.

(3) *Adjusting business relationships.* After the cascading effect occurs, there are two strategies for the upstream and downstream neighbour nodes to adjust business relationships. One is to strengthen the existing business relationships [15, 17, 19], and the other is to establish new business relationships with the competing nodes of the failed nodes with a certain probability [13, 17]. For example, if a node i fails which supplying node j , the node j hopes to establish new connections with node k which competing with node i . Considering its own load, node k will establish a new relationship with node j with a certain probability P_t . Then the load of node j changes to $L_j^{t+2} = L_j^{t+1} + \sum_{k \in \Gamma_j(in)} \Delta L_{kj}^{t+2} + \sum_{k \in \Gamma_i} M_j \Delta L_{kj}^{t+2}$, where ΔL_{kj}^{t+2} denotes the increase of load after node j strengthens the existing business relationships with node k or establishes a new relationship with node k , $\Gamma_j(in)$ refers to the set of upstream neighbour nodes of node j , Γ_i refers to the set of the competing nodes of node i , and $M_j \in \{0,1\}$. A random number R in range $(0, 1)$ is generated to compare with P_t . If $R > P_t$, $M_j = 1$ that is establishing a new relationship. Otherwise, $M_j = 0$. When the downstream node i of node j fail, the load of node j changes to $L_j^{t+2} = L_j^{t+1} + \sum_{k \in \Gamma_j(out)} \Delta L_{jk}^{t+2} + \sum_{k \in \Gamma_i} M_j \Delta L_{jk}^{t+2}$, $\Gamma_j(out)$ refers to the set of downstream neighbor nodes of node j , and the other process is the same as the upstream node failure of node j .

If the load of node j at time $t + 2$ is less than the lower limit of node load capacity, that is, $L_j^{t+2} < C_{j(min)}$, node j also fails and affects its neighbor nodes. We loop Stages 2–3 until no node failures occur.

3. Example application

3.1 Supply chain network and competition networks

This paper constructs two directed supply chain networks for simulation analysis including a real-world supply chain network with 37 nodes according to literature [32] and a scale-free network with 1000 nodes. The first network includes 11 suppliers, 3 factories, 5 warehouses and 18 markets and the second network includes 375 suppliers, 18 manufacturers, 16 wholesalers and 591 retailers as shown in Fig. 4. The competition network is constructed for the above two networks, in which the competition network is undirected.

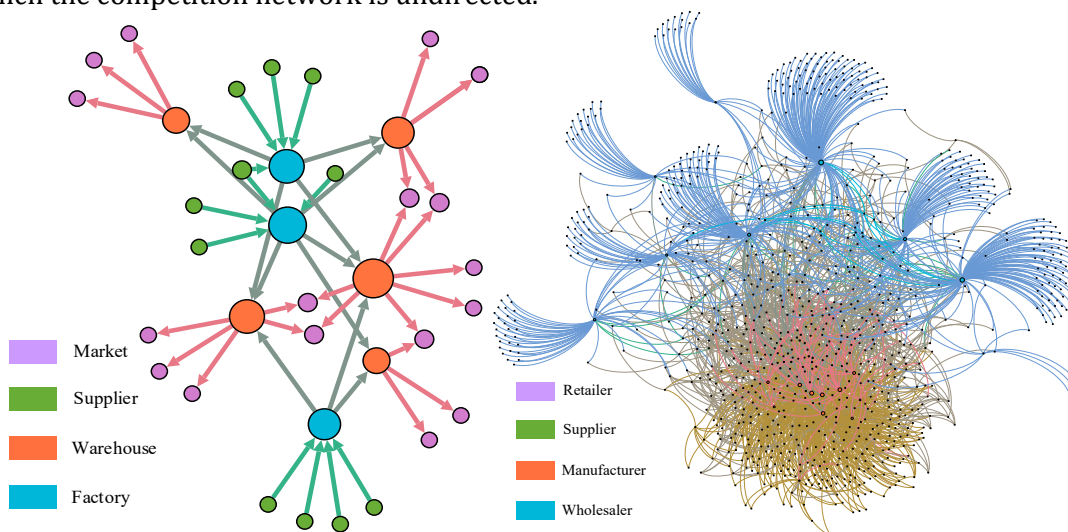


Fig. 4 Visualization of the network

$G_r(V, E)$ denotes the real-world supply chain network and $G_s(V, E)$ denotes the synthetic supply chain network. Competition network of the real-world supply chain network is denoted by $G_{r'}(V, E)$ and competition network of the synthetic supply chain network is denoted by $G_{s'}(V, E)$. $v_i \in V$ denotes the node i in the network and $e_{ij} \in (G_r \cup G_s)$ represents the directed edge from v_i to v_j , indicating that v_i is a supplier of v_j . $e'_{ij} \in (G_{r'} \cup G_{s'})$ represents the undirected edge between v_i and v_j , indicating that v_i and v_j are competitive.

3.2 Performances of memetic algorithm

In this section, the initial population size is set to 30. We simulate the supply chain network resilience under different recovery resources and disruption scenarios. According to Fig. 5(a-c), when the recovery resource is set to 50, the supply chain network resilience becomes smaller as the number of nodes removal increases. According to Fig. 5(d-f), when the number of nodes removal is set to 50, the supply chain network resilience becomes larger as the recovery resource increases. The above six simulations all show that with the increase of iterations, the results tend to be stable which verifies the effectiveness of the memetic algorithm.

As shown in Fig. 5, in most cases, the convergence results obtained by memetic algorithm are similar to those obtained by genetic algorithm, but the convergence speed and initial value of memetic algorithm are better than those of memetic algorithm. This indicates that the memetic algorithm converges faster and can achieve the optimal result through fewer iterations.

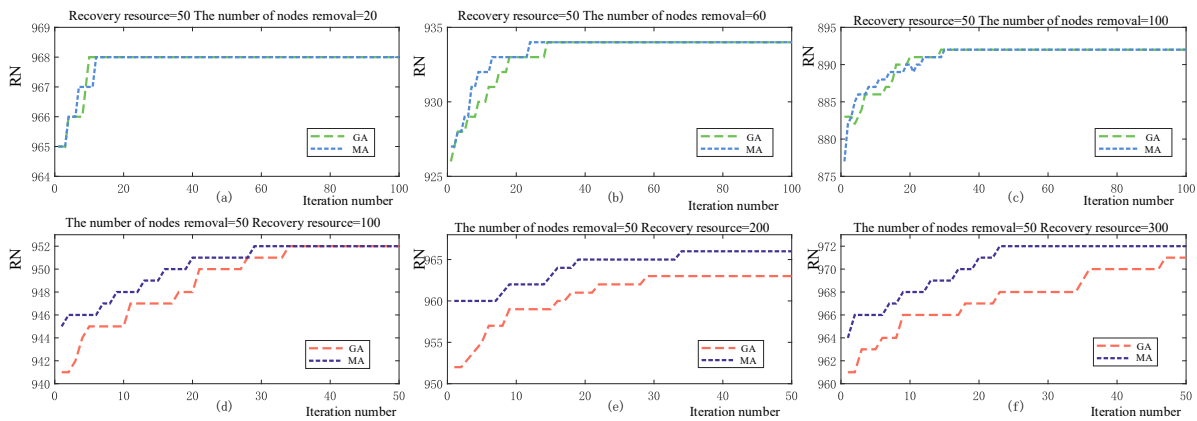


Fig. 5 The relationship between RN and iteration number

3.3 The relationship between network resilience and recovery resources

Fig. 6(a-c) shows the change of the synthetic supply chain network resilience with the increase of the recovery resources against different nodes removal. Fig. 6(d-f) shows the change of the real-world supply chain network resilience with the increase of the recovery resources against different nodes removal. It can be seen that although the change rules of the supply chain network resilience with different structures are not exactly the same, they all conform to the trend that the network resilience increases slowly with the increase of recovery resources. When $C_R = 0$, $RN = 926$, when $C_R = 50$, $RN = 943$, and $\Delta RN = 17$. When $C_R = 100$, $RN = 952$ and $\Delta RN = 9$. These results suggest that the relationship between network resilience and recovery resources also presents the characteristics of diminishing marginal utility. When the recovery resources are gradually increased, the increase of network resilience becomes slower. If the supply chain network is fully restored, it needs large amount of recovery resources.

As shown in Fig. 6(a-c), $RN = 926$ when the number of nodes removal is 50 and $C_R = 0$. $RN = 685$ when the number of nodes removal is 100 and $C_R = 0$. $RN = 13$ when the number of nodes removal is 150 and $C_R = 0$. The results indicate that if we don't take any recovery resources, the supply chain network resilience drops dramatically with the increase of the number of nodes removal. Therefore, the importance of recovery strategies under large-scale unexpected events is much higher than that under small-scale unexpected events. When a large-scale unexpected event occurs, effective recovery strategies can reduce the possibility of supply chain network crash.

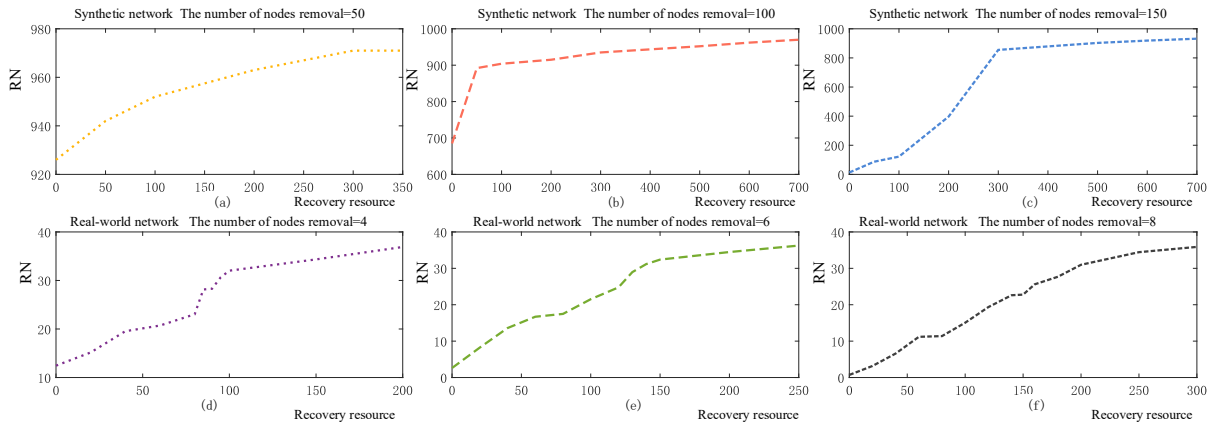


Fig. 6 The relationship between network resilience and recovery resources

4. Comparison of different recovery strategies

We compare the different recovery strategies, including strategies by memetic algorithm (strategy 1), according to the ascending order of the load of failure node (strategy 2), according to the descending order of the load of failure node (strategy 3), according to the ascending order of the node degree (strategy 4), according to the descending order of the node degree (strategy 5). As shown in Fig. 7, the recovery strategy obtained by using the memetic algorithm is always superior to other strategies, and the advantage tends to be obvious with the increase of the number of nodes removal.

According to the simulation results, the recovery strategy obtained through the memetic algorithm is closer to that according to the ascending order when the number of nodes removal is small. For example, when the number of nodes removal is 20, $RN = 968$ obtained by strategy 1; $RN = 968$ obtained by strategy 2; $RN = 956$ obtained by strategy 3; $RN = 964$ obtained by strategy 4; $RN = 956$ obtained by strategy 5. The results indicate that priority will be given to recover small-scale enterprises when the number of attacked enterprises is small.

However, the recovery strategies according to the ascending order no longer have an advantage when the number of nodes removal becomes larger. For example, as shown in Fig. 7(b), when the number of nodes removal is 120, $RN = 895$ obtained by strategy 1; $RN = 384$ obtained by strategy 2; $RN = 531$ obtained by strategy 3; $RN = 92$ obtained by strategy 4; $RN = 188$ obtained by strategy 5. The results indicate that the recovery strategy obtained by memetic algorithm has greater advantage, and the recovery strategies based on the node degree have the worst resilience.

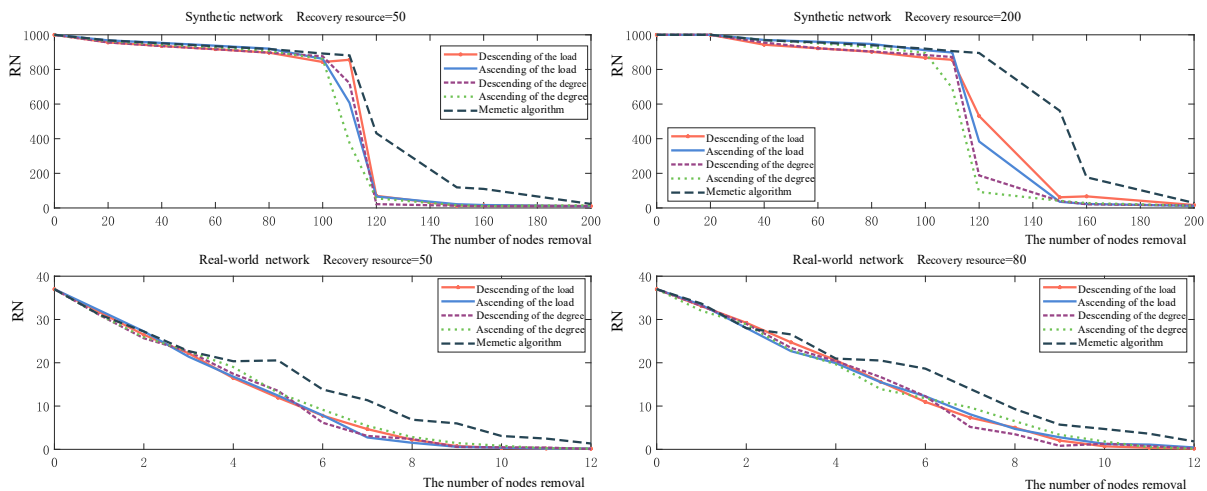


Fig. 7 Comparison of different recovery strategies

As shown in Fig. 7, when the number of nodes removal is 120 and the recovery resource is 200, The order of RN under different recovery strategies is that the memetic algorithm is better than the algorithm according to the order of the load of failure node, and the algorithm according to the order of the load of failure node is better than that of the degree of failure node. The restored nodes of different recovery strategies as shown in Table 1. The recovery strategy obtained by memetic algorithm can recover all kinds of enterprises in the supply chain network. The above results suggest that when a large number of nodes are removed, the strategies of recovery small-scale enterprises first and large-scale enterprises first are not optimal. The optimal strategy is to comprehensively analyse the characteristics of the business scale and supply-demand relationships of the disrupted enterprises, and to recover all kinds of enterprises in the supply chain network in a balanced way.

Table 1 Comparison of different recovery strategies

C_R	The number of nodes removal	Recovery strategies	The characteristic of restored nodes					
			Number	S:M:W:R	S (%)	M (%)	W (%)	R (%)
200	120	Memetic algorithm	44	16:0:1:27	31.4	0	50	42.9
		Genetic algorithm	46	15:2:2:27	29.4	50	100	42.9
		The ascending order of the load	62	14:2:1:45	27	50	50	71.4
		The descending order of the load	6	5:0:0:1	9.8	0	0	1.5
		Ascending order of the node degree	40	8:0:0:32	15.7	0	0	50.8
		Descending order of the node degree	9	5:2:2:0	9.8	50	100	0

Note: S:M:W:R refers to the ratio of the number of restored nodes of suppliers, manufacturers, wholesalers and retailers; S refers to the proportion of restored supplier nodes to disrupted supplier nodes; M refers to the proportion of restored manufacturer nodes to disrupted manufacturer node; W refers to the proportion of restored wholesaler nodes to disrupted wholesaler node; R refers to the proportion of restored retailer nodes to disrupted retailer node.

5. Conclusion

We innovatively construct the disruption propagation model considering the recovery strategy based on the characteristics of the competitiveness, time delay and underload cascading failure in the supply chain network. This model uses the memetic algorithm to determine the set of recovery nodes among all disruption nodes, which can minimize the impact of disruption propagation. The conclusions are as follows:

- The relationship between network resilience and recovery resources presents the characteristics of diminishing marginal utility. When the recovery resources are gradually increased, the increase of network resilience becomes slower. If the supply chain network is fully restored, it needs large amount of recovery resources.
- The recovery strategy obtained by memetic algorithm has greater advantage compared with other recovery strategies. What's more, the advantage tends to be obvious with the increase of the number of attacked enterprises.
- The priority will be given to recover small-scale enterprises when the number of attacked enterprises is small with a certain amount of recovery resources. However, this strategy no longer has an advantage when the number of attacked enterprises becomes larger. The strategy of recovery suppliers first is better than that of recovery retailers, but the optimal strategy is to comprehensively analyse the characteristics of the business scale and supply-demand relationships of the disrupted enterprises, and to recover all kinds of enterprises in the supply chain network in a balanced way.

Acknowledgments

This research was supported by the Social Science Planning Research Project of Shandong Province [grant numbers 20CGLJ11].

References

- [1] Ivanov, D., Pavlov, A., Sokolov, B. (2014). Optimal distribution (re)planning in a centralized multi-stage supply network under conditions of the ripple effect and structure dynamics, *European Journal of Operational Research*, Vol. 237, No. 2, 758-770, doi: [10.1016/j.ejor.2014.02.023](https://doi.org/10.1016/j.ejor.2014.02.023).
- [2] Duan, W., Ma, H., Xu, D.S. (2021). Analysis of the impact of COVID-19 on the coupling of the material flow and capital flow in a closed-loop supply chain, *Advances in Production Engineering & Management*, Vol. 16, No. 1, 5-22, doi: [10.14743/apem2021.1.381](https://doi.org/10.14743/apem2021.1.381).
- [3] Ivanov, D., Dolgui, A. (2020). Viability of intertwined supply networks: Extending the supply chain resilience angles towards survivability. A position paper motivated by COVID-19 outbreak, *International Journal of Production Research*, Vol. 58, No. 10, 2904-2915, doi: [10.1080/00207543.2020.1750727](https://doi.org/10.1080/00207543.2020.1750727).
- [4] Khan, I.A., Rahman, S. (2021). Review and analysis of blockage of Suez canal region due to giant container ship, *Marine Technology Society Journal*, Vol. 55, No. 5, 39-43, doi: [10.4031/MTSJ.55.5.5](https://doi.org/10.4031/MTSJ.55.5.5).
- [5] Rasi, R.E., Hatami, D. (2019). Environmental risk and innovation in supply chain: Analysis of influence of supply chain agility, *Journal of System and Management Sciences*, Vol. 9, No. 3, 1-25, doi: [10.33168/ISMS.2019.0301](https://doi.org/10.33168/ISMS.2019.0301).
- [6] Kim, M., Chai, S. (2022). The role of agility in responding to uncertainty: A cognitive perspective, *Advances in Production Engineering & Management*, Vol. 17 No. 1, 57-74, doi: [10.14743/apem2022.1.421](https://doi.org/10.14743/apem2022.1.421).
- [7] Moghaddas, Z., Vaez-Ghasemi, M., Hosseinzadeh Lotfi, F. (2021). A novel DEA approach for evaluating sustainable supply chains with undesirable factors, *Economic Computation and Economic Cybernetics Studies and Research*, Vol. 55, No. 2, 177-192, doi: [10.24818/18423264/55.2.21.11](https://doi.org/10.24818/18423264/55.2.21.11).
- [8] Kim, Y.-J., Ha, B.-C. (2022). Logistics service supply chain model, *Journal of Logistics, Informatics and Service Science*, Vol. 9, No. 3, 284-300, doi: [10.33168/LISS.2022.0320](https://doi.org/10.33168/LISS.2022.0320).
- [9] Fu, C., Wang, Y., Wang, X.-Y. (2017). Research on complex networks' repairing characteristics due to cascading failure, *Physica A: Statistical Mechanics and its Applications*, Vol. 482, 317-324, doi: [10.1016/j.physa.2017.04.086](https://doi.org/10.1016/j.physa.2017.04.086).
- [10] Posedaru, B.-S., Bologa, R., Toma, A., Pantelimon, F.-V. (2022). The influence of Covid-19 pandemic on online retail prices, *Economic Computation and Economic Cybernetics Studies and Research*, Vol. 56, No. 1, 289-304, doi: [10.24818/18423264/56.1.22.18](https://doi.org/10.24818/18423264/56.1.22.18).
- [11] Zeng, Y., Xiao, R. (2014). Modelling of cluster supply network with cascading failure spread and its vulnerability analysis, *International Journal of Production Research*, Vol. 52, No. 23, 6938-6953, doi: [10.1080/00207543.2014.917769](https://doi.org/10.1080/00207543.2014.917769).
- [12] Basole, R.C., Bellamy, M.A. (2014). Supply network structure, visibility, and risk diffusion: A computational approach, *Decision Sciences*, Vol. 45, No. 4, 753-789, doi: [10.1111/decl.12099](https://doi.org/10.1111/decl.12099).
- [13] Ledwoch, A., Brintrup, A., Mehnen, J., Tiwari, A. (2018). Systemic risk assessment in complex supply networks, *IEEE Systems Journal*, Vol. 12 No. 2, 1826-1837, doi: [10.1109/ISYST.2016.2596999](https://doi.org/10.1109/ISYST.2016.2596999).
- [14] Watts, D.J. (2002). A simple model of global cascades on random networks, *Applied Mathematics*, Vol. 99, No. 9, 5766-5771, doi: [10.1073/pnas.082090499](https://doi.org/10.1073/pnas.082090499).
- [15] Li, Z., Zhao, P., Han, X. (2022). Agri-food supply chain network disruption propagation and recovery based on cascading failure, *Physica A: Statistical Mechanics and its Applications*, Vol. 589, Article No. 126611, doi: [10.1016/j.physa.2021.126611](https://doi.org/10.1016/j.physa.2021.126611).
- [16] Jing, K., Du, X., Shen, L., Tang, L. (2019). Robustness of complex networks: Cascading failure mechanism by considering the characteristics of time delay and recovery strategy, *Physica A: Statistical Mechanics and its Applications*, Vol. 534, Article No. 122061, doi: [10.1016/j.physa.2019.122061](https://doi.org/10.1016/j.physa.2019.122061).
- [17] Wang, Y., Zhang, F. (2018). Modeling and analysis of under-load-based cascading failures in supply chain networks, *Nonlinear Dynamics*, Vol. 92, No. 3, 1403-1417, doi: [10.1007/s11071-018-4135-z](https://doi.org/10.1007/s11071-018-4135-z).
- [18] Ivanov, D., Dolgui, A., Sokolov, B., Ivanova, M. (2017). Literature review on disruption recovery in the supply chain, *International Journal of Production Research*, Vol. 55, No. 20, 6158-6174, doi: [10.1080/00207543.2017.1330572](https://doi.org/10.1080/00207543.2017.1330572).
- [19] Wang, Y., Xiao, R. (2016). An ant colony based resilience approach to cascading failures in cluster supply network, *Physica A: Statistical Mechanics and its Applications*, Vol. 462, 150-166, doi: [10.1016/j.physa.2016.06.058](https://doi.org/10.1016/j.physa.2016.06.058).
- [20] Duan, D.-L., Ling, X.-D., Wu, X.-Y., OuYang, D.-H., Zhong, B. (2014). Critical thresholds for scale-free networks against cascading failures, *Physica A: Statistical Mechanics and its Applications*, Vol. 416, 252-258, doi: [10.1016/j.physa.2014.08.040](https://doi.org/10.1016/j.physa.2014.08.040).
- [21] Crucitti, P., Latora, V., Marchiori, M. (2004). Model for cascading failures in complex networks, *Physical Review E*, Vol. 69, No. 4, Article No. 045104, doi: [10.1103/PhysRevE.69.045104](https://doi.org/10.1103/PhysRevE.69.045104).
- [22] Motter, A.E., Lai, Y.-C. (2002). Cascade-based attacks on complex networks, *Physical Review E*, Vol. 66, No. 6, Article No. 065102, doi: [10.1103/PhysRevE.66.065102](https://doi.org/10.1103/PhysRevE.66.065102).
- [23] Kim, Y., Chen, Y.-S., Linderman, K. (2015). Supply network disruption and resilience: A network structural perspective, *Journal of Operations Management*, Vol. 33-34, 43-59, doi: [10.1016/j.jom.2014.10.006](https://doi.org/10.1016/j.jom.2014.10.006).
- [24] Kazemian, I., Torabi, S.A., Zobel, C.W., Li, Y., Baghersad, M. (2021). A multi-attribute supply chain network resilience assessment framework based on SNA-inspired indicators, *Operational Research*, Vol. 22, No. 3, 1853-1883, doi: [10.1007/s12351-021-00644-3](https://doi.org/10.1007/s12351-021-00644-3).
- [25] Li, Y., Zobel, C.W. (2020). Exploring supply chain network resilience in the presence of the ripple effect, *International Journal of Production Economics*, Vol. 228, Article No. 107693, doi: [10.1016/j.ijpe.2020.107693](https://doi.org/10.1016/j.ijpe.2020.107693).
- [26] Cardoso, S.R., Barbosa-Póvoa, A.P., Relvas, S., Novais, A.Q. (2015). Resilience metrics in the assessment of complex supply-chains performance operating under demand uncertainty, *Omega*, Vol. 56, 53-73, doi: [10.1016/j.omega.2015.03.008](https://doi.org/10.1016/j.omega.2015.03.008).
- [27] Zhao, K., Kumar, A., Harrison, T.P., Yen, J. (2011). Analyzing the resilience of complex supply network topologies against

- random and targeted disruptions, *IEEE Systems Journal*, Vol. 5, No. 1, 28-39, doi: [10.1109/JSYST.2010.2100192](https://doi.org/10.1109/JSYST.2010.2100192).
- [28] Zhao, K., Scheibe, K., Blackhurst, J., Kumar, A. (2019). Supply chain network robustness against disruptions: topological analysis, measurement, and optimization, *IEEE Transactions on Engineering Management*, Vol. 66, No. 1, 127-139, doi: [10.1109/TEM.2018.2808331](https://doi.org/10.1109/TEM.2018.2808331).
- [29] Brintrup, A., Ledwoch, A. (2018). Supply network science: Emergence of a new perspective on a classical field, *Chaos*, Vol. 28, No. 3, Article No. 033120, doi: [10.1063/1.5010766](https://doi.org/10.1063/1.5010766).
- [30] Hasani, A., Khosrojerdi, A. (2016). Robust global supply chain network design under disruption and uncertainty considering resilience strategies: A parallel memetic algorithm for a real-life case study, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 87, 20-52, doi: [10.1016/j.tre.2015.12.009](https://doi.org/10.1016/j.tre.2015.12.009).
- [31] Hasani, A., Zegordi, S.H., Nikbakhsh, E. (2015). Robust closed-loop global supply chain network design under uncertainty: The case of the medical device industry, *International Journal of Production Research*, Vol. 53, No. 5, 1596-1624, doi: [10.1080/00207543.2014.965349](https://doi.org/10.1080/00207543.2014.965349).
- [32] Yang, Q., Scoglio, C.M., Gruenbacher, D.M. (2021). Robustness of supply chain networks against underload cascading failures, *Physica A: Statistical Mechanics and its Applications*, Vol. 563, Article No. 125466, doi: [10.1016/j.physa.2020.125466](https://doi.org/10.1016/j.physa.2020.125466).

Human-robot collaboration assembly line balancing considering cross-station tasks and the carbon emissions

Li, Y.C.^{a,*}, Wang, X.^b

^aSchool of Economics and Management, Beijing University of Technology, P.R. China

^bSchool of Economics and Management, Beijing University of Technology, P.R. China

ABSTRACT

With the growth of industrialization, the global manufacturing industry is continually evolving and reforming in the direction of intelligence and green production. Industrial robots have replaced human workers because of the benefit of production efficiency. However, the large-scale application of robots requires a large amount of energy consumption and generates a large amount of CO₂, which will lead to energy waste and environmental pollution. In addition, in term of performing some particular tasks, current robot technology cannot achieve the same level of intelligence as human. Therefore, the design trend of assembly lines in industry has shifted from traditional configuration to human-robot collaboration to achieve higher productivity and flexibility. This paper investigates the human-robot collaboration (HRC) assembly line balancing problem, taking cycle time and carbon emission as primary and secondary objectives. A new mixed-integer programming model that features a cross-station design is formulated. A particle swarm algorithm (PSO) with two improvement rules is designed to solve the problems. The comparative experiments on ten benchmark datasets are conducted to assess the performance of the proposed algorithm. The experimental results indicate that the improved particle swarm algorithm is superior to the other two heuristics: simulated annealing (SA) and the late acceptance hill-climbing heuristic (LAHC).

ARTICLE INFO

Keywords:

Assembly line balancing problem;
Human-robot collaboration;
Cross-station tasks;
Carbon emissions;
Collaborative robot (cobot);
Particle swarm algorithm (PSO)

*Corresponding author:

15831029622@163.com
(Li, Y.C.)

Article history:

Received 26 February 2024

Revised 17 April 2024

Accepted 19 April 2024



Content from this work may be used under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

Assembly line (AL) plays an essential role in industrial production. The assembly line balancing problem (ALBP) refers to an actual production scheduling problem of assigning the assembly sequence, distributing the production tasks, and dispatching agents to workstations appropriately to meet production targets in the manufacturing process. With the development of intelligent manufacturing technology, the application of industrial robots can not only improve product quality but also improve production efficiency. Robot assembly lines (RAL) have been widely used in various production fields of the manufacturing industry.

In recently years, we find that even though the robot is highly advanced, it can still not carry out some production jobs with high accuracy or flexibility. Furthermore, robots consume a large amount of electricity during operation, leading to energy consumption and the generation of greenhouse gases such as CO₂, exacerbating the trend of global warming. To address these issues, human-robot collaboration (HRC) is evolving into a new mode of production, directing the continuous development regarding the intelligent and environmental friendly manufacturing.

This study considers a design of “cross-station task” which has already been used in the actual production process of manufacturing enterprises. It is possible for a “cross-station task” to be processed simultaneously at multiple stations. By employing this design, it can reduce the idle time of the workstations on the AL, boost the workstations’ production efficiency and reduce the energy consumption and carbon emissions.

In this paper, we build a multi-objective programming model for balancing collaborative assembly lines between humans and robots while considering carbon emissions. The model considers the carbon emissions of various types of robots when optimizing two objectives, production efficiency represented by cycle time and carbon emissions. The contributions are twofold. Firstly, an ALBP-HRC considering a “cross-station task” design, which has never been discussed in studies of human-robot interactions on AL, is well studied. Secondly, this research utilizes and enhances the particle swarm optimization (PSO) algorithm to solve medium-scale and large-scale problems. Two improvement rules, “Task exchange” and “Set expansion,” are designed for the proposed PSO algorithm, which can reduce the cycle time or find more Pareto solutions. The proposed PSO are compared with two other heuristics, simulated annealing (SA) and the late acceptance hill-climbing heuristic (LAHC) on solving a set of benchmark problems. The results show that PSO outperforms the other algorithms in terms of three different metrics.

The remainder of this paper is organized as follows. Section 2 describes the current research progress in the literature. We propose a mixed-integer programming model in Section 3. An algorithm, based on the particle swarm optimization algorithm (PSO) is developed in Section 4 to find the optimal solution. Experimental studies are conducted in Section 5. Section 6 concludes the paper. The notations are defined in Table 1 and used throughout the paper.

Table 1 Notations

N	Total number of the agent types
n	Total number of the tasks
m	Total number of the workstations
a	Index of agent
i, j	Index of task ($\forall i, j = 1, 2, \dots, n$)
s, h	Index of workstation ($\forall s, h = 1, 2, \dots, m$)
t_{ia}	The operation time of task i by agent a
W_{ias}	The workload of agent a at station s
$P_r(i)$	Index of the immediate predecessors of task i
OPC_a	The operation energy consumption of the agent a per unit time
SEC_a	The standby energy consumption of the agent a per unit time
ECF_{elec}	Carbon emissions per unit of electricity consumption
TCF	The total carbon emissions
c	The cycle time
γ	The maximum time that can be shared between two stations
\emptyset	A large positive number
x_{is}	0-1 variables, $x_{is} = 1$ if task i is allocated to workstation s and 0 otherwise
y_{as}	0-1 variables, $y_{as} = 1$ if agent a is allocated to workstation s and 0 otherwise
z_{ias}	0-1 variables, $z_{ias} = 1$ if task i operated by agent a is allocated to workstation s and 0 otherwise
k_{sh}	0-1 variables, $k_{sh} = 1$ if workstation s utilizes the cycle time of workstation h and 0 otherwise
u_{as}	A non-negative value depicts the idle time of the agent a at workstation s
v_{sh}	A non-negative value indicates how much of workstation h 's cycle time that workstation s occupies

2. Literature review

This part reviews relevant research on robotic assembly line balancing problems (RALBPs) in Section 2.1, the ALBP-HRC studies in Section 2.2, and the studies on ALBPs considered the carbon emissions in Section 2.3.

2.1 Review of robotic assembly line balancing problem

The research on RALBP can be traced back to the 1990s. Rubinovitz *et al.* [1] first proposed the concept of the robot assembly line in 1993, and they proposed a linear programming model for the RALBP-I problem. Yoosefelahi *et al.* [2] studied the RALBP-II problem to simultaneously optimize multiple objectives, namely, robot costs, and cycle time. A mixed-integer programming (MILP) model was established. As the problem is NP-hard, the article proposed three metaheuristic methods to solve the above problem. More recently, Nilakantan [3] were the first to solve the RALBP-II problem in a U-shaped assembly line layout, they proposed a 0-1 IP model for small-scale problems. Wang *et al.* [4] improved traditional particle swarm optimization algorithms by introducing several extension operators, enabling them to solve process planning (PP) problems such as robot assembly lines. Borba *et al.* [5] proposed two new algorithms, namely a “branch-bound and remember” algorithm and an “iterative beam search” algorithm with problem-specific dominance rules to minimize the cycle time in RALBP. Li *et al.* [6] proposed a mixed integer linear programming model to minimize the cycle time. Raatz *et al.* [7] proposed a RALBP that minimized the cost. A multi-objective optimization method, namely the genetic algorithm, was used to solve the proposed problem. Jiang *et al.* [8] proposed an improved genetic algorithm to optimize the balance problem in the clothing production line in response to the low balance rate and the uneven work intensity of employees. The effectiveness of the algorithm was verified through simulation experiments. Şahin *et al.* [9] studied a robot stochastic assembly line balancing problem (RSALBP) based on the assumption that the task time is fixed. They proposed a mixed-integer second-order cone programming model and a constraint programming model to solve the problem given the number of workstations and robots.

Similar to RALBP, the disassembly line balancing of RAL is also NP-hard. Intelligent optimization algorithms have performed well in solving this type of problem. Based on analyzing the disassembly information of automotive components, Yu *et al.* [10] established a disassembly model for automotive components. The optimal disassembly sequence was obtained by considering the mapping between the Floyd Warhill algorithm and car disassembly patterns. Wang *et al.* (2021) [11] studied the multi-objective disassembly line balance problem and proposed an improved genetic algorithm to solve the model.

2.2 Review of human-robot collaboration assembly line balancing problem

The flexibility of HRC-AL is much higher than that of traditional robotic or manual assembly lines, which can improve the efficiency of ALs, enhance the work enthusiasm of workers, and become the primary production mode of ALs chosen by many enterprises today. Ding *et al.* [12] adopted a human and robot collaborative hybrid assembly cell to develop an automatic subtask allocation strategy. Nikolakis *et al.* [13] adopted a two-level breakdown where tasks were translated to specific operations, being carried out by humans or robots to realize an adaptive and more efficient execution of the production schedule. Mura *et al.* [14], who developed a genetic algorithm to reduce the cost of the assembly line, the number of qualified workers needed, and the variation in worker energy loads. Zanchettin *et al.* [15] used a fuzzy-timed Petri net with uncertain task time to minimize idle time. Vieira *et al.* [16] developed a novel optimization simulation based on the Recursive Optimization-Simulation Approach (ROSA) methodology to prioritize reducing costs and the makespan. Aljinovic *et al.* [17] utilized a systematic framework with the mathematical model to minimize cycle time. Riedel *et al.* [18] provided the architecture and implementation details of an assembly assistance system based on object detection models using deep learning and machine learning algorithms. Nourmohammadi *et al.* (2022) [19] studied ALBP with HRC, where human workers and robots share the same workplace to process tasks simultaneously. They developed a new mixed-integer linear programming model and proposed a neighborhood-search simulated annealing algorithm to solve the problem.

2.3 Review of assembly line balancing problem with carbon emissions

Global climate change has triggered a series of ecological, social, and economic problems, gradually increasing people’s environmental awareness. Enterprises have also begun to explore low-

carbon and low-energy production methods, taking the path of sustainable development. Many existing studies on ALs also incorporate the goal of reducing carbon emissions and energy consumption. Lin *et al.* [20] developed an integrated model for parameter optimization and process workshop scheduling to minimize the number of workstations and carbon emissions during the production process. Li *et al.* [21] presented a restarted simulated annealing algorithm to minimize energy consumption and cycle time simultaneously. Nilakantan *et al.* [22] proposed a multi-objective coevolutionary algorithm to minimize the carbon emissions of RAL. It is the first time carbon emissions have been considered in a robot assembly line system. Zhang *et al.* [23] established a multi-objective mathematical model with energy consumption and line balance rate. They proposed a multi-objective algorithm combining cellular strategy and local search to solve this multi-objective problem. Sun *et al.* [24] proposed an energy-efficient robot assembly line balancing (EERALB) problem to minimize cycle time and total energy consumption. They proposed a multi-objective mathematical model and a boundary-oriented mixed distribution estimation algorithm to solve this problem. Zhou *et al.* [25] considered robots with different efficiency and energy consumption rates in their programming model, with the total energy consumption and workload as optimization objectives. They proposed an improved MOEA/D algorithm and evaluated its superiority through computational experiments.

With the continuous development of industrialization, the relevant literature on human-robot collaborative assembly lines is constantly increasing, but there are still some research gaps worth exploring and investigating. For example, current research emphasizes traditional production objectives such as the number of stations, cycle time and cost but neglects carbon emissions. The related research which seeks the balance between carbon emissions and production efficiency in collaborative assembly lines is still absent. To this end, we propose a mixed-integer programming model for optimizing the cycle time and total carbon emissions for HRC-AL with “cross-station task” design.

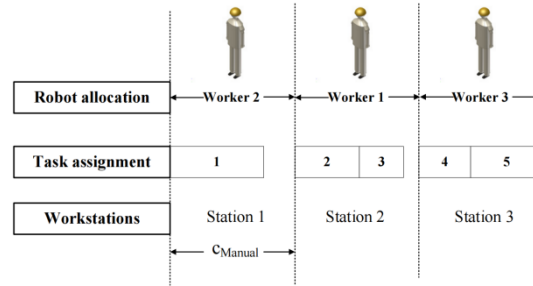
3. Mathematical model

In this section, we explain the structure of the research problem and establish a mixed-integer programming model to formulate the above problem.

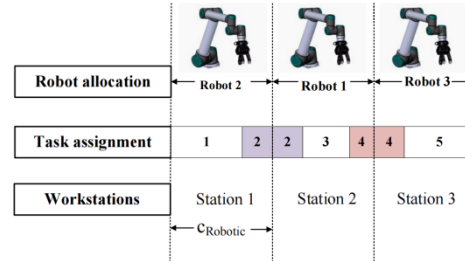
3.1 Problem description

In this paper, we consider a homogeneous product. Assume that there are N types of agents where the former $(N - 1)_{th}$ agents refer to robots and the N_{th} agent refers to human worker. The task operation time varies depending on the agent type. In the “cross-station task” design, one task might be operated concurrently at the assigned station and its rear station or the front station (if it exists). Because the industrial robotic arms contain multiple joints that act as axes controlling movement, the application of industrial robotic arms allows a task can be processed at a pair of stations simultaneously. On the contrary, the human workers cannot process more than one task. We provide some examples to illustrate our statements. In Fig. 1(a), tasks cannot be shared between stations resulting in some idle time which may lead to low production efficiency. In Fig. 1(b), task 2 and task 4 can be processed by the robots in advance at the previous station. In Fig. 1(c), cycle time cannot be shared between station 1 and 2 due to the assignment of human worker, but task 4 can be processed in advance by the robot at station 3. Given the number of stations and the task processing sequence, we can observe that the “cross-station task” design can reduce cycle time by comparing c_{Manual} , $c_{Robotic}$ and c_{HRC} .

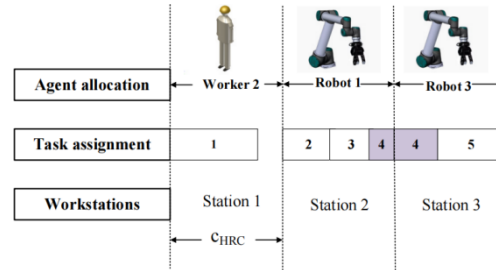
The objective of the optimization problem is to minimize the cycle time and the total carbon emissions. To address the aforementioned problem, we propose a mixed-integer programming model called “ALBP-HRC-CS,” aiming at minimizing the two objectives in the Pareto optimization framework.



(a) The assembly line with only human workers



(b) The assembly line with only robots and cross-station task design



(c) The assembly line with the human-robot collaboration and cross-station task design

Fig. 1 The layout of the assembly line

3.2 Mathematical formulation

The formulation of the mathematical model is as follows.

$$\min c \quad (1)$$

$$\min TCF \quad (2)$$

$$EC = OEC + SEC \quad (3)$$

$$OEC = \sum_{s=1}^m OEC_s \quad (4)$$

$$SEC = \sum_{s=1}^m SEC_s \quad (5)$$

$$OEC_s = \sum_{a=1}^N \sum_{i=1}^n OPC_a \times z_{ias} \times t_{ia}, \forall s = 1, \dots, m \quad (6)$$

$$OEC_s = \sum_{a=1}^N u_{as} \times SPC_a, \forall s = 1, \dots, m \quad (7)$$

The objective function (Eq. 1) minimizes the cycle time of all product models. The objective function (Eq. 2) minimizes the total carbon emissions. Robots in operation or on standby consume electricity and produce carbon emissions. Therefore, the total energy consumption is equal to the sum of all robotic workstations' standby and operation energy consumption. The total

carbon emissions are equal to the product of total energy consumption and carbon emission coefficient. Eqs. 3 to 7 compute the overall energy consumption for all workstations and the individual energy consumption of each workstation. It should be noted that the agent's OPC_a and SPC_a values are 0 if it is a human worker.

Now we first introduce the basic constraints of the proposed model.

$$\sum_{s=1}^m x_{is} = 1, \forall s = 1, \dots, m, \forall i = 1, \dots, n \quad (8)$$

$$\sum_{a=1}^N y_{as} = 1, \forall s = 1, \dots, m, \forall a = 1, \dots, N \quad (9)$$

$$x_{is} + y_{as} \leq z_{ias} + 1, \forall s = 1, \dots, m, \forall i = 1, \dots, n, \forall a = 1, \dots, N \quad (10)$$

$$x_{is} + (1 - y_{as}) \leq (1 - z_{ias}) + 1, \forall s = 1, \dots, m, \forall i = 1, \dots, n, \forall a = 1, \dots, N \quad (11)$$

$$(1 - x_{is}) + (1 - y_{as}) \leq (1 - z_{ias}) + 1, \forall s = 1, \dots, m, \forall i = 1, \dots, n, \forall a = 1, \dots, N \quad (12)$$

$$(1 - x_{is}) + y_{as} \leq (1 - z_{ias}) + 1, \forall s = 1, \dots, m, \forall i = 1, \dots, n, \forall a = 1, \dots, N \quad (13)$$

$$\sum_{s=1}^m s \times x_{js} \leq \sum_{s=1}^m s \times x_{is}, \forall j \in P_r(i), \forall s = 1, \dots, m, \forall i = 1, \dots, n \quad (14)$$

Constraints (Eqs. 8 and 9) are indivisibility of tasks/agents, meaning that a task/agent can only be allocated to one workstation. Constraints (Eqs. 10 to 13) ensure that the task can be executed, that is, when a task and an agent are allocated to the same workstation, the agent must process the task. Constraint (Eq. 14) is the task precedence constraints.

Next, the constraints related to the cross-station design are illustrated.

$$u_{as} \leq c + v_{s,s+1} + v_{s,s-1} - v_{s+1,s} - v_{s-1,s} - \sum_{a=1}^N \sum_{i=1}^n z_{ias} \times t_{ia} + \emptyset(1 - y_{as}), \quad (15)$$

$$\forall s = 1, \dots, m, \forall i = 1, \dots, n, \forall a = 1, \dots, N$$

$$u_{as} \geq c + v_{s,s+1} + v_{s,s-1} - v_{s+1,s} - v_{s-1,s} - \sum_{a=1}^N \sum_{i=1}^n z_{ias} \times t_{ia} - \emptyset(1 - y_{as}), \quad (16)$$

$$\forall s = 1, \dots, m, \forall i = 1, \dots, n, \forall a = 1, \dots, N$$

$$v_{s,s+1} \leq \gamma, \forall s = 1, \dots, m \quad (17)$$

$$v_{s+1,s} \leq \gamma, \forall s = 1, \dots, m \quad (18)$$

$$k_{s,s+1} \leq \sum_{a=1}^{N-1} y_{as}, \forall s = 1, \dots, m - 1 \quad (19)$$

$$k_{s,s-1} \leq \sum_{a=1}^{N-1} y_{as}, \forall s = 1, \dots, m \quad (20)$$

$$k_{s,s+1} \leq 1 - y_{Ns}, \forall s = 1, \dots, m - 1 \quad (21)$$

$$k_{s,s-1} \leq 1 - y_{Ns}, \forall s = 1, \dots, m \quad (22)$$

$$k_{s,s+1} + k_{s+1,s} \leq 1, \forall s = 1, \dots, m \quad (23)$$

$$v_{s,s+1} \leq \emptyset \cdot k_{s,s+1}, \forall s = 1, \dots, m \quad (24)$$

$$v_{s+1,s} \leq \emptyset \cdot k_{s+1,s}, \forall s = 1, \dots, m \quad (25)$$

$$v_{s,s+1} \geq 0.1 \cdot k_{s,s+1}, \forall s = 1, \dots, m \quad (26)$$

$$v_{s+1,s} \geq 0.1 \cdot k_{s+1,s}, \forall s = 1, \dots, m \quad (27)$$

Constraints (Eqs. 15 and 16) are in effect when agent a is allocated to workstation s ($y_{as} = 1$), and they calculate the agent's idle time if that agent is allocated. Constraints (Eqs. 17 to 18) restrict the maximum amount of time that the robot at one station consumes the cycle time of another station. Constraints (Eqs. 19 and 20) ensure that the ability to utilize the cycle time of the front or rear station is only possible if the station is furnished with robots. Constraints (Eqs. 21 to 22) ensure that human workers cannot utilize the cycle time of the front or rear station. Constraints (Eqs. 23 to 27) denote that tasks shared between adjacent stations can only occur once.

$$v_{s,s+1}, v_{s+1,s} \geq 0, \forall s = 1, \dots, m \quad (28)$$

$$u_{as} \geq 0, \forall s = 1, \dots, m, \forall a = 1, \dots, N \quad (29)$$

$$v_{0,1}, v_{1,0}, v_{m,m+1}, v_{m+1,m} = 0 \quad (30)$$

Constraints (Eqs. 28 to 30) are the domain constraints.

4. Methods

In this section, a particle swarm optimization algorithm (PSO) with two improvement rules is adapted to address the "ALBP-HRC-CS". The general framework of PSO is introduced in section 4.1. Variable neighborhood search mechanism is developed in section 4.2. Encoding and decoding schemes are designed in section 4.3. Finally, two improvement rules are delineated in section 4.4.

4.1 The general framework of PSO

Kennedy *et al.* [26] developed the simplified version of PSO in 1995 by the insight of birds' feeding behavior. Birds can locate most food locations through collective information sharing. In the mathematical model of "ALBP-HRC-CS", the PSO algorithm minimizes the whole assembly line's cycle time and carbon emissions by changing agent sequences and task sequences. The relevant parameters are displayed in Table 2.

Table 2 Related parameters

$iter, iter'$	The iteration index
$iter_{max}$	The maximal number of iterations
k_1, k_2	The index for agent particles and task particles
X_{k_1}, X_{k_2}	The position for agent particles and task particles
V_{k_1}, V_{k_2}	The velocity for agent particles and task particles
P_{best}, G_{best}	The local and global best of the particles
l_1, l_2	Learning coefficients
U_1, U_2	Uniform random numbers between [0,1]

In PSO, each particle in the search domain is constantly moving. During the movement, the particle's speed is changing, resulting in the change of the final position of the particle. The position update for the agent and task follow Eq. 31 and Eq. 32, respectively.

$$X_{k_1}^{iter+1} = X_{k_1}^{iter} + V_{k_1}^{iter+1} \quad (31)$$

$$X_{k_2}^{iter+1} = X_{k_2}^{iter} + V_{k_2}^{iter+1} \quad (32)$$

The velocity updates for the agent and task follow Eqs. 33 and 34, respectively.

$$V_{k_1}^{iter+1} = V_{k_1}^{iter} + l_1 \times U_1 \times (P_{best_{iter}} - X_{k_1}^{iter}) + l_2 \times U_2 \times (G_{best_{iter}} - X_{k_1}^{iter}) \quad (33)$$

$$V_{k_2}^{iter+1} = V_{k_2}^{iter} + l_1 \times U_1 \times (P_{best_{iter}} - X_{k_2}^{iter}) + l_2 \times U_2 \times (G_{best_{iter}} - X_{k_2}^{iter}) \quad (34)$$

PSO is an intelligent optimization algorithm with high local convergence and a "premature" phenomenon. If the PSO falls into a local extreme value early in the iteration, it will be difficult to jump out of it later, dramatically reducing the efficiency of the algorithm and significantly reduc-

ing the probability that the algorithm will find a globally optimal solution. To address this problem, “Task exchange” and “Set expansion” rules are developed to enhance the quality of the Pareto set’s solutions. The PSO also features a variable neighborhood search (VNS) mechanism.

The main body of the improved PSO is shown in Fig. 2.

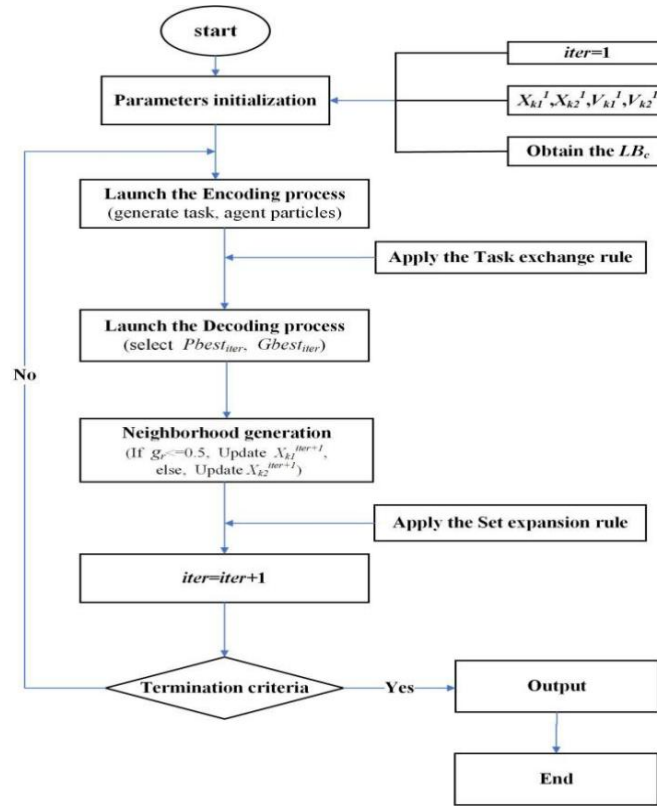


Fig. 2 Main body of the improved PSO

4.2 Encoding and decoding

Encoding

In this section, the task sequence Sq_{task} and agent sequence Sq_{agent} according to the constraints are arrayed. First, we obtain the initial Sq_{task} and Sq_{agent} as follows.

Initial Sq_{task}

Step 1: Set $Sq_{task} = []$

Step 2: Set $X_{k_1}^1 = 1 + random[0,1] * n$, where $[0,1]$ is a randomly generated number that conforms to the uniform distribution $[0,1]$

Step 3: The task is arrayed in a non-increasing order according to the $X_{k_1}^1$ under the constraint of task priority, and the Sq_{task} is attained

Step 4: If a task violates the task priority relationship when it is assigned, skip the task, assign the next task, and return to Step 3; otherwise, go to Step 5

Step 5: Continue Steps 3 and 4 until all tasks have been assigned

Initial Sq_{agent}

Step 1: Set $Sq_{agent} = []$

Step 2: Set $X_{k_2}^1 = 1 + random[0,1] * N$, where $[0,1]$ is a randomly generated number that conforms to the uniform distribution $[0,1]$

Step 3: Assign the agent with the largest values to the workstation

Step 4: Repeat Step 3 until all the stations are equipped with an agent

Decoding

As we know, lower bounds for the objective variables is crucial for the decoding process. We develop the lower bounds for c and TCF (LB_c and LB_{TCF}). It is straightforward that the lower bound for TCF : $LB_{TCF} = 0$ (i.e., the manual assembly line). Next, according to the necessary conditions for the feasibility of the model, LB_c can be determined as: $LB_c = [\max(\sum_{i=1}^n \min_{a=1}^N t_{ia}/m, \max_{i=1}^n \min_{a=1}^N t_{ia})]$. The decoding scheme is listed below.

- Step 1: Load $t_{ia}, OPC_a, SPC_a, ECF_{elc}, LB_c, P(i), m, \gamma, Sq_{task}, Sq_{agent}$, set $s = 1, ll = 1$
- Step 2: Arrange agents for each station according to Sq_{agent}
- Step 3: Arrange the ll_{th} task in the Sq_{task} to station s and compute the workload of station s , Wl_{as} , go to the Step 4; otherwise, if there is no task to be assigned, go to Step 13
- Step 4: If $Wl_{as} \geq LB_c + v_{s,s-1} - v_{s-1,s}$ AND $s < m$, go to Step 5; if $Wl_{as} \geq LB_c + v_{s,s-1} - v_{s-1,s}$ AND $s == m$, go to Step 12; otherwise, go back to Step 3, and $ll = ll + 1$
- Step 5: If $OPC_a(a_s(s)) == 0$, the agent assigned to the station s is a human worker who cannot utilize the cycle time of the front or rear station, go to Step 9; otherwise, go to step 6
- Step 6: If $Wl_{as} + t_{ia} \leq LB_c + v_{s,s-1} - v_{s-1,s} + \gamma$ AND $t_{ia} \leq LB_c + v_{s,s-1} - v_{s-1,s} - \min(\gamma, LB_c + v_{s,s-1} - v_{s-1,s} - Wl_{as}) - Wl_{as} + t_{ia}$, go to Step 7; otherwise, go to Step 9
- Step 7: Task $i = Sq_{task}(ll)$ is assigned to the station s , update Wl_{as} and $ll = ll + 1$, and then, repeat this step; otherwise, go to Step 8
- Step 8: Set $v_{s,s+1} = Wl_{as} + t_{ia} - (LB_c + v_{s,s-1} - v_{s-1,s}), v_{s+1,s} = 0$, and go to Step 11
- Step 9: Task $i = Sq_{task}(ll)$ is assigned to the station $s + 1$, update Wl_{as} and $ll = ll + 1$, go to Step 10
- Step 10: Set $v_{s+1,s} = \min(\gamma, LB_c + v_{s,s-1} - v_{s-1,s} - \sum_{j=1}^n z_{jas} t_{ja}), v_{s,s+1} = 0$, go to Step 11
- Step 11: OEC_s and SEC_s are computed according to Eqs. 5 to 10, $s = s + 1$, and then, return to Step 3
- Step 12: The total task time is computed for all unassigned tasks TR , and set $LB_c = LB_c + [TR/m], s = 1, ll = 1$, return to the Step 3
- Step 13: Compute TCF and $c = LB_c$

In the decoding process, the initial parameters are loaded in Step 1. The agent is assigned to each station according to Sq_{agent} in Step 2. Step 3 arranges tasks and computes the workload of the station. In Step 4, three situations may occur: if the current station is overloaded and is not the last station, go to Step 5; if the current station is overloaded but is the last station, go to Step 12; otherwise, return to Step 3. Step 5 shows that the station equipped with a human worker cannot utilize the time from the adjacent stations. Step 6 schedules this task to the current station through the “cross-station task” design and goes to Step 7; if it does not, go to Step 9. Step 7 updates the workload of the current station. Step 8 calculates the time it takes from its rear station. Step 9 assigns this task to the next station and updates the workload of the current station; Step 10 calculates the idle time of the current station (which can be used by other stations). The total energy consumption for the workstation is computed in Step 11. In Step 12, LB_c is increased to maintain feasibility. The above steps are designed to obtain a feasible task sequence. Based on it, we can calculate the values of the objective functions (TCF and c) in step 13.

4.3 Variable Neighborhood Search

Variable Neighborhood Search (VNS) is an enhanced local search mechanism. For alternating searches, it utilizes the neighborhood structure of various actions, striking a suitable balance between concentration and dispersion. First, the initial solution is generated and its neighborhoods is defined. Given the initial solution, VNS systematically selects new solution between two neighborhoods related to task and agent, respectively. A random number g_r is generated to select the part of the particle subject to change during each iteration. For instance, if $g_r \leq 0.5$, the positions and velocities are updated for the agent sequences while the task sequences remain unchanged. The new solution becomes the best if it can lead to a better objective. The algorithm terminates when the termination criterion (e.g., runtime limit) is satisfied.

4.4 Improvement rules

This section considers “Task exchange” and “Set expansion” as two enhanced mechanisms to find more solutions. “Task exchange” can enhance the efficiency of the PSO algorithm by swapping the processing orders of two different tasks operated by two stations within a feasible task sequence. “Set expansion” is proposed to address problem where the PSO algorithm is vulnerable to “premature convergence” and “local optimal solutions”. “Set expansion” can be used to discover more Pareto solutions.

Task exchange

The “task exchange” mechanism can be executed as follows: after obtaining a feasible solution, exchange the processing order and stations of two different tasks; assume that task a and task b are processed by agent r and r^* and assigned to station s and s^* , respectively (agents here can be human workers or robots); the tasks can be exchanged if the following criteria are met.

1. The exchange between task a and task b does not violate the precedence relationship;
2. “ $t_{ar} + t_{br^*} > t_{ar^*} + t_{br}$ ” or “ $t_{ar} + t_{br^*} == t_{ar^*} + t_{br}$ and $TCF_{ar} + TCF_{br^*} > TCF_{ar^*} + TCF_{br}$ ”;

The above criteria ensure that following task exchange, the cycle time can be decreased or the overall carbon emissions can be reduced without increasing the cycle time. When task exchange is applied, the decoding process is executed again to find a new cycle time and a new TCF .

Set expansion

Among the non-dominated solutions in the Pareto set, some solutions are far away from each other, resulting in a blank area in the middle of two solutions. The appearance of the blank area indicates some solutions in Pareto set are not available. To address that issue, the search area shall be expanded. The rule is explained as follows.

- Step 1: Input the initial Pareto set
- Step 2: Calculate the cycle time difference between two adjacent feasible solutions respectively, and calculate the mean value of these differences to obtain c_{diff}
- Step 3: Select the two solutions with the largest differentials of cycle times, called c_1 and c_2 (assume $c_1 > c_2$)
- Step 4: If $c_1 - c_2 > \lambda * c_{diff}$, go to Step 5, otherwise, stop the PSO
- Step 5: Set $c_3 = (c_1 + c_2)/2$, and calculate the TCF_3
- Step 6: Update the Pareto solution set and return to Step 2

In the “Set extension”, the following points should be noted: (1) The “Set extension” is added to expand the search neighborhood and explore better feasible solutions; for example, if $TCF_2 > TCF_3 > TCF_1$, it means that we have found a new feasible solution, and the Pareto set will be updated; (2) λ is a constant value which determines the termination criterion. The smaller the λ is chosen, the newer feasible solutions can be obtained, but it also takes more computational time.

5. Experimental design

5.1 Experimental setting

The benchmark datasets are extracted from Otto *et al.* [27]. According to the number of tasks, the experiments are divided into medium-scale ($n = 100$) and large-scale ($n = 1000$) problems. The values of OPC_a and SPC_a of different types of robots can be referred to Nilakantan *et al.* [22]. The number of agents (i.e., robot or human worker) ranges from 6 to 50. This experiment includes 10 instances, each is provided with 5 different values of m . Therefore, there is a total of 50 independent tests, and the corresponding results are presented.

5.2 Parameter setting

To show the superiority of the proposed algorithm, the PSO is compared with the other two algorithms, namely the simulated annealing algorithm (SA), which is extended on the SA proposed by Li *et al.* [21] and the late acceptance hill climbing heuristic algorithm (LAHC) which is extended on the LAHC developed by Yuan *et al.* [28]. All algorithms are coded in MATLAB R2022a and executed on a computer with inter Core i5, 4.7 GHz. Since the meta-heuristic algorithms are stochastic, each algorithm is run ten times for each instance to find the average result, and the algorithm runtime is limited to rt ($rt = n \times n$) milliseconds. The algorithm's parameters are obtained from the related literature. Table 3 shows the parameters and their values used in the three algorithms.

To evaluate the effectiveness of the three algorithms, three appropriate metrics, namely the ratio of non-dominated solutions (R_p), the non-dominated solution's convergence (C_p), and the spread measure (S_p) have been applied. R_p measures the solution in the Pareto set that is not dominated by other solutions; the higher the value of R_p is, the more effective the algorithm will be. C_p describes the difference between one Pareto-optimal set and the true Pareto-optimal set; the smaller the value of C_p is, the better the algorithm's performance. S_p captures the distributions of the solutions of the Pareto-optimal set; the algorithm with lower S_p performs better. Li *et al.* [29] introduced the formulae for these three metrics.

Table 3 Parameter settings for PSO, LAHC, and SA algorithms

Parameters	PSO	LAHC	SA
The number of task particles in medium (large) instances	100(30)	---	---
The number of agent particles in medium (large) instances	50(30)	---	---
The learning coefficient $l_1(l_2)$	2(2)	---	---
The expansion parameter λ	1.5	---	---
The initial temperature	---	---	100
The cooling rate	---	---	0.9
The length of the cost list	---	100	---

5.3 Results and analysis

The results for the three assessment indicators in medium and large-scale situations are shown in Tables 4 and 5, and the best results are denoted in bold.

The PSO outperforms the other two algorithms in Table 4 for medium-scale instances. For 24 out of 25 instances, PSO finds the best Pareto solutions based on the R_p metric, showing that the PSO algorithm generates more non-dominated solutions in the Pareto solution set than the other two algorithms. The PSO algorithm finds all of the best Pareto solutions based on the C_p metric, demonstrating that the optimal set of the PSO algorithm converges more quickly than those of other algorithms; for 22 out of 25 instances, PSO also generates the best Pareto solutions based on the S_p metric, demonstrating that the spread of Pareto set is superior.

For large-scale problem, as shown in Table 5, PSO provides the best Pareto solutions for 19 out of 25 problems using the R_p metric. The PSO algorithm finds all of the best Pareto solutions based on the C_p metric. For the S_p metric, PSO finds the best Pareto solutions for 23 out of 25 instances.

We also display the performance of the three algorithms on representative examples by using scatter plots. The results are shown in Fig. 3. The first number denotes the problem scale (M for medium problems; L for large problems); the second number denotes the index of instance in Table 4 and Table 5; and the third number denotes the number of stations.

Table 4 Computational results on the Medium-scale problems

Problem	m	SA			LAHC			PSO		
		R_p	C_p	S_p	R_p	C_p	S_p	R_p	C_p	S_p
1	6	0.70	0.01	1.09	0.00	0.02	0.78	0.79	0.00	0.77
	8	0.48	0.02	0.94	0.00	0.02	0.60	0.84	0.00	0.68
	10	0.58	0.01	0.98	0.00	0.01	0.75	0.79	0.00	0.59
	12	0.50	0.01	0.85	0.09	0.02	0.70	0.75	0.00	0.57
	14	0.53	0.01	1.02	0.00	0.01	0.81	0.86	0.00	0.73
2	6	0.30	0.01	0.96	0.17	0.01	0.64	0.83	0.00	0.78
	8	0.31	0.00	1.03	0.21	0.01	0.92	0.88	0.00	0.64
	10	0.64	0.01	0.98	0.00	0.02	0.81	0.86	0.00	0.60
	12	0.57	0.01	0.98	0.06	0.01	0.84	0.85	0.00	0.55
	14	0.37	0.02	0.76	0.14	0.01	0.69	0.77	0.00	0.52
3	6	0.86	0.01	1.01	0.00	0.02	0.97	0.85	0.00	0.76
	8	0.23	0.02	0.89	0.03	0.01	1.02	0.92	0.00	0.68
	10	0.29	0.01	0.78	0.00	0.01	0.92	0.87	0.00	0.59
	12	0.40	0.01	0.85	0.00	0.01	0.70	0.84	0.00	0.58
	14	0.25	0.02	0.65	0.10	0.01	0.58	0.79	0.00	0.48
4	6	0.22	0.02	0.76	0.37	0.01	0.73	0.81	0.00	0.72
	8	0.39	0.01	0.90	0.04	0.01	0.69	0.82	0.00	0.67
	10	0.35	0.02	0.84	0.00	0.01	0.93	0.84	0.00	0.56
	12	0.45	0.01	0.78	0.04	0.01	0.94	0.78	0.00	0.74
	14	0.08	0.01	0.77	0.42	0.00	0.83	0.79	0.00	0.54
5	6	0.50	0.01	0.94	0.00	0.01	0.83	0.84	0.00	0.77
	8	0.52	0.01	1.07	0.00	0.01	0.94	0.83	0.00	0.67
	10	0.45	0.01	0.85	0.00	0.01	0.77	0.86	0.00	0.62
	12	0.45	0.01	0.81	0.03	0.02	0.69	0.73	0.00	0.76
	14	0.26	0.02	0.87	0.00	0.02	0.73	0.77	0.00	0.52

Table 5 Computational results on the Large-scale problems

Problem	m	SA			LAHC			PSO		
		R_p	C_p	S_p	R_p	C_p	S_p	R_p	C_p	S_p
1	42	0.66	0.01	0.92	0.33	0.01	0.79	0.57	0.00	0.72
	44	0.69	0.00	0.80	0.85	0.00	0.90	0.48	0.00	0.75
	46	0.33	0.01	0.75	0.00	0.02	0.83	0.85	0.00	0.69
	48	0.32	0.01	0.69	0.47	0.01	0.76	0.64	0.00	0.67
	50	0.31	0.01	0.65	0.65	0.00	0.75	0.29	0.00	0.64
2	42	0.57	0.01	0.81	0.24	0.01	0.83	0.63	0.00	0.68
	44	0.00	0.02	0.80	0.18	0.02	0.83	0.99	0.00	0.61
	46	0.38	0.01	0.77	0.41	0.01	0.79	0.59	0.00	0.52
	48	0.45	0.01	0.81	0.31	0.02	0.76	0.77	0.00	0.72
	50	0.36	0.01	0.80	0.58	0.01	0.90	0.64	0.01	0.73
3	42	0.46	0.01	0.71	0.11	0.01	0.80	0.71	0.00	0.66
	44	0.37	0.01	0.85	0.16	0.01	0.70	0.67	0.00	0.78
	46	0.53	0.01	0.77	0.27	0.01	0.82	0.59	0.00	0.67
	48	0.48	0.01	0.68	0.31	0.01	0.91	0.51	0.00	0.68
	50	0.41	0.01	0.72	0.91	0.00	0.74	0.23	0.00	0.68
4	42	0.35	0.01	0.81	0.00	0.02	0.76	0.71	0.00	0.64
	44	0.24	0.01	0.74	0.39	0.01	0.83	0.64	0.00	0.66
	46	0.31	0.01	0.31	0.00	0.02	0.85	0.85	0.00	0.67
	48	0.69	0.00	0.78	0.41	0.01	0.81	0.34	0.00	0.56
	50	0.46	0.01	0.76	0.71	0.00	0.85	0.53	0.00	0.68
5	42	0.38	0.01	0.80	0.09	0.02	0.84	0.70	0.00	0.63
	44	0.44	0.01	0.82	0.00	0.02	0.81	0.83	0.00	0.50
	46	0.08	0.01	0.83	0.00	0.02	0.92	0.99	0.00	0.56
	48	0.33	0.01	0.71	0.50	0.02	0.88	0.59	0.00	0.67
	50	0.26	0.01	0.78	0.00	0.02	0.95	0.92	0.00	0.64

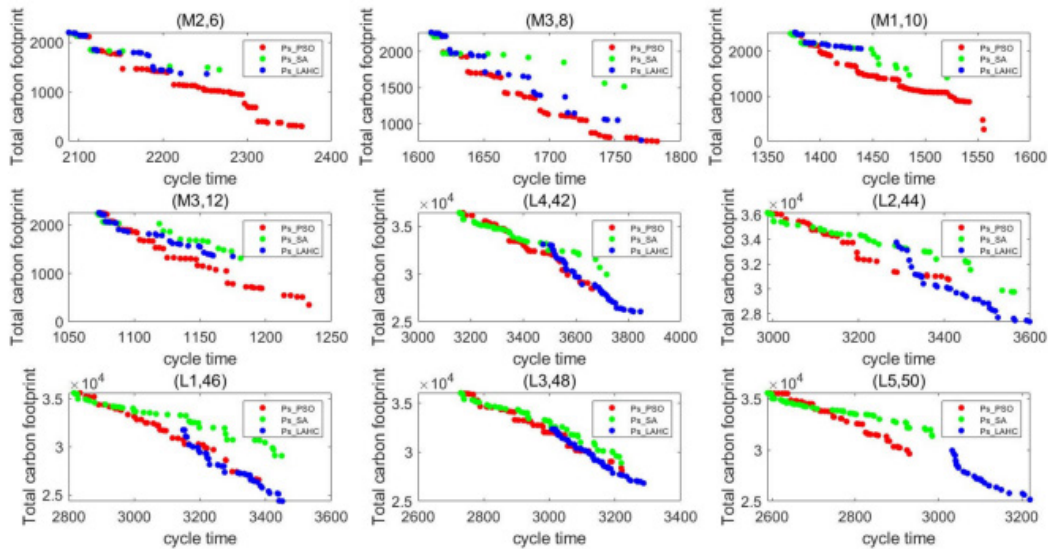


Fig. 3 Some Pareto fronts of different instances

6. Conclusion

In this paper, a multi-objective human-robot collaborative assembly line balancing problem that takes into account production efficiency and carbon emissions is investigated. A mixed-integer programming model is proposed which considers a cross-station task design. This design is intended to improve the assembly line's overall production efficiency and flexibility. A particle swarm optimization algorithm is devised to solve the problem. The performance of the proposed particle swarm optimization algorithm is validated by comparing against simulated annealing and late acceptance hill-climbing algorithms.

The model and algorithm proposed provide some management insights for production in the real world: (1) for assembly line designers, this research can provide a reasonable workstation configurations for HRC production in terms of the goals of production efficiency and carbon emissions; (2) for production managers, the model proposed in this paper can enable them to have a clear understanding of the constitution of carbon emissions in the production process.

In future work, the design for the collaboration of human workers and robots within the same station can be studied. Further, the ALBP with various levels of automation can be examined and contrasted, and management recommendations for the automation transition of small and medium-sized businesses will be offered. Additionally, the research problem can be transformed into a hybrid ALBP-HRC problem by eliminating the assumptions of the single product and the straight AL structure. Finally, the solution methods, such as deep-learning and machine learning algorithms are worth researched.

Acknowledgement

This work was partially supported by the National Natural Science Foundation of China (71901006).

Disclosure and conflicts of interest

No potential conflict of interest was reported by the author(s).

Data availability statements

The data that support the findings of this study are openly available in Wang, Wang xin (2023), "data and result for the "CS-ALBP-HRC", Mendeley Data, V1, doi: 10.17632/xb5mmtvz6j.1.

Ethics statements

There are no human subjects in this article and informed consent and ethics statement is not applicable.

References

- [1] Rubinovitz, J., Bukchin, J., Lenz, E. (1993). RALBP - A heuristic algorithm for design and balancing of robotic assembly lines, *CIRP Annals*, Vol. 42, No. 1, 497-500, doi: [10.1016/S0007-8506\(07\)62494-9](https://doi.org/10.1016/S0007-8506(07)62494-9).
- [2] Yoosefelahi, A., Aminnayeri, M., Mosadegh, H., Davari Ardakani, H. (2012). Type II robotic assembly line balancing problem: An evolution strategies algorithm for a multi-objective model, *Journal of Manufacturing Systems*, Vol. 31, No. 2, 139-151, doi: [10.1016/j.jmsy.2011.10.002](https://doi.org/10.1016/j.jmsy.2011.10.002).
- [3] Mukund Nilakantan, J., Ponnambalam, S.G. (2016). Robotic u-shaped assembly line balancing using particle swarm optimization, *Engineering Optimization*, Vol. 48, No. 2, 231-252, doi: [10.1080/0305215X.2014.998664](https://doi.org/10.1080/0305215X.2014.998664).
- [4] Wang, J.F., Kang, W.L., Zhao, J.L., Chu, K.Y. (2016). A simulation approach to the process planning problem using a modified particle swarm optimization, *Advances in Production Engineering & Management*, Vol. 11, No. 2, 77-92, doi: [10.14743/apem2016.2.211](https://doi.org/10.14743/apem2016.2.211).
- [5] Borba, L., Ritt, M., Miralles, C. (2018). Exact and heuristic methods for solving the robotic assembly line balancing problem, *European Journal of Operational Research*, Vol. 270, No. 1, 146-156, doi: [10.1016/j.ejor.2018.03.011](https://doi.org/10.1016/j.ejor.2018.03.011).
- [6] Li, Z., Janardhanan, M.N., Tang, Q.H., Ponnambalam, S.G. (2019). Model and metaheuristics for robotic two-sided assembly line balancing problems with setup times, *Swarm and Evolutionary Computation*, Vol. 50, Article No. 100567, doi: [10.1016/j.swevo.2019.100567](https://doi.org/10.1016/j.swevo.2019.100567).
- [7] Raatz, A., Blankemeyer, S., Recker, T., Pischke, D., Nyhuis, P. (2020). Task scheduling method for HRC workplaces based on capabilities and execution time assumptions for robots, *CIRP Annals*, Vol. 69, No. 1, 13-16, doi: [10.1016/j.cirp.2020.04.030](https://doi.org/10.1016/j.cirp.2020.04.030).
- [8] Jiang, L., Duan, J.J., Zheng, R.P., Shen, H.N., Li, H., Xu, J. (2023). Optimization and simulation of garment production line balance based on improved GA, *International Journal of Simulation Modelling*, Vol. 22, No. 2, 303-314, doi: [10.2507/IJSIMM22-2-CO6](https://doi.org/10.2507/IJSIMM22-2-CO6).
- [9] Şahin, M.C., Tural, M.K. (2023). Robotic stochastic assembly line balancing, *Flexible Services and Manufacturing Journal*, Vol. 35, No. 4, 1076-1115, doi: [10.1007/s10696-023-09494-x](https://doi.org/10.1007/s10696-023-09494-x).
- [10] Yu, B., Wu, E., Chen, C., Yang, Y., Yao, B.Z., Lin, Q. (2017). A general approach to optimize disassembly sequence planning based on disassembly network: A case study from automotive industry, *Advances in Production Engineering & Management*, Vol. 12, No. 4, 305-320, doi: [10.14743/apem2017.4.260](https://doi.org/10.14743/apem2017.4.260).
- [11] Wang, Y.J., Wang, N.D., Cheng, S.M., Zhang, X.C., Liu, H.Y., Shi, J.L., Ma, Q.Y., Zhou, M.J. (2021). Optimization of disassembly line balancing using an improved multi-objective Genetic Algorithm, *Advances in Production Engineering & Management*, Vol. 16, No. 2, 240-252, doi: [10.14743/apem2021.2.397](https://doi.org/10.14743/apem2021.2.397).
- [12] Ding, H., Schipper, M., Matthias, B. (2014). Optimized task distribution for industrial assembly in mixed human-robot environments - Case study on IO module assembly, In: *Proceedings of 2014 IEEE International Conference on Automation Science and Engineering (CASE)*, New Taipei, Taiwan, 19-24, doi: [10.1109/CoASE.2014.6899298](https://doi.org/10.1109/CoASE.2014.6899298).
- [13] Nikolakis, N., Kousi, N., Michalos, G., Makris, S. (2018). Dynamic scheduling of shared human-robot manufacturing operations, *Procedia CIRP*, Vol. 72, 9-14, doi: [10.1016/j.procir.2018.04.007](https://doi.org/10.1016/j.procir.2018.04.007).
- [14] Dalle Mura, M., Dini, G. (2019). Designing assembly lines with humans and collaborative robots: A genetic approach, *CIRP Annals*, Vol. 68, No. 1, 1-4, doi: [10.1016/j.cirp.2019.04.006](https://doi.org/10.1016/j.cirp.2019.04.006).
- [15] Zanchettin, A.M., Casalino, A., Piroddi, L., Rocco, P. (2019). Prediction of human activity patterns for human-robot collaborative assembly tasks, *IEEE Transactions on Industrial Informatics*, Vol. 15, No. 7, 3934-3942, doi: [10.1109/TII.2018.2882741](https://doi.org/10.1109/TII.2018.2882741).
- [16] Vieira, M., Moniz, S., Gonçalves, B.S., Pinto-Varela, T., Barbosa-Póvoa, A.P., Neto, P. (2022). A two-level optimisation-simulation method for production planning and scheduling: The industrial case of a human-robot collaborative assembly line, *International Journal of Production Research*, Vol. 60, No. 9, 2942-2962, doi: [10.1080/00207543.2021.1906461](https://doi.org/10.1080/00207543.2021.1906461).
- [17] Aljinovic, A., Gjeldum, N., Bilic, B., Mladineo, M. (2022). Optimization of industry 4.0 implementation selection process towards enhancement of a manual assembly line, *Energies*, Vol. 15, No. 1, Article No. 30, doi: [10.3390/en15010030](https://doi.org/10.3390/en15010030).
- [18] Riedel, A., Gerlach, J., Dietsch, M., Herbst, S., Engelmann, F., Brehm, N., Pfeifroth, T. (2021). A deep learning-based worker assistance system for error prevention: Case study in a real-world manual assembly, *Advances in Production Engineering & Management*, Vol. 16, No. 4, 393-404, doi: [10.14743/apem2021.4.408](https://doi.org/10.14743/apem2021.4.408).
- [19] Nourmohammadi, A., Fathi, M., Ng, A.H.C. (2022). Balancing and scheduling assembly lines with human-robot collaboration tasks, *Computers & Operations Research*, Vol. 140, Article No.105674, doi: [10.1016/j.cor.2021.105674](https://doi.org/10.1016/j.cor.2021.105674).
- [20] Lin, W., Yu, D.Y., Zhang, C., Liu, X., Zhang, S., Tian, Y., Liu, S., Xie, Z. (2015). A multi-objective teaching-learning-based optimization algorithm to scheduling in turning processes for minimizing makespan and carbon footprint, *Journal of Cleaner Production*, Vol. 101, 337-347, doi: [10.1016/j.jclepro.2015.03.099](https://doi.org/10.1016/j.jclepro.2015.03.099).
- [21] Li, Z., Tang, Q., Zhang, L.P. (2016). Minimizing energy consumption and cycle time in two-sided robotic assembly line systems using restarted simulated annealing algorithm, *Journal of Cleaner Production*, Vol. 135, 508-522, doi: [10.1016/j.jclepro.2016.06.131](https://doi.org/10.1016/j.jclepro.2016.06.131).
- [22] Nilakantan, J.M., Li, Z., Tang, Q., Nielsen, P. (2017). Multi-objective co-operative co-evolutionary algorithm for minimizing carbon footprint and maximizing line efficiency in robotic assembly line systems, *Journal of Cleaner Production*, Vol. 156, 124-136, doi: [10.1016/j.jclepro.2017.04.032](https://doi.org/10.1016/j.jclepro.2017.04.032).
- [23] Zhang, B., Xu, L., Zhang, J. (2020). A multi-objective cellular genetic algorithm for energy-oriented balancing and sequencing problem of mixed-model assembly line, *Journal of Cleaner Production*, Vol. 244, Article No. 118845, doi: [10.1016/j.jclepro.2019.118845](https://doi.org/10.1016/j.jclepro.2019.118845).

- [24] Sun, B.-Q., Wang, L., Peng, Z.-P. (2020). Bound-guided hybrid estimation of distribution algorithm for energy-efficient robotic assembly line balancing, *Computers & Industrial Engineering*, Vol. 146, Article No. 106604, doi: [10.1016/j.cie.2020.106604](https://doi.org/10.1016/j.cie.2020.106604).
- [25] Zhou, B., Wu, Q. (2020). Decomposition-based bi-objective optimization for sustainable robotic assembly line balancing problems, *Journal of Manufacturing Systems*, Vol. 55, 30-43, doi: [10.1016/j.jmsy.2020.02.005](https://doi.org/10.1016/j.jmsy.2020.02.005).
- [26] Kennedy, J., Eberhart, R. (1995). Particle swarm optimization, In: *Proceedings of ICNN'95 - International Conference on Neural Networks*, Perth, Australia, 1942-1948, doi: [10.1109/ICNN.1995.488968](https://doi.org/10.1109/ICNN.1995.488968).
- [27] Otto, A., Otto, C., Scholl, A. (2013). Systematic data generation and test design for solution algorithms on the example of SALBPGen for assembly line balancing, *European Journal of Operational Research*, Vol. 228, No. 1, 33-45, doi: [10.1016/j.ejor.2012.12.029](https://doi.org/10.1016/j.ejor.2012.12.029).
- [28] Yuan, B., Zhang, C., Shao, X. (2015). A late acceptance hill-climbing algorithm for balancing two-sided assembly lines with multiple constraints, *Journal of Intelligent Manufacturing*, Vol. 26, 159-168, doi: [10.1007/s10845-013-0770-x](https://doi.org/10.1007/s10845-013-0770-x).
- [29] Li, Y., Peng, R., Kucukkoc, I., Tang, X., Wei, F. (2020). System reliability optimization for an assembly line under uncertain random environment, *Computers & Industrial Engineering*, Vol. 146, Article No. 106540, doi: [10.1016/j.cie.2020.106540](https://doi.org/10.1016/j.cie.2020.106540).

Predicting the deep drawing process of TRIP steel grades using multilayer perceptron artificial neural networks

Sevšek, L.^a, Vilkovský, S.^b, Majerníková, J.^b, Pepelnjak, T.^{a,*}

^aForming Laboratory, Faculty of Mechanical Engineering, University of Ljubljana, Ljubljana, Slovenia

^bDepartment of Technologies, Materials and Computer Aided Production, Technical University of Košice, Košice, Slovakia

ABSTRACT

TRIP (Transformation Induced Plasticity) steels belong to the group of advanced high-strength steels. Their main advantage is their excellent strength combined with high ductility, which makes them ideal for deep drawing processes. The forming of TRIP steels in the deep drawing process enables the production of a thin-walled final product with superior mechanical properties. For this reason, this study presents comprehensive research into the deep drawing of cylindrical cups made from TRIP steel. The research focuses on three main aspects of the deep drawing process, namely the sheet metal thinning, the maximum force value and the ear height as a result of the anisotropic material behaviour. Artificial neural networks (ANNs) were built to predict all the mentioned output parameters of the part or the process itself. The ANNs were trained using data obtained from a sufficient number of simulations based on the finite element method (FEM). The ANN models were developed based on variable material properties, including anisotropic parameters, blank holding force, blank diameter, and friction coefficient. A good agreement between simulation, ANN and experimental results is evident.

ARTICLE INFO

Keywords:

Forming;
Deep drawing;
TRIP steel;
Artificial neural network (ANN);
Finite element methods (FEM);
Modelling;
Simulation

*Corresponding author:

Tomaz.Pepelnjak@fs.uni-lj.si
(Pepelnjak, T.)

Article history:

Received 25 March 2024

Revised 24 April 2024

Accepted 26 April 2024



Content from this work may be used under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

The development of materials for vehicle production faces the challenge of solving diverse and often conflicting requirements. These include the need to reduce vehicle weight while providing high levels of weldability, formability, joinability, and sufficient stiffness or mechanical strength to increase the safety of passengers. In addition, it is important to optimize the simple installation of components, maintain surface resistance in demanding conditions, and preserve aesthetic properties. In today's automotive environment, meeting fuel efficiency and environmental standards within economic constraints is critical [1-3]. Cost reduction can be considered one of the main interests for companies in the modern era [4]. Different advanced forming methods have been developed to make custom made goods at a reasonable price [5]. Given these complex requirements, the choice of materials plays a key role in shaping the car's structure. High-strength steels (HSS) have demonstrated the ability to strike a balance between cost-effectiveness, weight consideration and favourable mechanical properties [6, 7].

Current research is focused on advanced high-strength steels (AHSS) and ultra-high-strength steels (UHSS), mainly composed of martensitic steels. The AHSS category includes two-phase steels and TRIP (Transformation Induced Plasticity) steels. These groups of materials exhibit mechanical properties that meet stringent environmental standards and customer requirements while remaining economically affordable for manufacturers [1-3].

1.1 Fundamentals of TRIP steels

TRIP steels hold significant appeal for the automotive industry in constructing the body-in-white due to their enhanced mechanical properties, formability, and exceptional energy absorption during a crash [8]. The microstructure of TRIP comprises distinct phases, including ferrite, austenite, martensite, and bainite, contributing to the high performance of the steel [9]. The TRIP effect, employed to improve formability and strength, is rooted in a lattice transformation. The face-centred cubic (fcc) austenite transforms, without diffusion, into either the body-centred cubic (bcc) martensite (α -martensite) or a hexagonal martensite phase (β -martensite). The bcc crystal phase is more stable, and the hexagonal phase transitions to bcc depending on the conditions [10, 11]. These diverse phases are not only beneficial for forming processes but also play a crucial role in the crash behaviour of components. However, as mechanical properties are directly linked to microstructural evolution, local stress and strain conditions impact both forming and failure behaviour. Therefore, understanding the interaction among microstructure, stress and strain conditions, and the occurrence of failure is essential [12].

1.2 Fundamentals of deep drawing

Deep drawing is a sheet metal forming process which are used to plastically deform the sheet metal into the desired shape of the final product [13]. The deep drawing process, widely utilized across industries, involves forming sheet metal for a variety of applications. Its uses span from crafting automotive components, products in the arms industry, and aerospace parts to forming tubes for medication and perfume, pots, pans, and various kitchen appliances. This method relies on forming production parts with either simple or intricate designs through substantial plastic deformation [14, 15]. Throughout the deep drawing process, potential defects may arise, such as surface scratches, wrinkling of walls and flanges, tearing, and earing [16, 17]. The effectiveness of the deep drawing process is heavily dependent on material properties, geometric considerations, and technological parameters. Key material factors encompass elasticity, plasticity, and anisotropy. Additionally, significant roles are played by parameters like punch velocity, blank holding pressure, and lubrication. The radius of the punch and die, blank thickness, and the clearance between the punch and die also contribute to the process. Incorrectly defining these parameters can result in common defects of the deep drawing process [18, 19].

Earing, a distinctive defect marked by the development of a wavy edge at the open end of the cup, is a notable issue in the deep drawing process. Numerous recent studies have addressed earing defects [20, 21]. Colgan and Monaghan [22] explored key parameters in the deep drawing process, including punch and die radii, punch velocity, friction, and draw depth. Using ANOVA software, they calculated the percentage contributions of each factor. Interestingly, punch velocity emerged as the fourth most crucial parameter, particularly impacting wall thickness deviation. Seth *et al.* [23] investigated the formability of steel sheets subjected to high-velocity impact from an electromagnetically launched punch at speeds of 50 and 220 m/s. Their experiments involved five different steel specimens with varying thicknesses, revealing failure strains ranging from 30% to 50%. Notably, the study observed that higher forming velocities correlate with increased formability. Huang *et al.* [24] examined the influence of blank thickness and fracture thickness on forming limits. Fracture strain, derived from a uniaxial tension test, served as the fracture criterion. The assumption was made that the blank fractures when its thickness reaches the fracture thickness in any section. The study involved the analysis of square cup drawing and elliptical hole flanging. Chalal *et al.* [25] determined forming limit diagrams using different localization criteria, such as the criterion based on the maximum second-time derivative of thickness strain, the criterion based on the ratio of equivalent plastic strain increment, the maximum punch force criterion, and the loss of ellipticity. ABAQUS/Explicit tool was employed for their analysis [25]. Gusel *et al.* [26]

focused on the forming of high-strength steel sheets (DP780 and DP1180HD) by deep drawing. The influence of yield stress, tensile strength, blank holder force and punch speed on the output parameter of cup height was evaluated [26]. The fracture problems during deep drawing were successfully predicted using genetic modelling.

1.3 Simulation using finite element method

Simulations allow the prediction of crucial outcomes within the system or process itself, using the computer environment. A correctly created simulation model of a forming process enables fast and reliable prediction of key output parameters. A small discrepancy between experimental and simulation results shows the general correctness of the set simulation model. Simulations were carried out as part of this study to predict three key output parameters of the deep drawing process of TRIP steel. A sufficient number of simulations with different input parameter values and corresponding output results were used as the basis for building an artificial neural network (ANN) for each target output parameter. The simulations based on the finite element method (FEM) were carried out in the ABAQUS simulation environment. The material properties, the geometric properties of the blank and the tools as well as the kinematics of the deep drawing process were simulated. FEM simulations can be carried out with different calculation methods, finite element types and in different simulation environments [27]. The FEM simulation of the deep drawing process can be very time-consuming [28]. The use of the explicit calculation method provides acceptable accuracy with faster calculation compared to the implicit method [29]. For this reason, the explicit method was chosen for all 50 simulations in the study presented here.

The study by Vrh *et al.* [30] focused on the constitutive modelling of anisotropic plates and the prediction of the earing for the round cup drawing using the FEM method. Shell finite elements with reduced integration were used in the ABAQUS/Explicit simulation environment [30]. Bandyopadhyay *et al.* [31] evaluated the limiting drawing ratio (LDR) of tailor welded blanks using a deep drawing test. FEM models of the deep drawing process considered the anisotropy of the sheets and the inhomogeneous properties in the welded zone of the tailor welded blanks [31]. The simulations were performed with the nonlinear solver Lsdyna-971 using shell elements [31]. The study by Dwivedi and Agnihotri [32] focused on testing different materials for deep drawing of cylindrical cups without using a blank holder. The ANSYS 14.0 simulation environment was used to determine the limit drawing ratio [32]. Magnesium alloys develop a crystallographic texture and plastic anisotropy during rolling, resulting in ear formation during deep drawing of such sheets [33]. This was investigated in the study by Walde and Riedel [33], who performed FEM simulations of the deep drawing process in ABAQUS/Explicit environment using C3D8R solid brick elements with four elements across the sheet thickness. It was found that the earing pattern depends on the initial texture and on the development of the texture during the forming process under investigation [33].

In the study by Engler and Aretz [34], S4R shell elements with reduced integration were used in the ABAQUS simulation environment for the FEM simulation of the deep drawing process of various anisotropic aluminium alloys. In the study by Luyen *et al.* [35], FEM simulations of the deep drawing of cold-rolled carbon steel were carried out in the ABAQUS simulation environment. The simulations were used to determine the fracture heights of cylindrical cups and were validated with experimental results [35]. Another aim of the simulations and experiments in the study by Luyen *et al.* [35] was to investigate the effects of the blank holder force, the punch radius and the drawing ratio on the fracture height. Simulations and experiments showed that increasing the blank holder force reduces the fracture height and increasing the punch radius increases the fracture height [35]. Jayahari *et al.* [36] investigated the formability of austenitic stainless steel 304 at different temperatures under warm conditions during the deep drawing process of a cylindrical cup. The explicit FEM analysis was performed in the LS-DYNA program for forming at room temperature and other temperatures up to 150°C [36]. Shell elements were used for the blank and the tools, as this results in a shorter calculation time [36]. Shell finite elements were also used in the study presented here.

1.4 Fundamentals of neural networks

Artificial intelligence can contribute to sustainable manufacturing [37]. With advances in manufacturing intelligence, information technology is part of the automated production technologies that are the main contributors to industrial change today [38]. Artificial intelligence models used in the field of metal forming use process parameters as input parameters and forming results as output parameters [37]. One special branch of such models are neural network models, which are expressed by a nonlinear function of the weighted sum of inputs [37]. Artificial neural network (ANN) is a set of neurons and connections between them with adjustable weights [39]. Commonly used type of ANNs are the multi-layer neural networks, which consist of an input layer, one or more hidden layers, and an output layer [39, 40]. Neurons of each layer are connected to each other by connection links with adjustable weights [39]. These weights are adjusted during the training process of the neural network, commonly through the backpropagation algorithm, where the input-output example patterns are presented to the neural networks [39]. The main advantage of ANNs is the data-driven self-adaptive capabilities that allow ANNs to adapt to the data they have been trained on [41]. A basic structure of ANN is presented in Fig. 1.

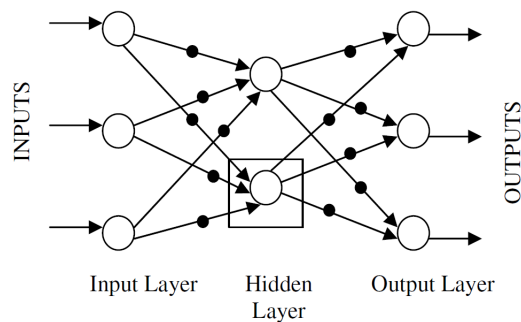


Fig. 1 Basic structure of ANN [39]

In order to teach the neural network, sufficient amount of data is needed. The data can either be provided by experiments or by simulations, including finite element method [39]. A large enough number of FEM simulations was also performed in here-presented study with a goal to provide data points on which the neural network is trained. One of the advantages of using ANN model is easy construction of said model on provided input and output data values with the goal of accurately predicting process dynamics [42].

Machine learning, deep learning, artificial intelligence and more specifically artificial neural networks are being used in many fields, including state of tools, defect detection, forming processes and material science [43-46]. In the study by Czinege and Harangozo [46], experimental data from tensile tests and Nakazima tests were used as input data for the ANN. The ANN models allowed estimation of points of the forming limit curve [46]. ANN models gave high correlation coefficient between predicted and measured values, which was better compared to capabilities of other linear and non-linear models [46]. Multi-layer perceptron (MLP) artificial neural networks are good for classification and for regression, which is also one of the reasons for being used in the study by Czinege and Harangozo [46]. MLP ANN requires a proper selection of the number of hidden layers and data splitting into training set, test set and validation set [46, 47]. In order to predict the flow curves of ZAM100 magnesium alloy sheets under hot-forming conditions as a function of process parameters, El Mehtedi *et al.* [48] developed an empirical model based on ANNs. ANN model predicted the flow stress as a function of strain, strain rate and temperature [48]. The ANN from the study by El Mehtedi *et al.* [48] had 6 input parameters and one output of equivalent stress, and it also had two hidden layers with 6 neurons each. The back-propagation training and validation of the multi-layer feed forward ANN was performed using MATLAB software [48]. Great capabilities of the set model were proven with excellent fitting between experimental and predicted curves [48]. For the reason of evaluating the prediction capabilities, the correlation coefficient (R) was considered, which compares predicted values and experimentally produced data [48]. The model for predicting flow curves requires a sufficient number of experiments to obtain the necessary data on which it can be trained, which is expensive and time-consuming

[48]. ANN models are often used also in the field of metal forming. In the study by Gondo and Arai [37] an ANN was developed for metal spinning using tool-path parameters, the size of the blank, size of the tools and the height and the thickness of the part. In the study by Xia *et al.* [45] multi-layer feed-forward perceptron ANN models with back-propagation were used to investigate the influence of rolling parameters on the rolling force, rolling power, and slip of tandem cold rolling. Different ANN architectures had been tested and the ANN configuration of one hidden layer with 9 neurons provided the best results [45]. In the study by Kazan *et al.* [39] a prediction model using ANN was developed in the field of wipe-bending. The training of the ANN model was done on the data provided by FEM simulations [39]. In the study by Sivasankaran *et al.* [42] a feed forward back-propagation neural network was presented for predicting and avoiding surface failure, including wrinkling, during pure aluminium sheet drawing through a conical die. The neural network had two hidden layers with different numbers of neurons for different grades of aluminium sheets tested [42].

The study by Babu *et al.* [49] focused on developing an expert system using ANN to predict the deep drawing behaviour of welded blanks made of steel grade and aluminium alloy. FEM code is used for forming simulation and data generation for ANN training [49]. ANNs with one hidden layer with 6 neurons, 6 neurons on the input layer and one neuron on the output layer are used for four different output parameters, including depth of drawing [49]. In the study by Manoochchri and Kolahan [50] an ANN was developed based on FEM results in the case of deep drawing, where important process parameters as inputs and process characteristics as outputs were considered. FEM models were verified with experimental tests using same parameter values, including stainless steel 304 (AISI 304) with 0.5 mm thickness with its mechanical properties and anisotropic coefficients [50]. In the study by Manoochchri and Kolahan [50] ABAQUS/Explicit software was used to develop FEM models of the deep drawing process, where the depth of drawing was 30 mm. The input parameters considered were blank holder force, punch radius, die radius, friction coefficient between punch and blank as well as die and blank, while the output parameter was minimum sheet thickness after forming [50]. An ANN with two hidden layers was developed in MATLAB software [50].

2. Material and methods

2.1 Used materials

TRIP steels prove highly effective in manufacturing automotive components that undergo significant work hardening during crash deformation and necessitate substantial energy absorption. Additionally, these steels are particularly well-suited for forming intricate and challenging to form parts due to their exceptional formability and hardening characteristics. In the experimental research, a double-sided galvanized steel sheet TRIP RAK40/70 Z100MBO with a thickness of 0.75 mm was utilized. To determine the mechanical properties, a uniaxial tensile test was conducted on the TIRAtest2300 machine (Fig. 2) according to the STN EN ISO 6892-1:2020 standard. During the examination, 5 test samples were assessed in three directions: 0°, 45°, and 90° degrees with respect to the rolling direction (see Fig. 3). The mechanical properties obtained using uniaxial tensile test are displayed in Table 1. The chemical composition of the experimental material TRIP RAK40/70 Z100MBO is shown in Table 2.

Table 1 Mechanical properties of TRIP RAK40/70 Z100MBO

RD (°)	$R_{p0.2}$ (MPa)	R_m (MPa)	A_{80} (%)	r (-)	r_m (-)	Δr (-)	n (-)
0	435	764	29	0.702			0.298
45	443	763	29	0.884	0.834	-0.100	0.294
90	449	764	31	0.867			0.279

Table 2 Chemical composition of TRIP RAK40/70 Z100MBO wt. %

C	Mn	Si	P	S	Al	Nb	Ti	V	Mo	Cr
0.204	1.683	0.199	0.018	-	1.731	0.004	0.009	0.004	0.008	0.055



Fig. 2 TIRAtest2300 testing machine

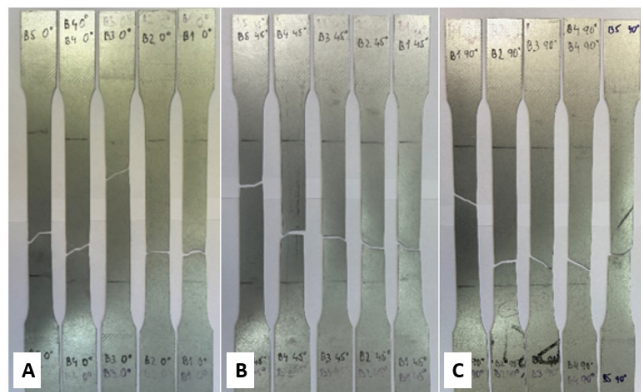


Fig. 3 Samples after uniaxial tensile test: a) in the RD direction 0°; b) in the RD direction 45°; c) in the RD direction 90°

2.2 Deep drawing test

Deep drawing is a sheet metal forming technique where a sheet metal blank is radially drawn into a forming die by mechanical energy. This process is notably influenced by material properties, as well as geometric and technological parameters. Key material parameters include elasticity, plasticity, and anisotropy. Moreover, parameters like punch velocity, blank holding force, and lubrication are crucial, besides the radius of the punch and die, the thickness of the blank, and the clearance between the punch and die. Improperly defining these parameters can lead to defects such as surface scratches, wrinkling, and tearing due to excessive thinning. The paper focuses on analysing how various technological and material parameters affect the key characteristics of the product and its forming process. These characteristics include deep-drawing force, earing, and thinning.

In the experimental research, three circular blanks with a diameter of 95 mm were cut from the sheet metal strip, from which cylindrical cups with a flat bottom were subsequently drawn on an experimental deep drawing tool (Fig. 4a) with the parameters shown in Fig. 4b. In the deep drawing test, oil was used as a lubricant to reduce the coefficient of friction. The holding force was set to 18000 N.

During the deep drawing process, it is essential to comprehend the maximum deep drawing force and the impact of input parameters on its magnitude. This understanding is fundamental for selecting the appropriate machine with the designated force. To measure the force during our experiment, we employed a load cell integrated into the forming press. During the deep drawing test, the load cell recorded and transmitted force data to a connected data acquisition system. The force was monitored and recorded throughout the entire process of deep drawing.

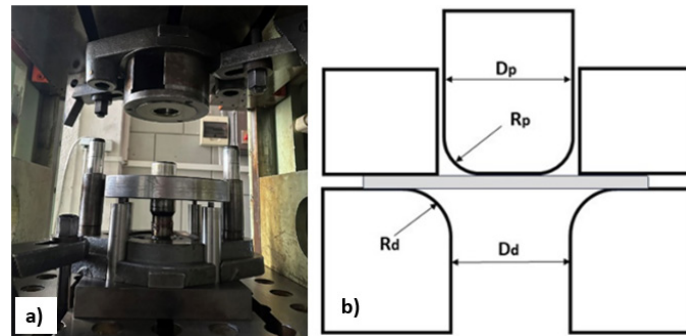


Fig. 4 a) Experimental deep drawing tool, b) parameters of the tool, where: D_p (punch diameter) = 50 mm, D_d (die diameter) = 52.4 mm, R_p (punch radius) = 5 mm, R_d (die radius) = 5.5 mm

Furthermore, we focused on understanding the influence of the parameters on earing, characterized by the development of a wavy edge at the open end of the cup. In this paper, earing was assessed as the difference between the highest (H_{max}) and lowest (H_{min}) measured heights on the cups, as defined by Eq. 1, where ΔH represents ear height.

$$\Delta H = H_{max} - H_{min} \quad (1)$$

According to the value of Δr , it is possible to determine the susceptibility of the sheet to the formation of ears during deep drawing. Ears are formed in the directions of the sheet where the value of the coefficient of normal anisotropy r is maximum, if:

- $\Delta r > 0$ ears will form in the directions of 0° and 90° to the rolling direction,
- $\Delta r = 0$ ears will not form,
- $\Delta r < 0$ ears will form in the direction of 45° .

The height of the ears (Fig. 5a) was measured with a sliding calliper at eight locations on the cups (Fig. 5b). Subsequently, according to Eq. 1, the maximum difference in the height of the cups was calculated.

We also focused on the thinning of the sheet, which was expressed by the t_{min} value, which represented the lowest measured value of the sheet thickness after forming. Thinning of the blank typically occurs at the transition point from the cylindrical part to the bottom of the blank. Excessive thinning in these areas can lead to the formation of cracks. The thinning itself is affected by the initial thickness of the sheet metal and by the diameter of the tool. When assessing the impact of these parameters on thinning, it is crucial not to overlook the material parameters.

After measuring the heights, the experimental cups were cut in half, allowing us to measure the thickness of the cups and obtain the values of thinning across the cross-section. The thinning of the sheet was measured using an optical microscope (Fig. 6a) at the transition point between the cylindrical part and the bottom of the cup as shown in Fig. 6b.

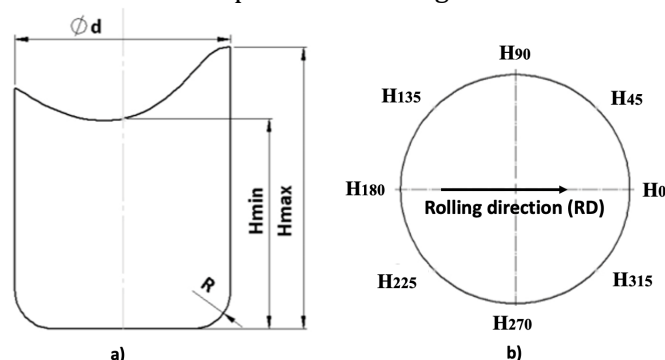


Fig. 5 a) Ear height on the cup, b) Measurement of ear heights on the cup

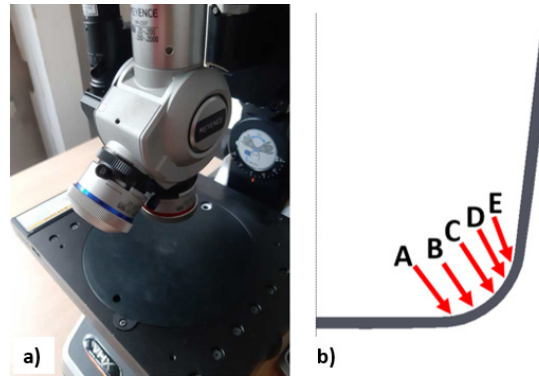


Fig. 6 a) Microscope used for thickness measurement, b) thickness measurement points

2.3 Finite element method setup

The simulation of the deep drawing process was performed using the ABAQUS simulation software environment, utilizing the explicit method. The CAD model of the drawing tool used in the simulation had identical dimensions to the experimental tool (see Fig. 4b). The blank model was given the characteristics of TRIP steel, which were acquired through the uniaxial tensile test (see Table 1). The material density was 7850 kg/m^3 , the Young's modulus was $210,000 \text{ MPa}$, and the Poisson ratio was 0.33 . During the simulations, the Hollomon approximation was used to describe the flow curve of the material. The sheet material was considered anisotropic using yield criteria according to Hill. The simulation time for the explicit simulation method was set to 0.5 seconds.

According to the finite element method, the simulation of technological processes requires the meshing of the entire set of 3D objects (rigid as well as deformable models) with finite elements. For all rigid bodies (punch, die, and blank holder), R3D4 quadrilateral shell finite elements and R3D3 transition triangle elements were used. In terms of workpiece meshing, concentric mapped quadrilateral finite elements S4R were chosen for the outer section, while a freely meshed combination of quadrilateral finite elements (S4R) and triangular finite elements (S3) was utilized for the central part.

Since the size of the blank was one of the parameters that was regularly changed in each simulation, the number of elements was different in each simulation. Table 3 shows the number and types of elements used in simulations for every 3D model used in simulations. Fig. 7 shows the mesh of the blank used in the simulation.

Table 3 Number and types of elements used in simulations

Object	Number of elements	Type of element
Punch	1634	R3D4, R3D3
Die	4200	R3D4
Blank Holder	2040	R3D4, R3D3
Blank	varied	S4R, S3

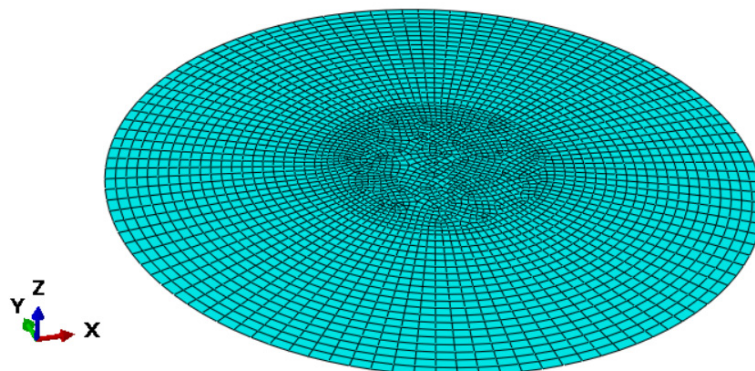


Fig. 7 Mesh of blank CAD model

The procedure involved conducting 50 simulations in which 8 independent input parameters were randomly varied. These parameters include blank diameter (D_0), blank holding force (F_{bh}), friction coefficient (f), material constant (C), hardening exponent (n), and normal anisotropy coefficient values (r) in the 0° , 45° , and 90° directions. Additionally, one dependent parameter, the yield strength (R_p). Table 4 displays the range of input parameters within which values were randomly selected for individual simulations. Values from these intervals were selected randomly for each simulation using the random function in Microsoft Excel. Table 5 shows the values of the input parameters used in the simulations.

Table 4 Range of input parameters used in simulations

Parameter	Minimal value	Maximal value
D_0 (mm)	86	102
F_{bh} (N)	16000	20000
f (-)	0.05	0.20
R_p (MPa)	300	700
C (MPa)	900	1800
n (-)	0.2	0.35
r_0 (-)	0.6	1
r_{45} (-)	0.6	1
r_{90} (-)	0.6	1

Table 5 Randomly selected input values used in individual simulations

Simulation number	D_0 (mm)	F_{bh} (N)	f (-)	R_p (MPa)	C (MPa)	n (-)	r_0 (-)	r_{45} (-)	r_{90} (-)
1	88.4	17778	0.11	491	1512	0.23	0.93	0.74	0.89
2	96.9	18078	0.15	308	1038	0.28	0.64	0.8	0.65
3	100.4	19788	0.07	418	1372	0.27	0.63	0.62	0.75
4	90.6	18082	0.19	308	1112	0.3	0.99	0.7	0.72
....
50	87.8	16560	0.1	385	1076	0.23	0.92	0.69	0.95

2.4 Artificial neural network

With the aim of making predictions of forming force value, sheet metal thinning and ear height for the deep drawing process and product, artificial neural networks (ANNs) had been set up. For this study, a separate ANN was set up for every output parameter that considered all 9 input parameters. In order to allow faster training of the neural network and good regression of the results, multilayer perceptron artificial neural networks (MLP ANN) had been used. This is an ANN model that used backpropagation learning algorithm. The basis of MLP ANN remains the same, meaning transformation of input values into output values or predictions. To better understand the process of MLP ANN, Eq. 2 shows the basic function, where y is the output value, x is the input value, w are weights and b are biases.

$$y = f(net) = f\left(\sum_{i=1}^n w_i x + b\right) \quad (2)$$

The basis for training an ANN is a sufficient number of data points, which was in our case provided by 50 finite element method simulations with different values of input parameter values and their corresponding values of output parameters of forming force, minimum sheet thickness and ear height. To set up the architecture for all three ANNs of three different output parameters, data had to be divided into subsets for training and for testing. A preliminary investigation showed that the best possible division of the data is 90 % for training and the rest for testing, meaning 45 points are used for training and five points out of 50 are used for testing.

Once the lengthy process of training is complete, the set ANNs can be used for prediction purposes. All three ANNs of three considered output parameters are based on the feed-forward multilayer perceptron (MLP) and are thus divided into layers. On every layer there is at least one neuron. The layers present within an MLP ANN are the input, output and at least one hidden layer. On the input layer there are as many neurons as there are input parameters, which in the case of

our study is equal to 9. In the case of our study there is only one neuron on the output layer of all three neural networks which corresponds to one output parameter of either forming force value, minimum sheet thickness or ear height. Within the preliminary study the number of hidden layers was set to five for every neural network and their corresponding output parameter. The same is true for the number of neurons on these five hidden layers, where 60 neurons had been chosen for all five hidden layers for all three neural networks. The configuration of hidden layers and corresponding number of neurons can be written as (60, 60, 60, 60, 60).

Within the neural network the value leaving a specific neuron is affected by the weight and the bias value before entering a specific neuron of the next layer. An important role is played by the activation function which removes certain values and maps modified remaining values. Many different activation functions had been tested in a preliminary study and a decision had been made to use sigmoid or logistic activation function on the neurons of the hidden layers and linear or identity activation function to be used on the output layer of all three set ANNs within this study. Eq. 3 represents sigmoid and Eq. 4 the linear activation function. Additionally, Adaptive Moment Estimation (ADAM) solver type had been used in MATLAB program environment, used for training of the ANNs and for predictions made used trained ANNs. Initial learning rate value was chosen to be 0.1 for all set ANNs and the value of the learning rate remained constant throughout the neural network training process. L2 regularisation was used, and its value chosen as 0.001 for all ANNs within the study. The essential hyperparameter values for all three ANNs are written in Table 6. For a better understanding of the ANN configurations for this study, Fig. 8 shows a schematic presentation of all the neurons within input, output and hidden layers. With all the chosen values of hyperparameters that define the ANN architecture, the training of three different ANNs of all three tested output parameters has been carried out.

$$f(x) = \frac{1}{1 + e^{-x}} \tag{3}$$

$$f(x) = x \tag{4}$$

Table 6 Hyperparameters chosen for three ANNs within the study for the corresponding output parameters of forming force, minimum sheet thickness and ear height

Hidden layer sizes	Activation function	Solver	Initial learning rate	Learning rate type	L2
(60 60 60 60 60)	Sigmoid	ADAM	0.100	Constant	0.001

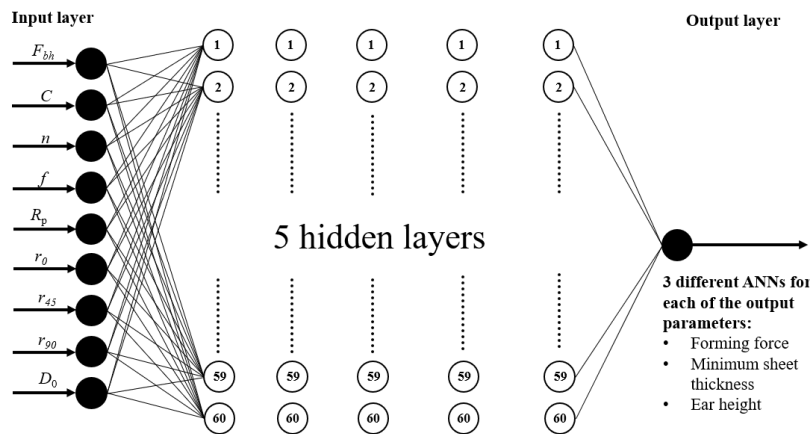


Fig. 8 Schematic representation of the ANNs used within here presented study; input layer with 9 neurons of the 9 tested input parameters, 5 hidden layers with 60 neurons each and an output layer with one neuron for one output parameter within a single ANN.

3. Results and discussion

As part of the deep drawing test, three cups with a flat bottom were drawn out of circular blanks with a diameter of 95 mm. Three parameters were examined: deep drawing force, minimal thickness, and ear height. Fig. 9 shows cups from TRIP RAK40/70.



Fig. 9 Cups made of TRIP RAK40/70 steel after deep drawing test

Within the framework of the experiment deep drawing forces with values of 84315 N for cup 1, 84490 N for cup 2 and 84490 N for cup 3 were measured. These values were measured by subtracting the holding force from the maximum force measured. The plot of the forces is shown in Fig. 10.

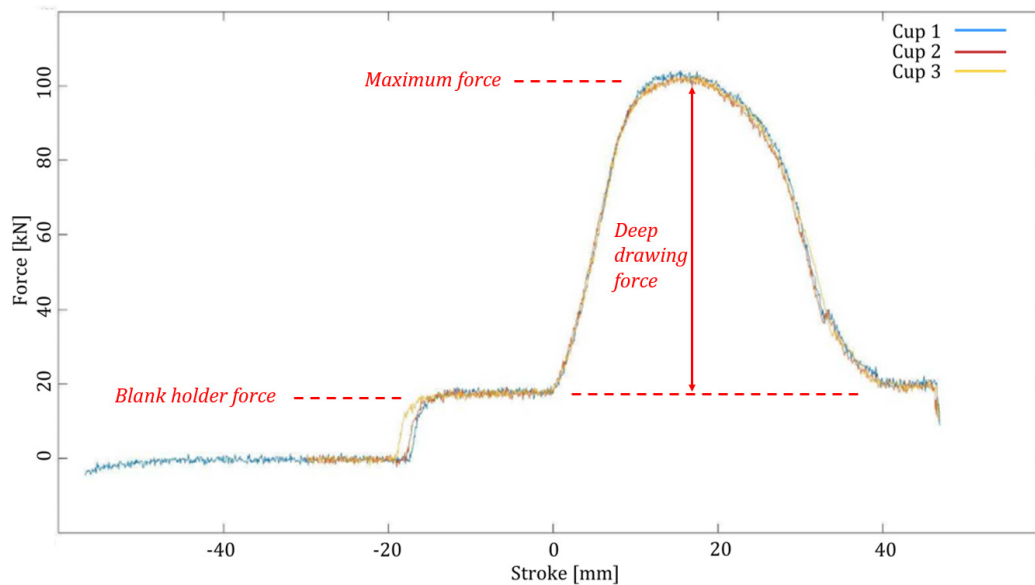


Fig. 10 The course of measured force during the experiment of deep drawing

The measured values of the ear heights in the eight measured directions are shown in Table 7. The ΔH value (the difference between the maximum and minimum ear height) was calculated for each cup, and the average value was subsequently determined.

As shown in Table 7, the most significant disparity in the measured heights occurred at cup 3, where the discrepancy between the maximum and minimum height reached 1.31 mm. The arithmetic average for the parameter ΔH was calculated to be 1.23 mm.

The sheet thickness after the deep drawing test was measured at five points along the cross-section of the cups at the transition from the cylindrical part to the bottom of the cups. The examined sheet thicknesses after the deep drawing tests are shown in Table 8. From the measured thicknesses, the smallest obtained value (t_{\min}) was evaluated. From the value t_{\min} , the arithmetic average was calculated within the three cups. Fig. 11 shows the measurement of sheet metal thickness at the bottom radius of the second cup.

Table 7 Ear height values of the cups in eight directions along with the calculated ΔH value

Cup number	H_0 (mm)	H_{45} (mm)	H_{90} (mm)	H_{135} (mm)	H_{180} (mm)	H_{225} (mm)	H_{270} (mm)	H_{315} (mm)	ΔH (mm)
1	33.61	34.11	33.07	33.17	32.93	33.25	32.86	33.84	1.25
2	32.95	33.96	33.58	34.09	33.22	33.58	32.98	33.65	1.14
3	33.65	34.52	33.67	34.02	33.53	33.95	34.02	34.84	1.31
								avg. ΔH (mm)	1.23

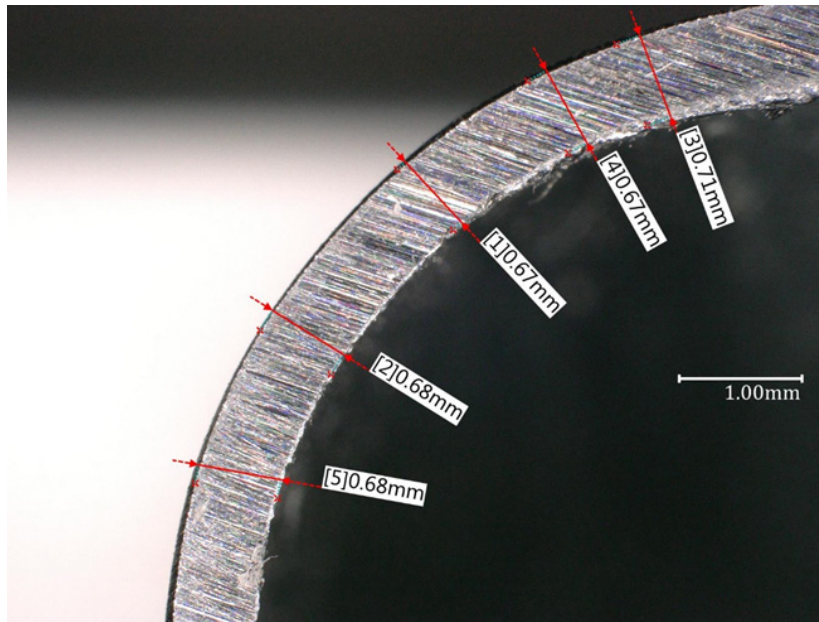


Fig. 11 Measurement of thinning at the bottom radius of the test cup

Table 8 Minimal thickness values of the cups in five measured points along with the t_{min} value

Cup number	t_A (mm)	t_B (mm)	t_C (mm)	t_D (mm)	t_E (mm)	t_{min} (mm)
1	0.65	0.64	0.64	0.64	0.67	0.64
2	0.71	0.67	0.67	0.68	0.68	0.67
3	0.65	0.62	0.62	0.62	0.65	0.62
avg. t_{min}						0.64
(mm)						

A deep drawing test simulation was performed with same parameter values as the performed experiments. The explicit method in the ABAQUS simulation program was used. The maximum deep drawing force, thinning value, and ear heights were measured in eight directions, from which the ΔH value was calculated using Eq. 1. Fig. 12 shows the cup shape after the performed simulation.

The results of the 50 simulations present the influence of 9 input parameters on the three monitored outputs, which are the maximum deep drawing force, minimum sheet thickness and earing. The results of 50 simulations in which the input parameters were randomly selected (see Table 5) from the interval of values (see Table 4) are shown in Table 9. As we can see, changing the 9 input parameters monitored significantly affects the observed output parameters.

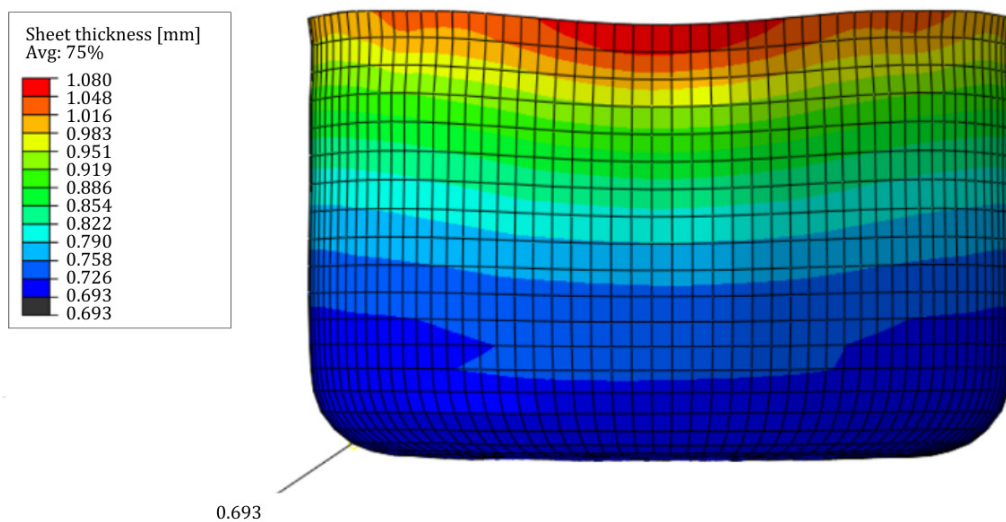


Fig. 12 Result of the simulation of deep drawing test using ABAQUS

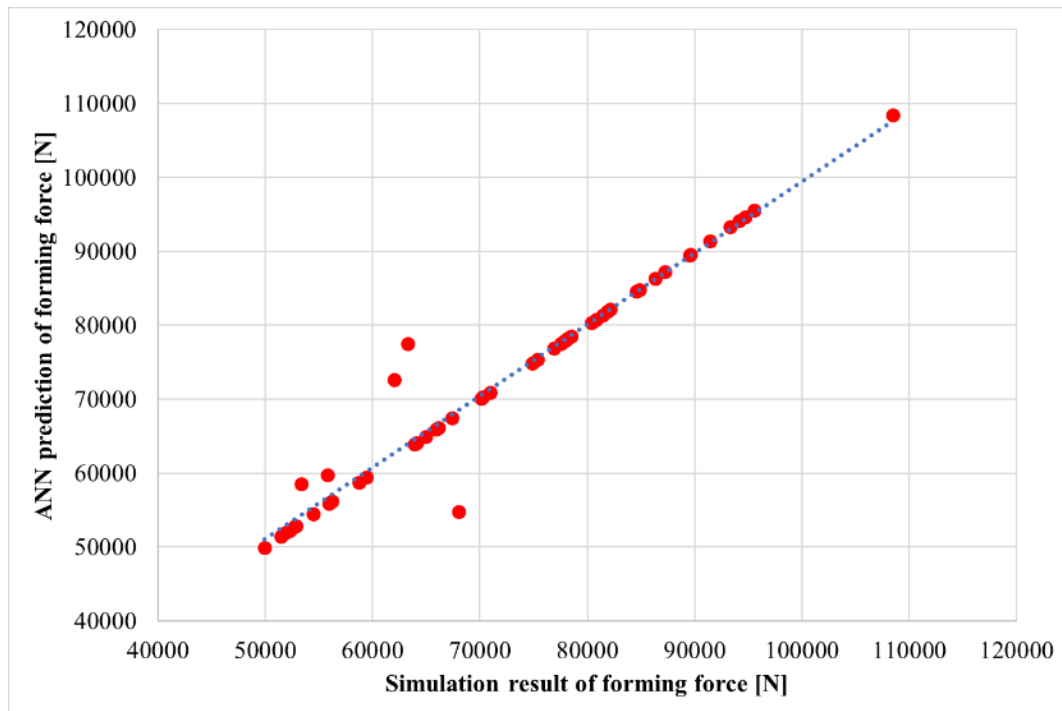
Table 9 Results of 50 simulations carried out in ABAQUS simulation software

Simulation number	F_{max} (N)	t_{min} (mm)	H_{max} (mm)	H_{min} (mm)	ΔH (mm)
1	77814	0.713	27.32	26.41	0.91
2	63907	0.641	35.26	33.92	1.34
3	84826	0.653	37.11	36.52	0.59
4	62045	0.672	30.01	28.79	1.22
....
50	54482	0.713	26.34	25.07	1.27

With all the chosen values of hyperparameters that define the ANN architecture, the training of three different ANNs of all three tested output parameters has been carried out. Trained neural networks allow for predictions to be carried out. Figs. 13, 14 and 15 show diagrams comparing original output values provided by simulations and predicted values using trained ANNs that were provided by using the same input parameter values. If the predictions perfectly matched the original values, then a straight line would be seen on the diagrams. But as we can see, a slight discrepancy is evident. To further understand the predictive capabilities of all three ANNs two coefficients were calculated, namely correlation coefficient R and coefficient of determination R^2 . Using the MATLAB program environment, the value for R squared was calculated automatically at the end of the ANN training process. R^2 value (Table 10) was calculated with the use of Eq. 5, where simulation output parameter values y_i are compared to ANN predicted values \hat{y}_i . For the predicted values using ANN, same input parameter values were used as in the case of the FEM simulations. To better understand the predictive capabilities of ANNs and to show the general correctness of the simulations performed, Fig. 16, Fig. 17 and Fig. 18 show a comparison of the values obtained with an experimental method, an explicit simulation and a neural network using the same input parameters (see Table 11).

Table 10 Evaluation of set ANNs for three different output parameters using R and R^2

Output parameter	R	R^2
Forming force	0.975	0.949
Sheet metal thickness	0.953	0.907
Ear height	0.926	0.844

**Fig. 13** Comparison of simulation results and predicted values by ANN for the output parameter of forming force value

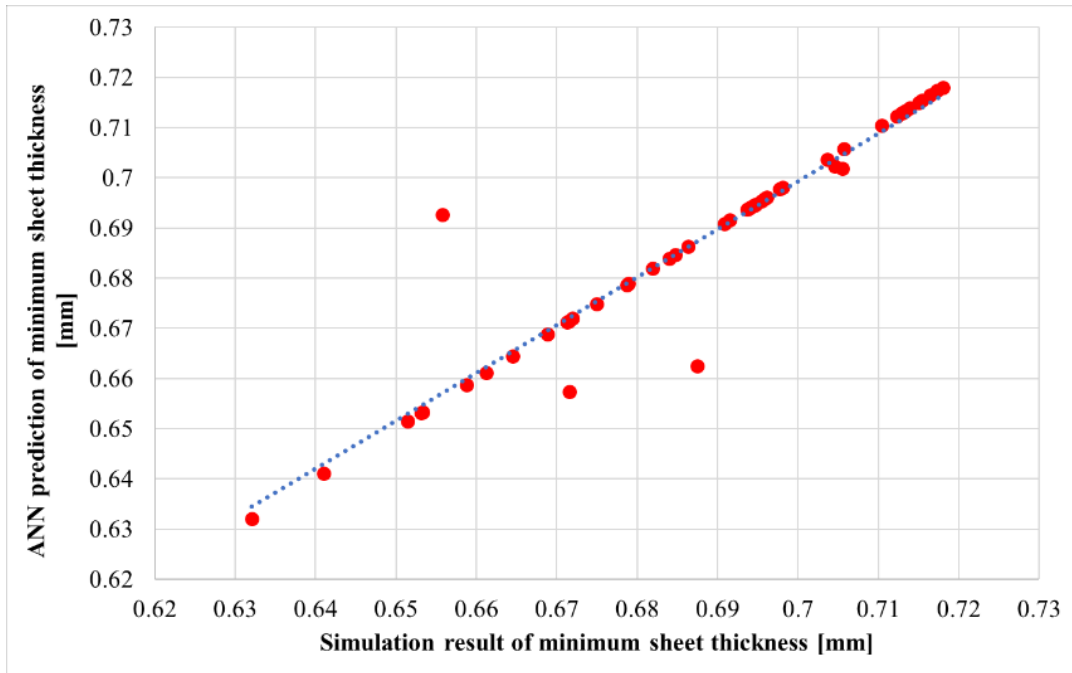


Fig. 14 Comparison of simulation results and predicted values by ANN for the output parameter of minimum sheet thickness

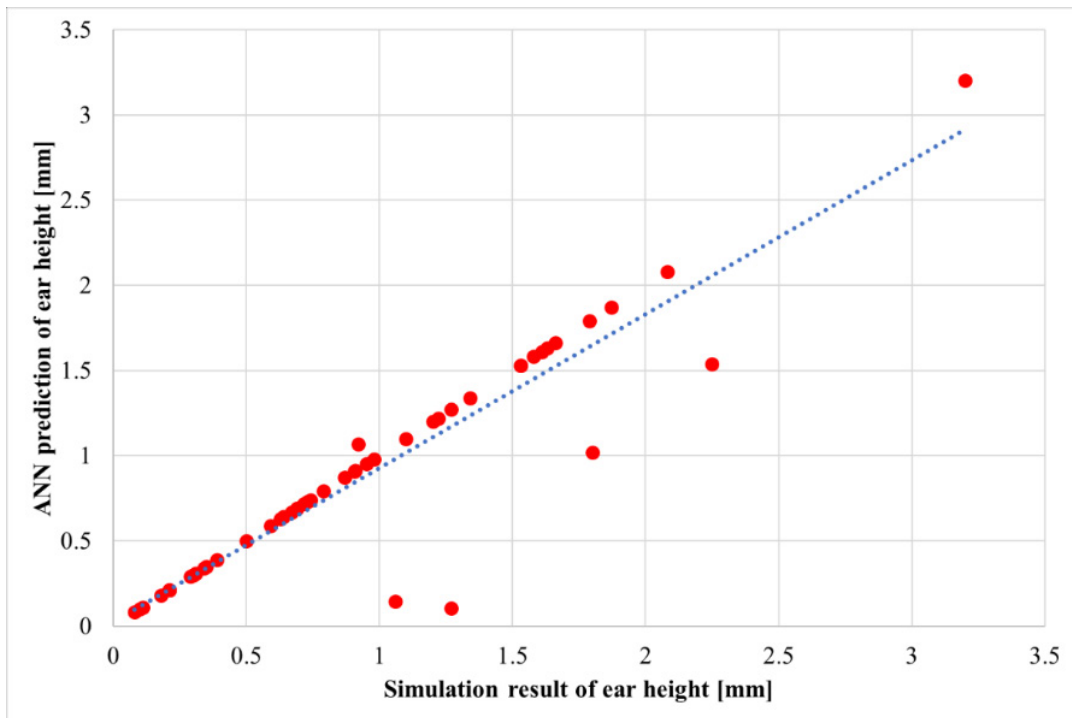


Fig. 15 Comparison of simulation results and predicted values by ANN for the output parameter of ear height

Table 11 Parameters used in the experiment, explicit simulation and ANN with results shown in Figs. 16 to 18

D_0 (mm)	F_{bh} (N)	f (-)	R_p (MPa)	C (MPa)	n (-)	r_0 (-)	r_{45} (-)	r_{90} (-)
95	18000	0.1	422.3	1488	0.28	0.702	0.884	0.867

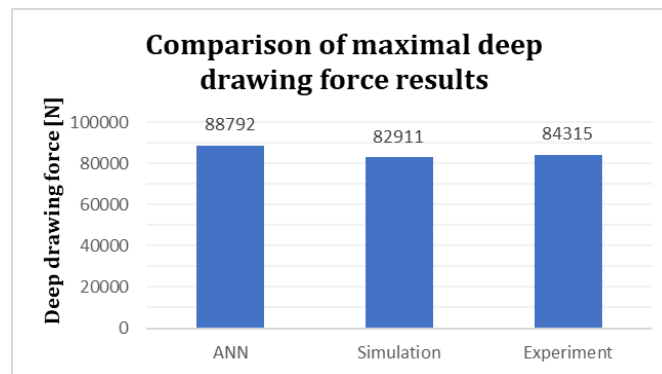


Fig. 16 Comparison of measured and calculated maximal deep drawing force values

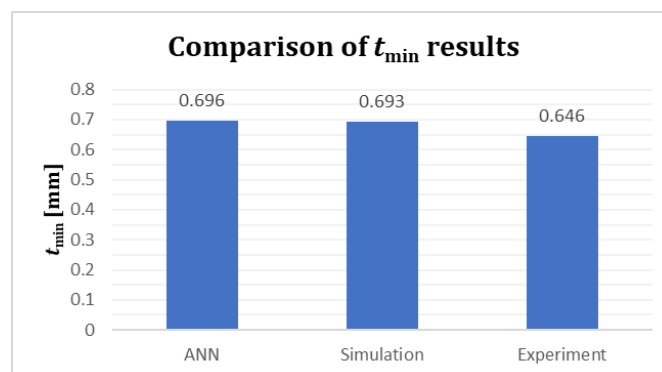


Fig. 17 Comparison of measured and calculated t_{\min} values

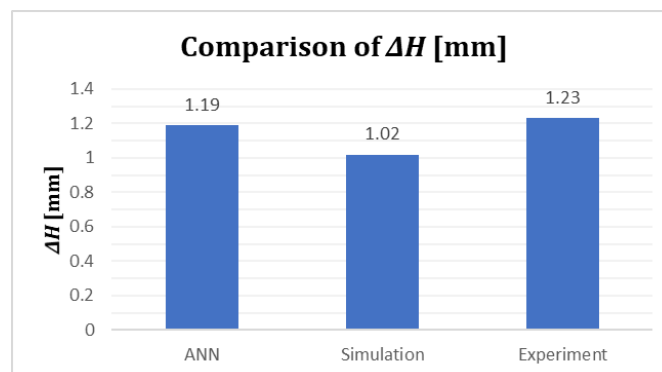


Fig. 18 Comparison of measured and calculated ΔH values

4. Conclusion

Optimisation of the deep drawing process is crucial due to its wide application in the automotive industry. Constantly increasing requirements for reducing emissions force designers to utilize high-strength materials, resulting in a reduction in the weight of structures. The demands on the industry are constantly growing, leading to the development of new methods for optimizing the deep drawing process. Artificial intelligence plays a pivotal role in this, with its increasingly frequent integration into technological processes. Thanks to constant improvement in this field, the output parameters of deep drawing can be predicted. In this study, ANN was used to predict the output parameters, the results of which were compared with experimental values and simulation using the explicit integration method.

The paper has focused on the analysis of various technological and material parameters that influence the key properties of the forming process. The results of the experiment demonstrate the potential for employing various methods and their integration to predict selected output parameters, such as evaluating the maximum deep drawing force, thinning and earing expressed

through the difference between the maximum and minimum yield height. In the experimental part of the research, a deep-drawing tool with defined geometry was used. Double-sided galvanized steel sheet TRIP RAK40/70 was chosen as the test material. TRIP steels represent a material that combines exceptional forming properties with strength.

In the simulations, eight independent input parameters and one dependent input parameter, which was the yield strength (R_p), were randomly varied. Independent input parameters include blank diameter (D_0), blank holding force (F_{bh}), coefficient of friction (f), material constant (C), strain hardening exponent (n) and normal anisotropy coefficient values (r) at 0° , 45° , and 90° directions. The experiment proved the correctness of the simulation of the deep drawing process for the prediction of the selected output parameters, considering the variability of the use of possible input parameters that affect the forming process. In order to predict the selected output parameters, artificial neural networks were developed. Each output parameter was modeled by its own dedicated ANN, considering all nine input parameters for this study. In order to provide a sufficient amount of data to train the ANNs, a sufficient number of simulations had to be carried out. Running a simulation took between one to two hours, which is a lengthy process. The training of the ANNs and the predictions with the trained ANNs were performed on a computer with an Intel Core i7-1065G7 processor with 3.6 GHz and 16 GB RAM. The training of the ANNs took several minutes, while the prediction process took less than one second after the input values of all input parameters were provided to the ANN.

As can be seen from Table 10, the prediction of the maximum deep drawing force using ANN achieved the best results within all evaluated parameters with a value of $R^2 = 0.949$. When entering the input parameters, which were identical to the performed experiment, the prediction using ANN reached the maximum deep drawing force of 88792 N. During the simulation, it was measured as 82911 N compared to the deep drawing force of 84315 N during the experiment. The same values of the input parameters used in the experiment and in the simulation and provided to the ANN are listed in Table 11.

One of the main difficulties in the deep drawing process lies in optimizing the thinning as it can lead to reduced mechanical strength in the final product. If the thinning reaches a critical value, it can lead to failure. As part of the experiments, thinning was evaluated on three cups. The measured minimal thickness (t_{min}) reached the value of 0.64 mm for cup 1, 0.67 mm for cup 2 and 0.62 mm for cup 3. When comparing the arithmetic mean of these three values, we register a difference of 0.047 mm compared to the results from the simulation using the explicit method. In all three experimental cups, a smaller wall thickness value was measured compared to the simulation, where t_{min} reached a value of 0.693 mm using same input parameter values listed in Table 11. When comparing original output values provided by 50 simulations and predicted values using trained ANNs that were provided same 50 sets of input parameter values, R^2 of 0.907 was calculated. The trained ANN predicted a minimum thickness of 0.696 mm when entering exact values of the experiment listed in Table 11. The most significant discrepancy in predicting the output parameters was measured in the case of ear height, evaluated using the ΔH parameter. When comparing the difference in maximum and minimum cup height, the deviation between simulation and experimental results was found to be 0.21 mm for the same input parameter values. ANN showed superior performance in this regard, with the difference narrowing to 0.04 mm. The calculated R^2 value for the ANN was 0.844 when comparing 50 results from simulations and predictions made with trained ANN using same input parameter values. With this value of R^2 parameter, we can state that this value was influenced by the manual measurement of heights after the experiments and simulations, that were utilized to train the ANN.

The conducted research provides a strong foundation for the continued utilization of the Finite Element Method and Artificial Neural Network techniques in the field of predicting the critical output parameters of deep drawing such as maximal deep drawing force, thinning and earing. With the aim of predicting the crucial output parameter values of the deep drawing process and providing additional control of the process itself, our future work will focus on the following:

- Analysis of different tool geometry for the production of a cylindrical cup with a smaller inner diameter. Experiments, simulations and ANNs will be set for the use of a smaller tool and compared with the results obtained with the tool in the study presented here.
- Analysis of different steel grades and their performance during the deep drawing process. Further extension of the predictive capabilities of ANN models for more extensive material parameter ranges.
- Analysis of deep drawing of more complex part shapes. Experiments and simulations for different tool and thus part shapes would enable the development of more flexible ANN models that would allow the prediction of important deep drawing output parameters for a larger number of selected part shapes.

Acknowledgement

This paper was written with the financial support of the granting agency APVV within the project solution APVV-21-0418 and support from the Slovenian Research Agency ARIS (research core funding No. P2-0248).

References:

- [1] Spišák, E., Majerníková, J. (2014). A study of thickness change of spherical cup made from TRIP steel after hydraulic bulge test, *Key Engineering Materials*, Vol. 635, 157-160, doi: [10.4028/www.scientific.net/KEM.635.157](https://doi.org/10.4028/www.scientific.net/KEM.635.157).
- [2] Lai, M., Brun, R. (2007). Latest developments in sheet metal forming technology and materials for automotive application: The use of ultra-high strength steels at fiat to reach weight reduction at sustainable costs, *Key Engineering Materials*, Vol. 344, 1-8, doi: [10.4028/www.scientific.net/KEM.344.1](https://doi.org/10.4028/www.scientific.net/KEM.344.1).
- [3] Bright, G.W., Kennedy, J.I., Robinson, F., Evans, M., Whittaker, M.T., Sullivan, J., Gao, Y. (2011). Variability in the mechanical properties and processing conditions of a High Strength Low Alloy steel, *Procedia Engineering*, Vol. 10, 106-111, doi: [10.1016/j.proeng.2011.04.020](https://doi.org/10.1016/j.proeng.2011.04.020).
- [4] Thomas, S.K., Ali, A., AlArjani, A., Attia, E.-A. (2022). Simulation based performance improvement: A case study on automotive industries, *International Journal of Simulation Modelling*, Vol. 21, No. 3, 405-416, doi: [10.2507/IJSIMM21-3-606](https://doi.org/10.2507/IJSIMM21-3-606).
- [5] Visagan, A., Ganesh, P. (2022). Parametric optimization of two point incremental forming using GRA and TOPSIS, *International Journal of Simulation Modelling*, Vol. 21, No. 4, 615-626, doi: [10.2507/IJSIMM21-4-622](https://doi.org/10.2507/IJSIMM21-4-622).
- [6] Spišák, E., Majerníková, J. (2013). Analysis of variance of mechanical properties of sheets as the input parameters for simulation of processes, *Acta Metallurgica Slovaca*, Vol. 18, No. 2-3, 109-116.
- [7] Dykeman, J., Hoydick, D., Link, T., Mitsuji, H. (2009). Material property and formability characterization of various types of high strength dual phase steel, *SAE Technical Paper 2009-01-0794*, SAE International, doi: [10.4271/2009-01-0794](https://doi.org/10.4271/2009-01-0794).
- [8] Takahashi, M. (2003). Development of high strength steels for automobiles, *Nippon Steel Technical Report*, No. 88, 2-7, from <https://www.nipponsteel.com/en/tech/report/nsc/pdf/n8802.pdf>, accessed February 7, 2024.
- [9] Bleck, W., Brühl, S., Gerdemann, F.L.H., Prahl, U. (2007). Gefüge-Engineering bei kaltumformbaren Stählen, *Umformtechnik: Stahl und NE-Werkstoffe; der Zukunft Form geben; Tagungsband / 22. ASK, Aachener Stahlkolloquium*, 08/09. März 2007, Eurogress Aachen. Institut für Bildsame Formgebung; Institut für Eisenhüttenkunde, RWTH, Rheinisch-Westfälische Technische Hochschule Aachen, 267-280.
- [10] Heinemann, G. (2004). *Virtuelle Bestimmung des Verfestigungsverhaltens von Bändern und Blechen durch verformungsinduzierte Martensitbildung bei metastabilen rostfreien austenitischen Stählen*, Dissertation, Eidgenössische Technische Hochschule Zürich.
- [11] Behrens, B.-A., Hübner, S., Voges-Schwieger, K., Weilandt, K. (2007). Verformungsinduzierte Martensitevolution zur lokalen Festigkeitssteigerung, *UTF Science II*, 1-4.
- [12] Papaefthymiou, S. (2005). Failure mechanisms of multiphase steels, Dissertation, RWTH Aachen, Aachen, Shaker Verlag.
- [13] Stefanovska, E., Pepelnjak, T. (2022). Development of a flexible tooling system for sheet metal bending, *Advances in Production Engineering & Management*, Vol. 17, No. 3, 311-325, doi: [10.14743/apem2022.3.438](https://doi.org/10.14743/apem2022.3.438).
- [14] Spišák, E., Greškovič, F., Maňková, I., Brezinová, J., Král, J., Slota, J., Draganovská, D., Viňas, J., Kaščák, L. (2011). *Strojárske technológie*, Strojnícka fakulta TU v Košiciach, Košice, Slovakia.
- [15] Dwivedi, R., Agnihotri, G. (2017). Study of deep drawing process parameters, *Materials Today: Proceedings*. Vol. 4, No. 2, Part A, 820-826, doi: [10.1016/j.matpr.2017.01091](https://doi.org/10.1016/j.matpr.2017.01091).
- [16] Semiati, S.L. (2006). Introduction to sheet-forming processes, In: Semiati (ed.), *Metalworking: Sheet Forming* Vol. 14B, ASM International, Almere, Netherlands, 319-333, doi: [10.31399/asm.hb.v14b.9781627081863](https://doi.org/10.31399/asm.hb.v14b.9781627081863).
- [17] Spišák, E., Slota, J., Majerníková, J., Kaščák, L., Malega, P. (2012). Inhomogeneous plastic deformation of tinplates under uniaxial stress state, *Chemické listy*. Vol. 106, 537-540.
- [18] Joshi, A.R., Kothari, K.D., Jhala, R.L. (2013). Effects of different parameters on deep drawing process: Review, *International Journal of Engineering Research & Technolog.* Vol. 2, No. 3.

- [19] Benke, M., Schweitzer, B., Hlavacs, A., Mertinger, V. (2020). Prediction of earing of cross-rolled al sheets from {h00} pole figures, *Metals*, Vol. 10, No. 2, Article No. 192, doi: [10.3390/met10020192](https://doi.org/10.3390/met10020192).
- [20] Hlavacs, A., Szucs, M., Mertinger, V., Benke, M. (2021). Prediction of earing of hot-rolled al sheets from pole figures, *Metals*, Vol. 11, No. 1, Article No. 99, doi: [10.3390/met11010099](https://doi.org/10.3390/met11010099).
- [21] Dong, W., Wang, Q., Wang, X., Wang, B. (2018). Stress analysis of cylindrical parts during deep drawing based on Dynaform, *IOP Conference Series Materials Science Engineering*, Vol. 423, Article No. 012166, doi: [10.1088/1757-899X/423/1/012166](https://doi.org/10.1088/1757-899X/423/1/012166).
- [22] Colgan, M., Monaghan, J. (2003). Deep drawing process: Analysis and experiment, *Journal of Materials Processing Technology*, Vol. 132, No. 1-3, 35-41, doi: [10.1016/S0924-0136\(02\)00253-4](https://doi.org/10.1016/S0924-0136(02)00253-4).
- [23] Seth, M., Vohnout, V.J., Daehn, G.S. (2005). Formability of steel sheet in high velocity impact, *Journal of Materials Processing Technology*, Vol. 168, No. 3, 390-400, doi: [10.1016/j.jmatprotec.2004.08.032](https://doi.org/10.1016/j.jmatprotec.2004.08.032).
- [24] Huang, Y.-M., Tsai, Y.-W., Li, C.-L. (2008). Analysis of forming limits in metal forming processes, *Journal of Materials Processing Technology*, Vol. 201, No. 1-3, 385-389, doi: [10.1016/j.jmatprotec.2007.11.279](https://doi.org/10.1016/j.jmatprotec.2007.11.279).
- [25] Chalal, H., Abed-Meraim, F. (2017). Determination of forming limit diagrams based on ductile damage models and necking criteria, *Latin American Journal of Solids and Structures*, Vol. 14, No. 10, 1872-1892, doi: [10.1590/1679-78253481](https://doi.org/10.1590/1679-78253481).
- [26] Gusel, L., Boskovic, V., Domitner, J., Ficko, M., Brezocnik, M. (2018). Genetic programming method for modelling of cup height in deep drawing process, *Advances in Production Engineering & Management*, Vol. 13, No. 3, 358-365, doi: [10.14743/apem2018.3.296](https://doi.org/10.14743/apem2018.3.296).
- [27] Sevšek, L., Baressi Šegota, S., Car, Z., Pepelnjak, T. (2023). Determining the influence and correlation for parameters of flexible forming using the random forest method, *Applied Soft Computing*, Vol. 144, Article No. 110497, doi: [10.1016/j.asoc.2023.110497](https://doi.org/10.1016/j.asoc.2023.110497).
- [28] Cwiekala, T., Brosius, A., Tekkaya, A.E. (2011). Accurate deep drawing simulation by combining analytical approaches, *International Journal of Mechanical Sciences*, Vol. 53, No. 5, 374-386, doi: [10.1016/j.ijmecsci.2011.02.007](https://doi.org/10.1016/j.ijmecsci.2011.02.007).
- [29] Milutinovic, M., Lendjel, R., Baloš, S., Labus Zlatanovic, D., Sevšek, L., Pepelnjak, T. (2021). Characterisation of geometrical and physical properties of a stainless steel denture framework manufactured by single-point incremental forming, *Journal of Materials Research and Technology*, Vol. 10, 605-623, doi: [10.1016/j.jmrt.2020.12.014](https://doi.org/10.1016/j.jmrt.2020.12.014).
- [30] Vrh, M., Halilović, M., Starman, B., Štok, B., Comsa, D.-S., Banabic, D. (2014). Capability of the BBC2008 yield criterion in predicting the earing profile in cup deep drawing simulations, *European Journal of Mechanics, A/Solids*, Vol. 45, 59-74, doi: [10.1016/j.euromechsol.2013.11.013](https://doi.org/10.1016/j.euromechsol.2013.11.013).
- [31] Bandyopadhyay, K., Panda, S.K., Saha, P., Padmanabham, G. (2015). Limiting drawing ratio and deep drawing behavior of dual phase steel tailor welded blanks: FE simulation and experimental validation, *Journal of Materials Processing Technology*, Vol. 217, No. 48-64, doi: [10.1016/j.jmatprotec.2014.10.022](https://doi.org/10.1016/j.jmatprotec.2014.10.022).
- [32] Dwivedi, R., Agnihotri, G. (2015). Numerical simulation of aluminum and brass material cups in deep drawing process, *Materials Today: Proceedings*, Vol. 2, No. 4-5, 1942-1950, doi: [10.1016/j.matpr.2015.07.159](https://doi.org/10.1016/j.matpr.2015.07.159).
- [33] Walde, T., Riedel, H. (2007). Simulation of earing during deep drawing of magnesium alloy AZ31, *Acta Materialia*, Vol. 55, No. 3, 867-874, doi: [10.1016/j.actamat.2006.09.007](https://doi.org/10.1016/j.actamat.2006.09.007).
- [34] Engler, O., Aretz, H. (2021). A virtual materials testing approach to calibrate anisotropic yield functions for the simulation of earing during deep drawing of aluminium alloy sheet, *Materials Science and Engineering: A*, Vol. 818, Article No. 141389, doi: [10.1016/j.msea.2021.141389](https://doi.org/10.1016/j.msea.2021.141389).
- [35] Luyen, T.-T., Tong, V.-C., Nguyen, D.-T. (2022). A simulation and experimental study on the deep drawing process of SPCC sheet using the graphical method, *Alexandria Engineering Journal*, Vol. 61, No. 3, 2472-2483, doi: [10.1016/j.aej.2021.07.009](https://doi.org/10.1016/j.aej.2021.07.009).
- [36] Jayahari, L., Balunaik, B., Gupta, A.K., Singh, S.K. (2015). Finite element simulation studies of AISI 304 for deep drawing at various temperatures, *Materials Today: Proceedings*, Vol. 2, No. 4-5, 1978-1986, doi: [10.1016/j.matpr.2015.07.166](https://doi.org/10.1016/j.matpr.2015.07.166).
- [37] Gondo, S., Arai, H. (2022). Data-driven metal spinning using neural network for obtaining desired dimensions of formed cup, *CIRP Annals*, Vol. 71, No. 1, 229-232, doi: [10.1016/j.cirp.2022.04.044](https://doi.org/10.1016/j.cirp.2022.04.044).
- [38] Tian, S., Zhang, Z., Xie, X., Yu, C. (2022). A new approach for quality prediction and control of multistage production and manufacturing process based on Big Data analysis and Neural Networks, *Advances in Production Engineering & Management*, Vol. 17, No. 3, 326-338, doi: [10.14743/apem2022.3.439](https://doi.org/10.14743/apem2022.3.439).
- [39] Kazan, R., Firat, M., Tiryaki, A.E. (2009). Prediction of springback in wipe-bending process of sheet metal using neural network, *Materials & Design*, Vol. 30, No. 2, 418-423, doi: [10.1016/j.matdes.2008.05.033](https://doi.org/10.1016/j.matdes.2008.05.033).
- [40] Afshari, D., Ghaffari, A., Barsum, Z. (2022). Optimization in the resistant spot-welding process of AZ61 magnesium alloy, *Strojnicki Vestnik/Journal of Mechanical Engineering*, Vol. 68, No. 7-8, 485-492, doi: [10.5545/sv-jme.2022.174](https://doi.org/10.5545/sv-jme.2022.174).
- [41] Berus, L., Klancnik, S., Brezocnik, M., Ficko, M. (2019) Classifying parkinson's disease based on acoustic measures using artificial neural networks, *Sensors*, Vol. 19, No. 1, Article No. 16, doi: [10.3390/s19010016](https://doi.org/10.3390/s19010016).
- [42] Sivasankaran, S., Narayanasamy, R., Jeyapaul, R., Loganathan, C. (2009). Modelling of wrinkling in deep drawing of different grades of annealed commercially pure aluminium sheets when drawn through a conical die using artificial neural network, *Materials & Design*, Vol. 30, No. 8, 3193-3205, doi: [10.1016/j.matdes.2009.01.020](https://doi.org/10.1016/j.matdes.2009.01.020).
- [43] Spaić, O., Krivokapić, Z., Kramar, D. (2020). Development of family of artificial neural networks for the prediction of cutting tool condition, *Advances in Production Engineering & Management*, Vol. 15, No. 2, 164-178, doi: [10.14743/APEM2020.2.356](https://doi.org/10.14743/APEM2020.2.356).

- [44] Nguyen, T.P.Q., Yang, C.L., Le, M.D., Nguyen, T.T., Luu, M.T. (2023). Enhancing automated defect detection through sequential clustering and classification: An industrial case study using the Sine-Cosine Algorithm, Possibilistic Fuzzy c-means, and Artificial Neural Network, *Advances in Production Engineering & Management*, Vol. 18, No. 2, 237-249, [doi: 10.14743/apem2023.2.470](https://doi.org/10.14743/apem2023.2.470).
- [45] Xia, J.S., Khaje Khabaz, M., Patra, I., Khalid, I., Alvarez, J.R.N., Rahmanian, A., Eftekhari, S.A., Toghraie, D. (2023). Using feed-forward perceptron Artificial Neural Network (ANN) model to determine the rolling force, power and slip of the tandem cold rolling, *ISA Transactions*, Vol. 132, 353-363, [doi: 10.1016/j.isatra.2022.06.009](https://doi.org/10.1016/j.isatra.2022.06.009).
- [46] Czinege, I., Harangozó, D. (2024). Application of artificial neural networks for characterisation of formability properties of sheet metals, *International Journal of Lightweight Materials and Manufacture*, Vol. 7, No. 1, 37-44, [doi: 10.1016/j.ijlmm.2023.08.003](https://doi.org/10.1016/j.ijlmm.2023.08.003).
- [47] Savkovic, B., Kovac, P., Rodic, D., Strbac, B., Klančnik, S. (2020). Comparison of artificial neural network, fuzzy logic and genetic algorithm for cutting temperature and surface roughness prediction during the face milling process, *Advances in Production Engineering & Management*, Vol. 15, No. 2, 137-150, [doi: 10.14743/APEM2020.2.354](https://doi.org/10.14743/APEM2020.2.354).
- [48] El Mehtedi, M., Forcellese, A., Greco, L., Pieralisi, M., Simoncini, M. (2019). Flow curve prediction of ZAM100 magnesium alloy sheets using artificial neural network-based models, *Procedia CIRP*, Vol. 79, No. 661-666, [doi: 10.1016/j.procir.2019.02.050](https://doi.org/10.1016/j.procir.2019.02.050).
- [49] Babu, K.V., Ganesh Narayanan, R., Saravana Kumar, G. (2010). An expert system for predicting the deep drawing behavior of tailor welded blanks, *Expert Systems with Applications*, Vol. 37, No. 12, 7802-7812, [doi: 10.1016/j.eswa.2010.04.059](https://doi.org/10.1016/j.eswa.2010.04.059).
- [50] Manoochchri, M., Kolahan, F. (2014). Integration of artificial neural network and simulated annealing algorithm to optimize deep drawing process, *The International Journal of Advanced Manufacturing Technology*, Vol. 73, 241-249, [doi: 10.1007/s00170-014-5788-5](https://doi.org/10.1007/s00170-014-5788-5).

The investigation of production variance in a module-based assembly system: A Markovian Arrival Process approach

Zhong, C.W.^a, Zhang, H.^a, Zhang, Z.W.^b, Wu, Z.Q.^c, Lu, M.L.^{a,*}

^aSchool of Economics and Management, Beijing Institute of Petrochemical Technology, Beijing, P.R. China

^bHenan Yuzhan Precision Technology Co.,Ltd, Zhengzhou, P.R. China

^cDepartment of Industrial Engineering, Tsinghua University, Beijing, P.R. China

ABSTRACT

This paper provides an in-depth study of the assembly production variance problem through the largest supplier of General Motors China. We focus on the production variance problem in an unreliable assembly (MOBA) system with a finite inter-station buffer, focusing on two of the central issues, namely the output variance as well as the delivery schedule variance. We model every subsystem's departure procedure in the MOBA system using the Markov Arrival Process (MAP) approach. Through the approximate use of MAP, we successfully shorten the time needed to calculate the output variance as well as delivery schedule variance of a large-scale MOBA system, which improves the efficiency of the system while ensuring that it meets the customer's needs. The relationship between production variance and system parameters is also studied, which is of substantial significance for optimizing the productivity of MOBA systems and improving customer satisfaction.

ARTICLE INFO

Keywords:

Manufacturing;
Assembly;
Production variance;
Markovian arrival process (MAP);
Module-based assembly (MOBA) system

*Corresponding author:

lumilin@bipt.edu.cn
(Lu, M.L.)

Article history:

Received 18 January 2024

Revised 28 January 2024

Accepted 7 March 2024



Content from this work may be used under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Background

The aim of this paper is to address the variance problem of an unreliable module-based assembly system (MOBA). The main assembly line, which includes M assembly stations and M sub-fabrication lines, is the component that makes up the MOBA system, each supplying modules to one assembly station with a specific fill rate. One assembly operation is performed at each assembly station, which involves attaching a module from a sub-line from a higher stream station to a subassembly of the stream level.

In China, most of suppliers for the car industry share this production mode. To illustrate it, a study was made by us for a famous bumper factory which is the biggest supplier for GM in Yan-tai City.

The factory uses a typical MOBA system after injection molding and painting. Each assembly line has 5 to 10 assembly stations, as well as sub-fabrication lines according to the parts of the bumpers. In today's business landscape, clients have increasingly high expectations for shorter delivery times. Given the fast-paced nature of the industry, we feel it is crucial to build models

that analyze the variability in output or delivery schedule of MOBA systems. This will assist factories in effectively managing fluctuations in production.

Some interrupts such as set-up and module-changing were not included in this study since these conditions are totally under control in manufacturing process with a fixed time and procedure.

2. Introduction

The field of analytical modeling of manufacturing systems has been the subject of significant research and interest since the early 1960s. Dallery and Gershwin's [1], and Papadopoulos and Heavey's [2] contributions to this field are notable. Most of those research examine manufacturing systems in a condition of equilibrium, which provide certain performance metrics such as the mean production rate and the mean buffer levels.

The existing body of literature mostly concentrates on the analysis of production variation in manufacturing systems that employ serial production lines. The variability in output of manufacturing lines equipped with buffers between stations has been examined by Miltenberg [3], Martinčević and Kozina [4], Carrascosa [5], and Tan [6, 7]. Miltenberg [3] provided a technique for calculating the asymptotic variance of the output per unit time. This approach utilizes the outcomes derived for the asymptotic averages as well as variance of the overall duration of stay in Markov chains. Martinčević and Kozina [4] proposed a technique for calculating the variability of the output from a single machine during a certain time frame. This approach relies on calculating the derivative functions for the likelihood of generating n components at a specific moment, and subsequently solving these equations by including certain boundary equations. Gershwin further suggested a decomposition technique for calculating the output variation in extended manufacturing lines. Carrascosa [5] expanded upon Gershwin's approach as well as provided extensive numerical and simulation findings that investigate the impact of system factors on the output variance. Tan [6, 7] determined the rate of asymptotic variation for production lines with restricted buffers by using the averages as well as variances of Markov incentives systems. In terms of the amount of operations, his technique outperforms Miltenberg's by a factor of a thousand. In his work on serial production lines without inter-station buffers, Tan [8, 9] discussed the output variance and found closed-form equations for the asymptotic variance of the output. Tan [10], and Behmanesh and Rahimi [11] investigated the delivery schedule variation of serial manufacturing lines. For a solitary workstation characterized by foreseeable processing times and unforeseeable downtimes, a predetermined lot size was determined by Kim and Alden [12] by analytically approximating the density function as well as the variance of the period. The number of commodities produced in time t , denoted as $N(t)$, and its asymptotic normalcy, was used by Tan [10] to estimate the delivery schedule variance. Tan [13] addressed the output variance of series-parallel production systems with no inter-station buffers. This study is an extension of Tan [8]. El Abbadi *et al.* [14] and Marinas *et al.* [15] studied production variance in serial manufacturing lines with unreliable machines in a Bernoulli reliability case.

The current research on production variation might not be enough for real-world applications, according to these evaluations. In addition, with the exception of Tan's work on series-parallel systems [13], all of the systems discussed in the aforementioned literatures are serial systems. The unreliability of a module-based assembly (MOBA) system is the subject of this paper's discussion of production variation. MOBA systems have become widely adopted in the modern manufacturing sector to achieve efficient production and adaptability. They combine agile manufacturing principles to optimize both product-focused and process-focused production modes. Several concerns must be considered while designing and controlling such systems, including variations in production, along steady-state throughput, and average buffer level. In this study, we introduced a novel and efficient technique known as the MAP approximation-based compression method.

The MAP is a helpful framework of math for modeling point processes having irregular, non-Markovian dynamics or a unique pattern. Notable instances are Neuts [16, 17], Lucantoni [18, 19], Montoro-Cazorla and Pérez-Ocón [20], Visagan and Ganesh [21], and Fan [22].

Because of the intricate nature and interconnections of MOBA systems, it is quite challenging to assess the variability in production of these systems. The MAP estimation method in this study offers a foundation for a viable strategy to assess production variation in complicated and unreliable assembly systems. This work examines two important issues related to production variance: (i) the issue of output variance; and (ii) the issue of delivery schedule variance.

The remaining portion of this paper is organized in the following manner. Section 2 provides a description of the MOBA system. Section 3 discusses the estimation using the MAP method used to estimate the production variance. In Section 4, two deviations for the production are estimated. Numerical examples are examined in Section 5. Lastly, we provide some last comments in Section 6.

3. Model description

This section provides a description of a MOBA system, as seen in Fig. 1.

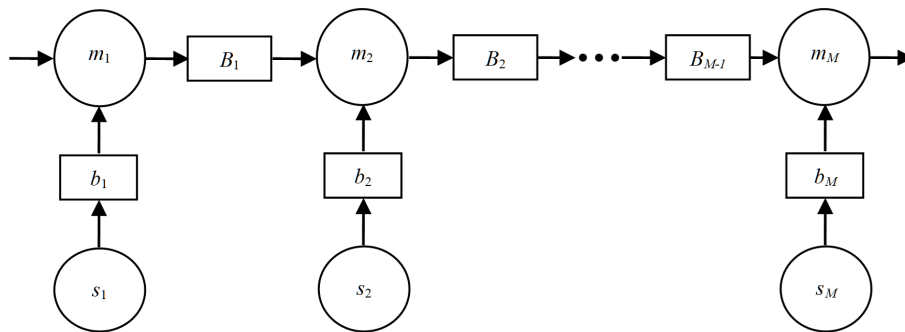


Fig. 1 A MOBA system

The MOBA system has a primary production lane equipped with M stations for assembly and M lines for sub-fabrication. Every assembly station carries out a single assembly operation by extracting a module from its sub-fabrication line and joining it with a subassembly obtained from the upstream station. The primary assembly line has $M-1$ inter-station buffers. The following assumptions define the assembly stations and the inter-station buffers.

- The assembly time it takes to assemble station m_i is distributed exponentially at a rate of μ_i , where $i = 1, \dots, M$. The assembly times of the assembly stations are not dependent on each other.
- When it comes to assembly activities, every single assembly station is prone to erratic failures. The duration till failure and the duration until repair for the station m_i follow exponential distributions with rates α_i and β_i , respectively, for $i = 1, \dots, M$.
- The inter-station buffer B_i has a capacity to store N_i components, where N_i is a positive integer more than or equal to 1 and less than infinity, for $i = 1, \dots, M-1$.
- If assembly station m_i completes work on a component while inter-station buffer B_{i-1} is empty, station m_i will not receive any new parts until m_i finishes a part, for $i = 2, \dots, M$. The first station is always well-supplied, as there is an endless amount of basic assembly available.
- If assembly station m_i completes work on a part while inter-station buffer B_i is at maximum capacity, station m_i is unable to proceed until an empty buffer space becomes vacant at B_i for $i = 1, \dots, M-1$. The final station remains unimpeded, as there is an endless capacity for transporting completed goods.

In an actual MOBA system, every single subsidiary manufacturing line may have several stations for either machining or assembly. To keep things simple, we suppose it is a subsidiary manufacturing line is made up of a solitary processing machine as well as a supply buffer designed to store the outputs generated by the processing machine. The processing machines and supply buffers are defined by the following assumptions.

- The sub-fabrication lines are managed separately to achieve rate of module filling θ_i at assembly station m_i , where $0 \leq \theta_i \leq 1$ for $i = 1, \dots, M$.
- When there is a shortage of stock in the buffer of supplies b_i , the time it takes for the next module to become available follows an exponential distribution with a rate of δ_i , where i ranges from 1 to M . In other words, there is an infinite supply of raw materials and components for the sub-fabrication line, as well as the machine's processing time s_i follows an exponential pattern.

4. The MAP approximation

Within this part, we initially present the effective assembly times. Based on that, the procedure by which every subsystem leaves the MOBA system can be modeled as a MAP. Then we describe Diamond and Alfa's [23] discussion on MAP approximation, which is used in our study.

Random failures and maintenance of assembly stations, and also fill rates and delays at sub-fabrication lines, have an impact on assembly operations in a MOBA system. The duration of station's effective assembly m_i is defined as the duration between the availability of an item of work or a station m_i and the commencement of an operation at station m_i , provided that a module may be accessed via the supply buffer. It includes the time needed to complete a procedure that involves the assembling of the workpiece.

The efficient assembly duration of the station m_i is proved by Liu *et al.* [24] to follow a PH-distribution with the representation (τ_i, T_i) , which is given by

$$\tau_i = (1 - \theta_i, \theta_i, 0, 0), T_i = \begin{pmatrix} T_i' - \alpha_i I_2 & \alpha_i I_2 \\ \beta_i I_2 & -\beta_i I_2 \end{pmatrix} \tag{1}$$

where I_2 is an identity matrix of order 2, $T_i' = \begin{pmatrix} -\delta_i & \delta_i \\ 0 & -\mu_i \end{pmatrix}$, for $i = 1, \dots, M$.

4.1 The MAP structure of a MOBA system

A MAP, or Markovian Arrival Process, is a type of Markov process that may be represented by a map of size m . The transitions in this process are categorized based on whether they result in an arrival or not. The symbol m refers to the order of the MAP. The associated rates are divided into two matrices E and F of size m . Matrix F has only nonnegative components, whereas matrix E comprises negative diagonal members and nonnegative off-diagonal elements. The infinitesimal generator Q , which is the sum of matrices E and F , is indivisible. Therefore, the Markov chain Q exhibits positive recurrence. The matrix that determines the likelihood of the system's phase transitioning the information on the time between two consecutive arrival epochs is given as $P = -E^{-1}F$. Let η be the stationary probability vector of P . It is clear that $\eta P = \eta$.

The MOBA system may be studied by examining the MAP of each subsystem's departure procedure, which is based on the PH-distribution of the effective assembly time of assembly stations. The information is explained as follows.

Subsystem 1 shown in Fig. 2 is a PH/PH/1/ N_1 queue. Its leaving procedure may be represented as a MAP₁ of order $(N_1 + 1) \times 16$ with the matrix identifier (E_1, F_1) , determined by

$$E_1 = \begin{bmatrix} T_1 \otimes I_4 & T_1^0 \tau_1 \otimes I_4 & \square & \square & \square \\ \square & T_1 \oplus T_2 & T_1^0 \tau_1 \otimes I_4 & \square & \square \\ \square & \square & \ddots & \ddots & \square \\ \square & \square & \square & T_1 \oplus T_2 & T_1^0 \tau_1 \otimes I_4 \\ \square & \square & \square & \square & I_4 \otimes T_2 \end{bmatrix} \tag{2}$$

$$F_1 = \begin{bmatrix} 0 & \square & \square & \square \\ I_4 \otimes T_2^0 \tau_2 & 0 & \square & \square \\ \square & \ddots & \ddots & \square \\ \square & \square & I_4 \otimes T_2^0 \tau_2 & 0 \end{bmatrix} \tag{3}$$

where I_4 is identity matrix of order 4, $T_i^0 = -T_i e$ for $i = 1, \dots, M$ and e is the column vector of ones, the operator \otimes denotes the Kronecker product, see Diamond *et al.* [23], the operator \oplus is defined as $A \oplus B = A \otimes I + I \otimes B$.

A MAP order reduction to 2 is required to make this paper's technique work for a big MOBA system. As a result, a MAP_1^* of order 2 with the matrix identifier should be used to approximate the MAP_1 of higher order; Section 4.2 will provide further information on this.

As depicted in Fig. 2, subsystem 2 regards the MAP_1^* as input. It may be referred to as a MAP/PH/1/ N_2 queue, where the process of departure is a newly constructed MAP_2 that includes the matrix identifier (E_2^*, F_2^*) , as follows:

$$E_2 = \begin{bmatrix} E_1^* \otimes I_4 & F_1^* \otimes I_4 & \square & \square & \square \\ \square & E_1^* \oplus T_3 & F_1^* \otimes I_4 & \square & \square \\ \square & \square & \ddots & \ddots & \square \\ \square & \square & \square & E_1^* \oplus T_3 & F_1^* \otimes I_4 \\ \square & \square & \square & \square & I_2 \otimes T_3 \end{bmatrix} \tag{4}$$

$$F_2 = \begin{bmatrix} 0 & \square & \square & \square \\ I_2 \otimes T_3^0 \tau_3 & 0 & \square & \square \\ \square & \ddots & \ddots & \square \\ \square & \square & I_2 \otimes T_3^0 \tau_3 & 0 \end{bmatrix}, \tag{5}$$

where I_2 is identity matrix of order 2.

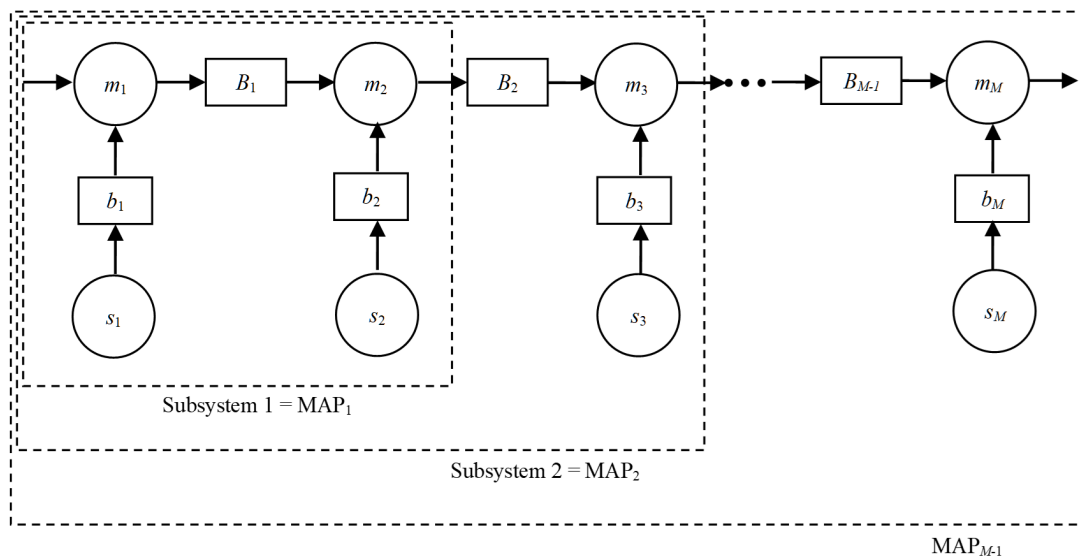


Fig. 2 The MAP approximation of a MOBA system

The MAP_2 's order is 8 times the product of N_2 and 1. The estimation of MAP_2 involves the utilization of a second-order MAP_2^* in conjunction with the matrix descriptor (E_2^*, F_2^*) , following an analogous discourse to that over the MAP_1 . Continuing in this manner, it is evident that the ultimate system may clearly be described by a MAP_{M-1} using the matrix identifier (E_{M-1}, F_{M-1}) , as follows:

$$E_{M-1} = \begin{bmatrix} E_{M-2}^* \otimes I_4 & F_{M-2}^* \otimes I_4 & \square & \square & \square \\ \square & E_{M-2}^* \oplus T_M & F_{M-2}^* \otimes I_4 & \square & \square \\ \square & \square & \ddots & \ddots & \square \\ \square & \square & \square & E_{M-2}^* \oplus T_M & F_{M-2}^* \otimes I_4 \\ \square & \square & \square & \square & I_2 \otimes T_M \end{bmatrix}, \tag{6}$$

and

$$F_{M-1} = \begin{bmatrix} 0 & \square & \square & \square \\ I_2 \otimes T_M^0 \tau_M & 0 & \square & \square \\ \square & \vdots & \vdots & \square \\ \square & \square & I_2 \otimes T_M^0 \tau_M & 0 \end{bmatrix} \quad (7)$$

The order of MAP_{M-1} is $8 \times (N_{M-1} + 1)$.

4.2 Higher order MAP approximation using MAP of order 2

In a MOBA system with several stations and substantial buffer capacity between them, the resultant MAP grows exponentially in order. It is required to approximate all MAPs of higher order, such $MAP_1, MAP_2, \dots, MAP_{M-2}$, to those of order 2, as mentioned in Section 4.1. Keep in mind that Diamond and Alfa [23] offered this approximation model.

5. Variance of output and delivery schedule

Two major issues with production variance are discussed here. The output variance as well as the delivery schedule variance are these. They may be calculated using two efficient techniques that are also provided.

Define O_{t1} as the number of arrivals in the interval $[0, t]$ for the Markovian Arrival Process (MAP) with the matrix descriptor (E, F) . According to Andersen and Nielsen [25], we have

$$Var[O_{t1}] = (\lambda^* - 2\lambda^{*2} + 2\varphi F(e\varphi - Q)^{-1}Fe)t - 2\varphi F(I - \exp\{Qt\})(e\varphi - Q)^{-2}Fe \quad (8)$$

the vector φ represents the stationary probabilities of the irreducible infinitesimal generator $Q = E + F$, $\lambda^* = \varphi Fe$ is the stationary arrival rate of the MAP.

The stationary sequence of the arrival times between events is denoted by X_n , and $D_{n1} = \sum_{i=1}^n X_i$. Utilizing Andersen and Nielsen [25], we get

$$Var[D_{n1}] = \left(\frac{2}{\lambda^*} \varphi(I + E^{-1}F + e\varphi)^{-1}(-E)^{-1}e - \frac{1}{\lambda^{*2}}\right)n - \frac{2}{\lambda^*} \varphi(I - (-E^{-1}F)^n)(I + E^{-1}F + e\varphi)^{-2}(-E^{-1}F)(-E)^{-1}e \quad (9)$$

$Var[O_{t1}]$ and $Var[D_{n1}]$ represent the output variance as well as the delivery schedule variance. It is important to acknowledge both Eq. 10 and Eq. 11 can produce valuable outcomes when time t and part number n approach infinity.

Because $\lim_{t \rightarrow +\infty} \exp\{Qt\} = e\varphi$, it follows from Eq. 8 that

$$\lim_{t \rightarrow +\infty} (Var[O_{t1}]/t) = \lambda^* - 2\lambda^{*2} + 2\varphi F(e\varphi - Q)^{-1}Fe \quad (10)$$

and as $\lim_{t \rightarrow +\infty} (E^{-1}F)^n = e\varphi$, it follows from Eq. 9 that

$$\lim_{t \rightarrow +\infty} (Var[D_{n1}]/n) = \frac{2}{\lambda^*} \varphi(I + E^{-1}F + e\varphi)^{-1}(-E)^{-1}e - \frac{1}{\lambda^{*2}} \quad (11)$$

The asymptotic variance rate of O_{t1} , as t approaches infinity, refers to the maximum rate at which the output variance may change over time. $Var[O_{t1}]/t$, as defined by Tan [9]. Fig. 3 demonstrates the computational methods for calculating the production variance, as outlined in Eqs. 8 and 9.

Fig. 3 may be utilized to examine the production variation through the calculation of the output variance as well as delivery schedule variance utilizing the following two algorithms (see Appendix A).

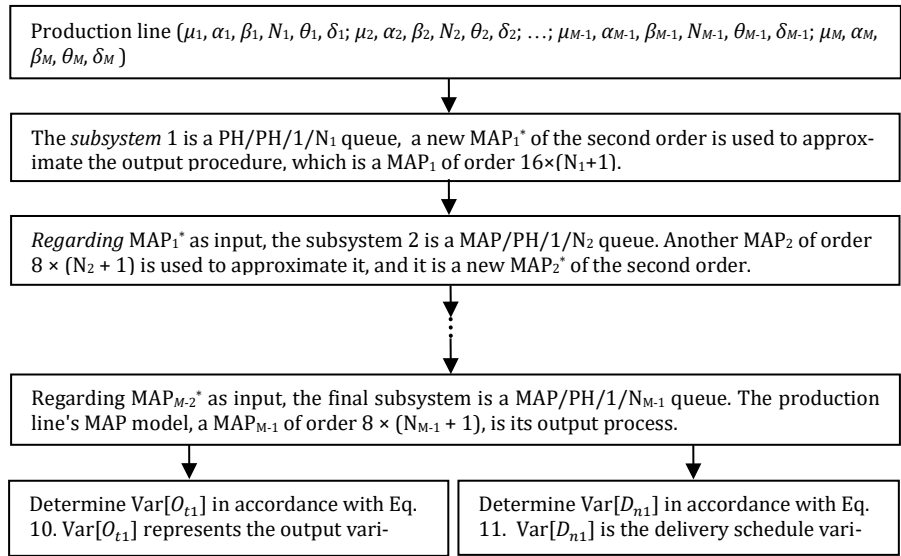


Fig. 3 Computing the output variance and delivery schedule variance

6. Numerical examples

The approach provided in the preceding section may be used to assess a broad spectrum of MOBA systems with finite inter-station buffers. In this section, an comparison between approximation and simulation results will be made to state the accuracy of this method. And then some numerical experiments show that this approach is computationally very efficient. And, some basic observations are also given to show how the production variance is influenced by system parameters. Finally, some discuss about the limit of this approach will be held.

6.1 Accuracy of the approximation

Diamond and Alfa [23] has proved the MAP approximation on higher order is acceptable and has a better performance than renewal approximation by using simulation method. In our case, the Figs. 4 and 5 depict the approximation versus simulation results. A five-station MOBA system with a certain set of system parameters is studied, where $\mu = 2.125$, $\alpha = 0.135$, $\beta = 1.452$, $\theta = 0.61$, $\delta = 2.031$, $N = 5$. All the data are obtained from the bumper factory. Curves are used to demonstrate the approximation result while the solid point as the simulation ones, as seen in the figure, the estimation yields satisfactory outcomes.

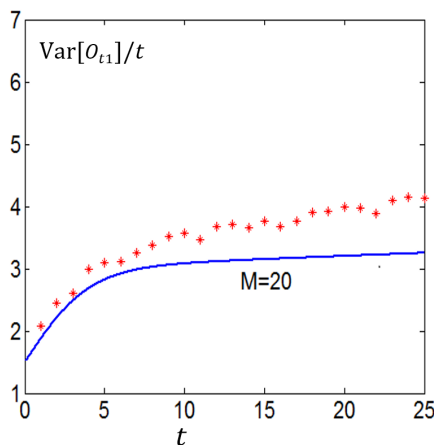


Fig. 4 Simulation and analytical approximation of output variance (Curve as approximation; point as simulation)

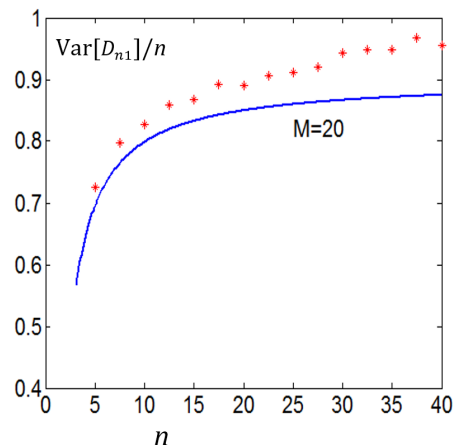


Fig. 5 Simulation and analytical approximation of delivery time (Curve as approximation; point as simulation)

6.2 Numerical examples for two/three-station MOBA system

First we investigate the relationship between the two production variances and the inter-station buffer capacity. A two-station MOBA system with two different sets of system parameters is studied, where $\mu_1 = 2.125, \alpha_1 = 0.135, \beta_1 = 1.452, \theta_1 = 0.61, \delta_1 = 2.031, \mu_2 = 3.725, \alpha_2 = 0.231, \beta_2 = 2.524, \theta_2 = 0.91, \delta_2 = 1.215$ in system 1 and $\mu_1 = 3.125, \alpha_1 = 0.135, \beta_1 = 1.452, \theta_1 = 0.8, \delta_1 = 2.725, \mu_2 = 6.725, \alpha_2 = 0.081, \beta_2 = 2.024, \theta_2 = 0.8, \delta_2 = 2.326$ in system 2. This two system are used for different bumpers. Figs. 6 and 7 show the output variance over time for various buffer capacity. It has been demonstrated that, in general, the output variance over time grows as the buffer capacity increases in both systems. Nevertheless, it is not feasible to establish a strong correlation between the delivery schedule variance as well as the capacity of the buffer. Figs. 8 and 9 show a clear correlation between the variability in delivery schedule and the size of the buffer for both systems.

Secondly, the effect of the assembly station's failure rate and repair rate on the production variance is studied. For a MOBA system with three identical assembly stations and sub-fabrication lines where $\mu = 2.726, \beta = 2.426, \theta = 0.8, \delta = 1.823$ and $N = 3$ for both inter-station buffers, Figs. 10 and 11 illustrate the relationship between the output variation for each time period as well as the delivery schedule variance for each item, with respect to the failure rate. It's shown that in this system, as the failure rate increases, the delivery schedule variance for each individual item increases, while output variance per unit time increases first and then decreases. For this three-station MOBA system, when failure rate of each assembly station is fixed as $\alpha = 0.715$, it can be seen from Figs. 12 and 13 that as the repair rate increases, the delivery schedule variance per unit part diminishes, because the rate of change in output variance increases with time, and this is dependent on the variable t , may fall somewhat in the middle.

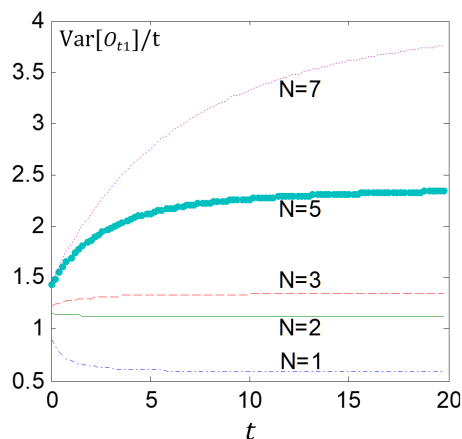


Fig. 6 Impact of buffer capacity on output variance in system 1

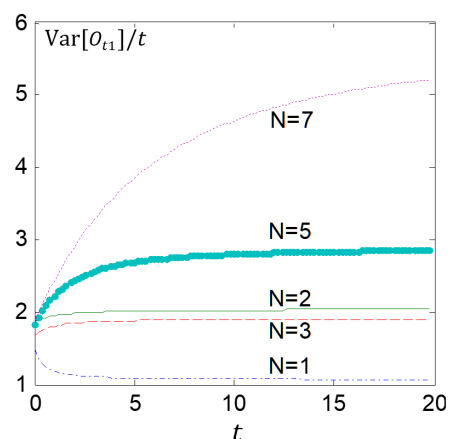


Fig. 7 Impact of buffer capacity on output variance in system 2

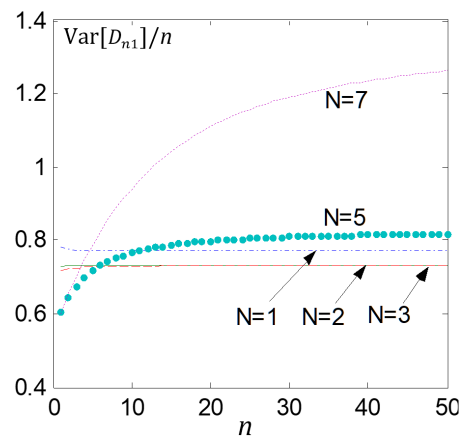


Fig. 8 Impact of buffer capacity on delivery time variance in system 1

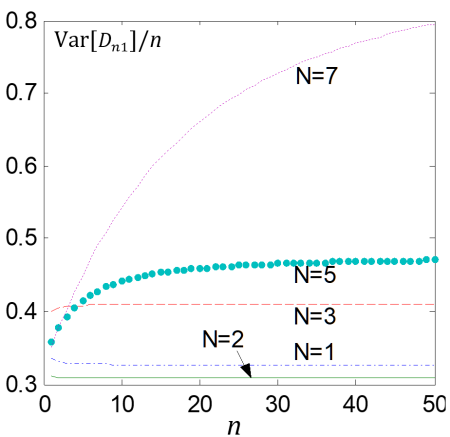


Fig. 9 Impact of buffer capacity on delivery time variance in system 2

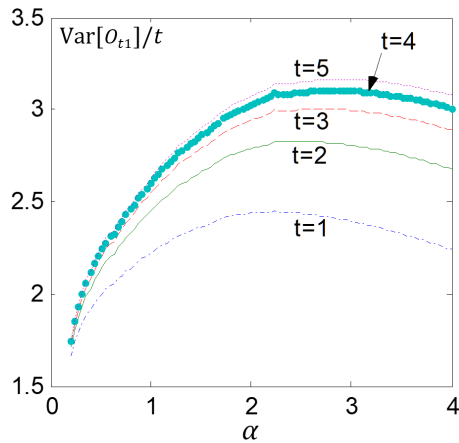


Fig. 10 Impact of failure rate on the output variance for a three-station MOBA system

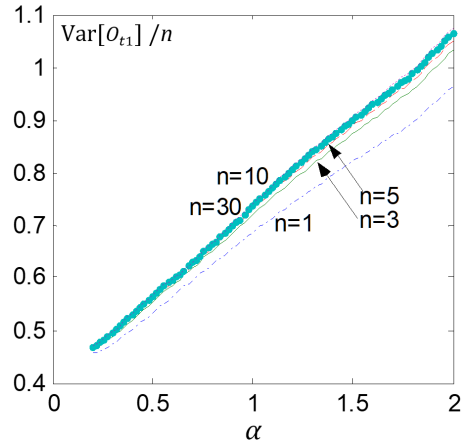


Fig. 11 Impact of failure rate on the delivery time variance for a three-station MOBA system

This observation may inspire the factory an efficient way to control the delivery schedule variance, that is, put in more effort on increasing the repair rate when it is in a lower level at the very beginning, and then focus on decreasing failure rate when the repair rate stays in a steady level. As for the output variance, failure rate may play a minor role after a certain level, but the increasing effect is quite obvious when it's in a lower level.

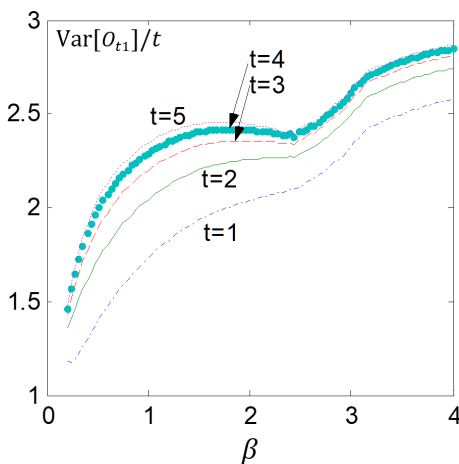


Fig. 12 Impact of repair rate on the output variance for a three-station MOBA system

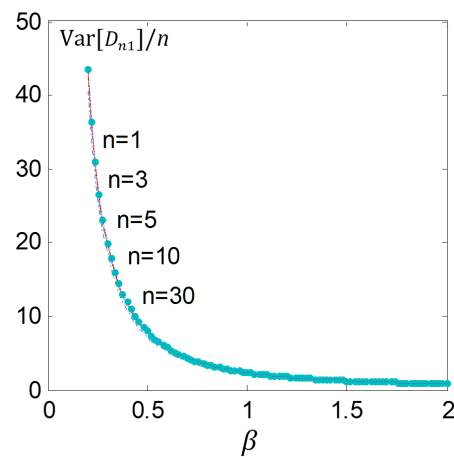


Fig. 13 Impact of repair rate on the delivery time variance for a three-station MOBA system

Thirdly, we investigate the effect of the module fill rate and the delay rate of sub-fabrication line on the production variance. A three-station MOBA system with $\mu = 2.726$, $\alpha = 0.715$, $\beta = 2.426$ for each assembly station and $N = 3$ for each inter-station buffer is studied. With each sub-fabrication line having a delay rate of $\delta = 1.823$, Figs. 14 and 15 show the output variance for each time period as well as the delivery schedule variance for each item as a function of the module fill rate for this MOBA system. It can be seen that as the module fill rate increases from 0.6 to 1, while the delivery schedule variance per unit component grows first and subsequently drops, the output variance for each time period falls initially as well as then increases. Figs. 16 and 17 depict the effect of the sub-fabrication line delay rate on the output variance per unit time as well as the delivery schedule variance for each item for this MOBA system when the module fill rate of each sub-fabrication line is $\theta = 0.9$. It's shown that as the delay rate of sub-fabrication line increases, the delivery schedule variance per unit component reduces, while the output variance for each time period decreases first, then increases and decreases again finally.

Some rules can be found from the Figs. 14-17. The trend of each variance is not monotonous as the module fill rate changes, and the effects are even opposite which means a balance point between them. By contrast, the effects of sub-fabrication line delay rate has a great influence on the delivery schedule variance at lower level which suggests the operators to increase it to a proper state as soon as possible.

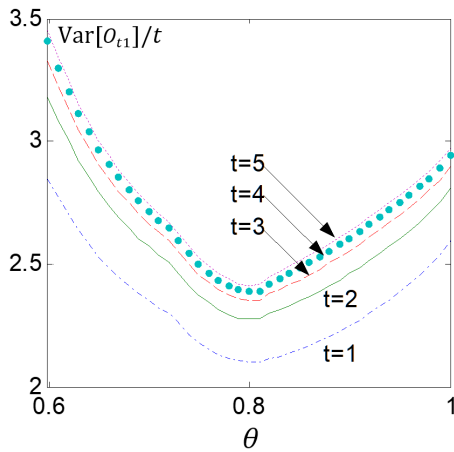


Fig. 14 Impact of module fill rate on the output variance for a three-station MOBA system

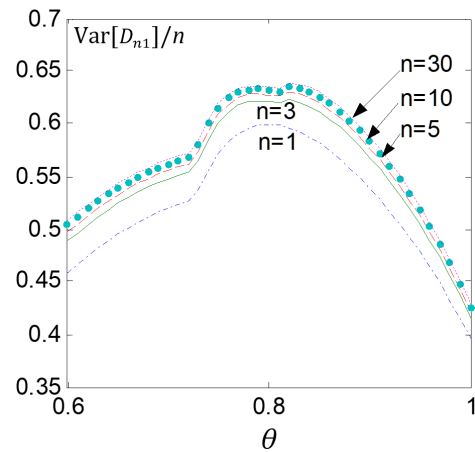


Fig. 15 Impact of module fill rate on the delivery-time variance for a three-station MOBA system

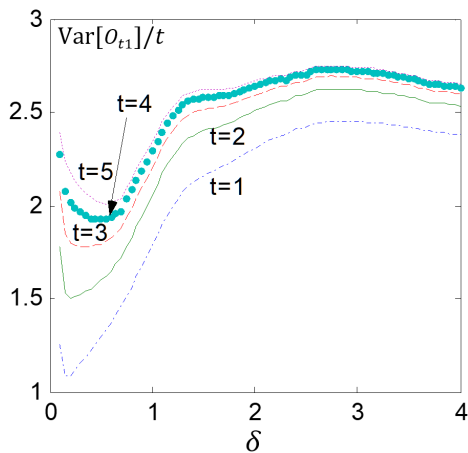


Fig. 16 Impact of sub-fabrication line delay rate on the output variance for a three-station MOBA system

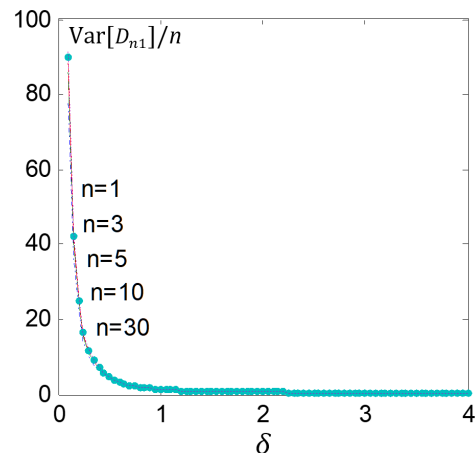


Fig. 17 Impact of sub-fabrication line delay rate on the delivery-time variance for a three-station MOBA system

6.3 Discuss about the limit of this method

The approach outlined in this research is executed via the use of software. Due to its reliance on recursion, the approach exhibits a high level of computing efficiency.

To explain the limitation of this strategy, we analyzed how the length of the line affected the production variance. We studied a MOBA system of 2, 4, 6, 10, 80 identical stations with $\mu = 2.726$, $\alpha = 0.715$, $\beta = 2.426$, $\theta = 0.8$, $\delta = 1.823$ and capacities for buffering between stations of $N = 4$. It's shown in Figs. 16 and 17, as the line length increases, output variance during each time period falls as delivery schedule variance per unit of component rises. In fact, very few assembly line can hold more than 80 stations which means this method can be used widely from a practical point of view.

Seen from the Figs. 8 and 9, the variance increases rapidly when the buffer capacity larger than 7 for both the output and the delivery schedule. Although our method can deal with a much larger buffer capacity, the result will be unacceptable for the factory.

This method for calculating production variance is quite effective. To compute the output variance as well as delivery schedule variance for big MOBA systems, it provides the findings instantly. By utilizing the MAP approximation on a per-subsystem basis, we can prevent a fast growth in the number of states as the number of stations and inter-station buffer capacity increase. Therefore, the overall computational effort for the production variance of the MOBA system is kept within a reasonable range.

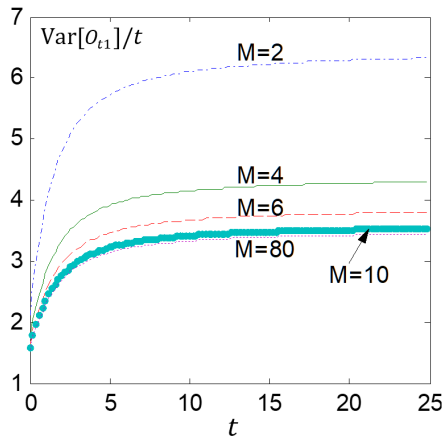


Fig. 18 Impact of line length on the output variance of MOBA system

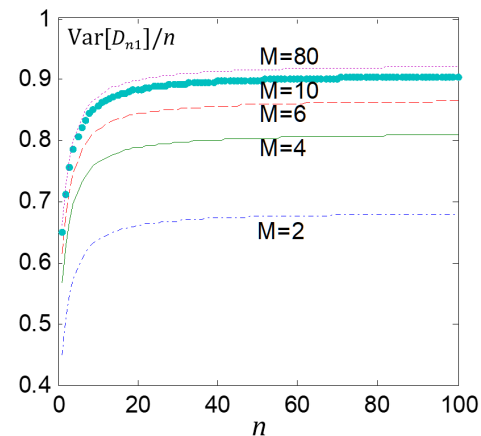


Fig. 19 Impact of line length on the delivery-time variance of MOBA system

7. Concluding remarks

The article covered the topic of production variance in MOBA systems that aren't trustworthy. For an unstable MOBA system with finite inter-station buffers, an efficient method is suggested to find the output variance as well as the delivery schedule variance using the MAP approximation. Using this method, the number of states in the manufacturing line's subsystems is drastically reduced. Therefore, the overall system's computational effort is kept within a tolerable range. When dealing with big MOBA systems, the numerical examples show that this method works. Additionally, numerical experiments are used to study the correlations between the system characteristics and the production variance. As indicated in section 1, it's known that study on production variance of assembly systems is imperfect. It is thought that this study would address a significant lack in the existing research, and the MAP approximation method discussed in this work offers a foundation for a potentially effective way to evaluating production variation in intricate and unreliable assembly systems.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The research was funded by the National Natural Science Foundation of China (71973051 and 71703009), and the Social Science Foundation of Beijing (16GLB019 and 16GLC063).

References

- [1] Dallery, Y., Gershwin, S.B. (1992). Manufacturing flow line systems: A review of models and analytical results, *Queueing Systems: Theory and Applications*, Vol. 12, No. 1, 3-94, doi: [10.1007/BF01158636](https://doi.org/10.1007/BF01158636).
- [2] Papadopoulos, H.T., Heavey, C. (1996). Queuing theory in manufacturing systems analysis and design: A classification of models for production and transfer lines, *European Journal of Operational Research*, Vol. 92, No. 1, 1-27, doi: [10.1016/0377-2217\(95\)00378-9](https://doi.org/10.1016/0377-2217(95)00378-9).
- [3] Miltenburg, G.J. (1987). Variance of the number of units produced on a transfer line with buffer inventories during a period of length T, *Naval Research Logistics*, Vol. 34, No. 6, 811-822, doi: [10.1002/1520-6750\(198712\)34:6<811::AID-NAV3220340606>3.0.CO;2-Z](https://doi.org/10.1002/1520-6750(198712)34:6<811::AID-NAV3220340606>3.0.CO;2-Z).
- [4] Martinčević, I., Kozina, G. (2021). Influence of digital technologies and its technological dynamics on company management, *Tehnički Vjesnik – Technical Gazette*, Vol. 28, No. 4, 1262-1267, doi: [10.17559/TV-20200924091906](https://doi.org/10.17559/TV-20200924091906).
- [5] Carrascosa, M. (1995). *Variance of the output in a deterministic two-machine line*, Master thesis, Massachusetts Institute of Technology, Cambridge, USA, from <https://web.mit.edu/manuf-sys/www/oldcell1/theses/carrascosa-ms.pdf>, accessed January 17, 2024.
- [6] Yeralan, S., Tan, B. (1997). Analysis of multistation production systems with limited buffer capacity part 1: The subsystem model, *Mathematical and Computer Modelling*, Vol. 25, No. 7, 109-122, doi: [10.1016/S0895-7177\(97\)00052-6](https://doi.org/10.1016/S0895-7177(97)00052-6).

- [7] Tan, B. (2000). Asymptotic variance rate of the output in production lines with finite buffers, *Annals of Operations Research*, Vol. 93, No. 1, 385-403, doi: [10.1023/A:1018992327521](https://doi.org/10.1023/A:1018992327521).
- [8] Tan, B. (1997). Variance of the throughput of an N-station production line with no intermediate buffers and time dependent failures, *European Journal of Operational Research*, Vol. 101, No. 3, 560-576, doi: [10.1016/S0377-2217\(96\)00191-9](https://doi.org/10.1016/S0377-2217(96)00191-9).
- [9] Tan, B. (1999). Asymptotic variance rate of the output of a transfer line with no buffer storage and cycle-dependent failures, *Mathematical and Computer Modelling*, Vol. 29, No. 7, 97-112, doi: [10.1016/S0895-7177\(99\)00065-5](https://doi.org/10.1016/S0895-7177(99)00065-5).
- [10] Tan, B. (1999). Variance of the output as a function of time: Production line dynamics, *European Journal of Operational Research*, Vol. 117, No. 3, 470-484, doi: [10.1016/S0377-2217\(98\)00266-5](https://doi.org/10.1016/S0377-2217(98)00266-5).
- [11] Behmanesh, R., Rahimi, I. (2021). Improved ant colony optimization for multi-resource job shop scheduling: A special case of transportation, *Economic Computation and Economic Cybernetics Studies and Research*, Vol. 55, No. 4, 277-294, doi: [10.24818/18423264/55.4.21.18](https://doi.org/10.24818/18423264/55.4.21.18).
- [12] Kim, D.S., Alden, J.M. (1997). Estimating the distribution and variance of time to produce a fixed lot size given deterministic processing times and random downtimes, *International Journal of Production Research*, Vol. 35, No. 12, 3405-3414, doi: [10.1080/002075497194156](https://doi.org/10.1080/002075497194156).
- [13] Tan, B. (1998). An analytic formula for variance of output from a series-parallel production system with no interstation buffers and time-dependent failures, *Mathematical and Computer Modelling*, Vol. 27, No. 6, 95-112, doi: [10.1016/S0895-7177\(98\)00031-4](https://doi.org/10.1016/S0895-7177(98)00031-4).
- [14] El Abbadi, L., Elrhanimi, S., El Manti, S. (2020). A literature review on the evolution of lean manufacturing, *Journal of System and Management Sciences*, Vol. 10, No. 4, 13-30, doi: [10.33168/JSMS.2020.0402](https://doi.org/10.33168/JSMS.2020.0402).
- [15] Marinas, M., Dinu, M., Socol, A.G., Socol, C. (2021). The technological transition of European manufacturing companies to Industry 4.0. Is the human resource ready for advanced digital technologies? The case of Romania, *Economic Computation and Economic Cybernetics Studies and Research*, Vol. 55, No. 2, 23-41, doi: [10.24818/18423264/55.2.21.02](https://doi.org/10.24818/18423264/55.2.21.02).
- [16] Neuts, M.F. (1984). Matrix-analytic methods in queuing theory, *European Journal of Operational Research*, Vol. 15, No. 1, 2-12, doi: [10.1016/0377-2217\(84\)90034-1](https://doi.org/10.1016/0377-2217(84)90034-1).
- [17] Chakravarthi, S.R., Neuts, M.F. (2014). Analysis of a multi-server queueing model with MAP arrivals of regular customers and phase type arrivals of special customers, *Simulation Modelling Practice and Theory*, Vol. 43, No. 1, 79-95, doi: [10.1016/j.simpat.2014.01.008](https://doi.org/10.1016/j.simpat.2014.01.008).
- [18] Lucantoni, D.M. (1991). New results on the single server queue with a batch Markovian arrival process, *Communications in Statistics. Stochastic Models*, Vol. 7, No. 1, 1-46, doi: [10.1080/15326349108807174](https://doi.org/10.1080/15326349108807174).
- [19] Lucantoni, D.M. (1993). The BMAP/G/1 queue: A tutorial, In: Donatiello, L., Nelson, R. (eds.), *Performance evaluation of computer and communication systems. Performance SIGMETRICS 1993, Lecture notes in computer science*, Vol. 729, Springer, Berlin, Germany, 330-358, doi: [10.1007/BFb0013859](https://doi.org/10.1007/BFb0013859).
- [20] Montoro-Cazorla, D., Pérez-Ocón, R. (2014). A reliability system under different types of shock governed by a Markovian arrival process and maintenance policy K, *European Journal of Operational Research*, Vol. 235, No. 3, 636-642, doi: [10.1016/j.ejor.2014.01.021](https://doi.org/10.1016/j.ejor.2014.01.021).
- [21] Visagan, A., Ganesh, P. (2022). Parametric optimization of two point incremental forming using GRA and TOPSIS, *International Journal of Simulation Modelling*, Vol. 21, No. 4, 615-626, doi: [10.2507/IJSIMM21-4-622](https://doi.org/10.2507/IJSIMM21-4-622).
- [22] Fan, Y.Y. (2022). Demand prediction of production materials and simulation of production management, *International Journal of Simulation Modelling*, Vol. 21, No. 4, 720-731, doi: [10.2507/IJSIMM21-4-CO20](https://doi.org/10.2507/IJSIMM21-4-CO20).
- [23] Diamond, J.E., Alfa, A.S. (2000). On approximating higher order MAPs with MAPs of order two, *Queueing Systems*, Vol. 34, No. 1, 269-288, doi: [10.1023/A:1019165221472](https://doi.org/10.1023/A:1019165221472).
- [24] Liu, L.M., Yuan, X.-M., Liu, J.J. (2004). Operational capacity allocation for unreliable module-based assembly systems, *European Journal of Operational Research*, Vol. 155, No. 1, 134-153, doi: [10.1016/S0377-2217\(02\)00875-5](https://doi.org/10.1016/S0377-2217(02)00875-5).
- [25] Andersen, A.T., Nielsen, B.F. (2002). On the use of second-order descriptors to predict queueing behavior of MAPs, *Naval Research Logistics*, Vol. 49, No. 4, 391-409, doi: [10.1002/nav.10015](https://doi.org/10.1002/nav.10015).
- [26] Wang, Y.J., Wang, N.D., Cheng, S.M., Zhang, X.C., Liu, H.Y., Shi, J.L., Ma, Q.Y., Zhou, M.J. (2021). Optimization of disassembly line balancing using an improved multi-objective Genetic Algorithm, *Advances in Production Engineering & Management*, Vol. 16, No. 2, 240-252, doi: [10.14743/apem2021.2.397](https://doi.org/10.14743/apem2021.2.397).
- [27] Altioek, T. (1985). On the phase-type approximations of general distributions, *IIE Transactions*, Vol. 17, No. 2, 110-116, doi: [10.1080/07408178508975280](https://doi.org/10.1080/07408178508975280).
- [28] Albu, A.V., Caciora, T., Berdenov, Z., Ilies, D.C., Sturzu, B., Sopota, D., Herman, G.V., Ilies, A., Kecse, G., Gherghes, C.G. (2021). Digitalization of garment in the context of circular economy, *Industria Textila*, Vol. 72, No. 1, 102-107, doi: [10.35530/IT.072.01.1824](https://doi.org/10.35530/IT.072.01.1824).
- [29] He, X.-F., Wu, S., Li, Q.-L. (2007). Production variability of production lines, *International Journal of Production Economics*, Vol. 107, No. 1, 78-87, doi: [10.1016/j.ijpe.2006.05.014](https://doi.org/10.1016/j.ijpe.2006.05.014).
- [30] Ojstersek, R., Javernik, A., Buchmeister, B. (2023). Optimizing smart manufacturing systems using digital twin, *Advances in Production Engineering & Management*, Vol. 18, No. 4, 475-485, doi: [10.14743/apem2023.4.486](https://doi.org/10.14743/apem2023.4.486).
- [31] Jia, C., Ding, H., Zhang, X. (2021). Reliability evaluation of direct current distribution system for intelligent buildings based on big data analysis, *Tehnički Vjesnik – Technical Gazette*, Vol. 28, No. 5, 1769-1781, doi: [10.17559/TV-20210507090202](https://doi.org/10.17559/TV-20210507090202).
- [32] Riedel, A., Gerlach, J., Dietsch, M., Herbst, S., Engelmann, F., Brehm, N., Pfeifroth, T. (2021). A deep learning-based worker assistance system for error prevention: Case study in a real-world manual assembly, *Advances in Production Engineering & Management*, Vol. 16, No. 4, 393-404, doi: [10.14743/apem2021.4.408](https://doi.org/10.14743/apem2021.4.408).

- [33] Zhao, Y., Wei, R., Zhong, C. (2021). Research on spatial spillover effects and regional differences of urban housing price in China, *Economic Computation and Economic Cybernetics Studies and Research*, Vol. 55, No. 2, 211-228, doi: 10.24818/18423264/55.2.21.13.
- [34] Kim, J.A. (2022). A case study of domain engineering in software product line engineering, *Journal of Logistics, Informatics and Service Science*, Vol. 9, No. 1, 97-115.

Appendix A

In this appendix we come up with two algorithms to calculate the output variance as well as due-time variance of production variance.

Algorithm 1. Calculation of the output variance.

INPUT: The parameters $\mu_1, \alpha_1, \beta_1, N_1, \theta_1, \delta_1; \mu_2, \alpha_2, \beta_2, N_2, \theta_2, \delta_2; \dots; \mu_{M-1}, \alpha_{M-1}, \beta_{M-1}, N_{M-1}, \theta_{M-1}, \delta_{M-1}; \mu_M, \alpha_M, \beta_M, \theta_M, \delta_M$.

OUTPUT: The output variance $\text{Var}[O_{t1}]$.

Calculation:

Calculate the matrix identifier (E_1, F_1) of MAP_1 in the first step.

Calculate the matrix identifier (E_1^*, F_1^*) of the MAP_1^* of the second order in the second step.

Calculate the matrix identifier (E_2, F_2) of MAP_2 in the third step.

Calculate the matrix identifier (E_2^*, F_2^*) of the MAP_2^* of the second order in the fourth step.

Continuously calculate the matrix identifier (E_{M-1}, F_{M-1}) of the MAP_{M-1} in the fifth step.

Calculate the variance of O_{t1} in the sixth step.

Algorithm 2. Calculation of the due-time variance.

INPUT: The parameters $\mu_1, \alpha_1, \beta_1, N_1, \theta_1, \delta_1; \mu_2, \alpha_2, \beta_2, N_2, \theta_2, \delta_2; \dots; \mu_{M-1}, \alpha_{M-1}, \beta_{M-1}, N_{M-1}, \theta_{M-1}, \delta_{M-1}; \mu_M, \alpha_M, \beta_M, \theta_M, \delta_M$.

OUTPUT: The due-time variance $\text{Var}[D_{n1}]$.

Calculation:

Calculate the matrix identifier (E_1, F_1) of MAP_1 in the first step.

Calculate the matrix identifier (E_1^*, F_1^*) of the MAP_1^* of the second order in the second step.

Calculate the matrix identifier (E_2, F_2) of MAP_2 in the third step.

Calculate the matrix identifier (E_2^*, F_2^*) of the MAP_2^* of the second order in the fourth step.

Continuously calculate the matrix identifier (E_{M-1}, F_{M-1}) of the MAP_{M-1} in the fifth step.

Calculate the variance of D_{n1} in the sixth step.

A comparative study of machine learning regression models for production systems condition monitoring

Jankovič, D.^a, Šimic, M.^{a,*}, Herakovič, N.^a

^aUniversity of Ljubljana, Faculty of Mechanical Engineering, Department of Manufacturing Technologies and Systems, Ljubljana, Slovenia

ABSTRACT

This research investigates the benefits of different Machine Learning (ML) approaches in production systems, with respect to the given use case of considering the forming process and different friction conditions on hydraulic press response in between the phases of the sheet metal bending cycle, i.e. bending, levelling and movement. A framework for enhancing production systems with ML facilitates the transition to smarter processes and enables fast, accurate predictions integrated into decision-making and adaptive control. Comparative ML analysis provides insights into predictive regression models for hydraulic press condition recognition, enhancing process improvement. Our results are supported by performance evaluation metrics of predictive accuracy RMSE, MAE, MSE and R^2 for Linear Regression (LR), Decision Trees (DT), Support Vector Machine (SVM), Gaussian Process Regression (GPR) and Neural Network (NN) models. Given the remarkable predictive accuracy of the regression models with R^2 values between 0.9483 and 0.9995, it is noteworthy that less complex models exhibit significantly shorter training times, up to 437 times shorter than more complex models. In addition, simpler models have up to 36 times better prediction rates, compared to more complex models. The fundamentals illustrate the trade-offs between model complexity, accuracy and computational training and prediction rate.

ARTICLE INFO

Keywords:

Hydraulic press;
Metal forming;
Machine Learning (ML);
Linear Regression (LR);
Decision Trees (DT);
Support Vector Machine (SVM);
Gaussian Process Regression (GPR);
Artificial Neural Networks (ANN)

*Corresponding author:
marko.simic@fs.uni-lj.si
(Šimic, M.)

Article history:

Received 1 February 2024
Revised 19 April 2024
Accepted 23 April 2024



Content from this work may be used under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

Challenges in manufacturing processes often require straightforward solutions or are so complex that they present technicians with seemingly unsolvable problems. Simulation approaches are often used to overcome these challenges in order to gain a more detailed understanding of the system under consideration and the manufacturing processes running on it [1]. However, the underlying problems are sometimes indirect when comparing the real environment with the virtual environment established in a simulation model [2]. With the advent of Industry 4.0, technological solutions have emerged in digitalization and automation processes, often using artificial intelligence (AI) approaches [3]. The synergy of edge computing (EC) and 5G networks leverages high bandwidths and low latency to process data closer to the source, enabling faster response times and more efficient, stable and secure data transfer and data management, which has a positive impact on the effectiveness and quality of production processes as well as sustain-

ability and circular economy improvement [4]. The ML paradigm enables a more in-depth system analysis that captures properties that cannot be evaluated by visual representation and graphical interpretation alone. Classification and regression models based on data-driven approaches have already been successfully implemented in many technical fields, offering high precision in predicting system behaviour and often overcoming 99 % prediction efficiency in real systems [5, 6].

Hydraulic presses play a crucial role in various forming processes such as bending, stamping, forging and drawing, where the aim is to convert hydraulic energy into deformation energy of the workpiece. The dynamic properties of hydraulic presses and bending processes, including material properties, fluid properties and friction behaviour, are interrelated and influence each other. These complex relationships require in-depth investigations, such as those presented in this study, to gain a comprehensive understanding of how the bending process affects the response of the hydraulic press. The quality of the forming process has a significant impact on the overall quality of the product, especially when disturbances, such as dynamic frictional properties, are present [7, 8]. Identifying altered conditions is a fundamental step in comprehending system behaviour, as highlighted in prior research on issues like hydraulic valve wear [9], pump malfunction [10] and leakage [11] has shown. To summarise, artificial intelligence offers data-driven modelling solutions tailored to different engineering challenges and based on different principles. Briefly speaking, ML provides solutions for the identification of faults and malfunctions of hydraulic components, which are crucial for decision making to adapt and improve the operation of hydraulic presses.

A detailed analysis in the field of hydraulic systems shows that there is no simple method for selecting the most appropriate ML approach, regardless of the input parameters. Given the use of numerous classification and regression methods in studies, a significant dilemma arises when choosing between less complex methods such as LR, DT, SVR and more complex methods such as GPR and NN for the study of specific research problems [12–17]. Another possible shortcoming of the previously referenced research refers to the fact the authors believe that certain derivations of ML methods are best suited for their particular use case due to their complexity. This means that the optimization and prediction of performance depends on the method chosen for the particular application, which requires careful consideration. Su *et al.* [18] demonstrated that when comparing the ML methods LR, K Nearest Neighbour (KNN), DT, SVM and NN, the NN method outperformed the others with an accuracy of 99.8 % in predicting valve flow. Highlighting the best ML method, it is worth noting that other ML models achieved robust prediction accuracy, with the worst performing LR model still achieving 99.1 %. Moreover, the same conclusions have been confirmed by research groups investigating other use cases [19–22]. On the other hand, Guo *et al.* [23] emphasize that training and prediction times are a crucial parameter for the efficient integration of ML models into decision algorithms. It determines how quickly the entire decision-making system will react to changing conditions and therefore allows real-time actions such as control and parameters set-up.

By pointing out the basic assumptions and an analysis of the available ML algorithms, the objective of this paper is to demonstrate the advantages of the five basic regression modelling types LR, DT, SVM, GPR, NN in the investigation of a hydraulic press under different intensities of the forming process and friction dynamics. Chapter 2 introduces the background and the research problem, supplemented by Design of Experiments (DOE), data extraction and data pre-processing. Section 3 introduces different types of regression models and defines the range of hyperparameters to obtain the best prediction from five ML approaches. Finally, Section 4 presents the best fitted regression models along with the most efficient selection of hyperparameters.

2. Experiment and dataset analysis

2.1 A case study of hydraulic press

This study focuses on a hydraulic press subjected to the conditions of the forming process, including the bending process and the constraints on the movement of the hydraulic cylinder resulting from the friction between the press guides and the hydraulic cylinder, as shown in Fig. 1. A hydraulic press is a sophisticated and robust mechanical device designed for various industrial

applications that utilises the principles of fluid mechanics to generate significant force for forming, shaping and compressing materials. The design of a hydraulic press combines

- hydraulic components, i.e. hydraulic valve, hydraulic cylinder, hydraulic power unit (hydraulic pump with electric motor and controller),
- mechanical components, i.e. guiding system, hydraulic press frame, mechanical system (bending process),
- control system, e.g. PLC Beckhoff Controller CX 9020, PID Moog Controller, Raspberry Pi.

The components listed above with their respective characteristics form a distributed network in which data collection and analysis takes place. They therefore form the basis for integration into the concept of edge computing and 5G communication technology and particular on 5G real-time data transfer. The integration of these components enables the hydraulic press to effectively convert hydraulic energy into mechanical power. As explained by Jankovic D. et. Al [24], these influences contribute to a delayed movement of the hydraulic cylinder, which is determined not only by the mentioned factors, but also by the intensity of these factors and the velocity of the hydraulic cylinder. To achieve this goal, 30 different scenarios were analysed in the experiments, focusing on the effects of the forming process. In addition, a further 46 scenarios were systematically carried out to evaluate the influence of different operating conditions on the reaction of the hydraulic press.

Forming processes require the mechanical energy of a hydraulic press to change the shape of the specimen. During this process, the deformation of the specimen restricts the movement of the hydraulic cylinder. Consequently, the applied hydraulic energy increases to ensure that the movement of the hydraulic cylinder measured by the displacement sensor X_C matches the referenced movement X_{ref} controlled by a PLC controller and given on the basis of a predefined sheet metal bending cycle. However, the ability of the controller to fully compensate for the described causes is not given, as shown by the response error ΔX_C of the hydraulic press, which is evaluated as the difference between the measured X_C and the referenced X_{ref} displacement of the hydraulic cylinder. In addition, the pressure sensors are integrated to monitor the pressure conditions p_A, p_B in both hydraulic cylinder chambers. In addition, a force sensor is positioned between the hydraulic cylinder and the pressing plate to monitor the generated hydraulic cylinder pressing force F_C . The integrated LVDT sensor allows to foresee the opening of hydraulic valve X_V . The intensity of the bending process was controlled by applying combinations of hydraulic cylinder velocities (5 mm/s, 15 mm/s, 25 mm/s) and the width of the samples (10 mm, 20 mm, 30 mm, 40 mm). To capture the analog signals from the sensors, Beckhoff modules were used to import the data into the digital datasets. Each scenario was experimentally performed three times to confirm the repeatability of the experiments.

The friction conditions in the press guides arise due to the friction between two sliding surfaces. Stribeck friction, which is characterised by its dynamic nature, is defined by the correlation between frictional force and velocity of movement [7]. The choice of sealing and guiding technology in hydraulic cylinders is influenced by the quality and stability of the dynamic friction of hydraulic cylinders [25]. Advanced guiding and sealing systems enable a more stable friction profile over the entire range (static, hydrodynamic and mixed friction). In contrast, conventional sealing and guiding systems exhibit characteristics in which the static friction phase is significantly higher than the hydrodynamic phase. In addition, higher frictional forces occur in conventional sealing and guiding systems under the same velocity conditions. The implementation of a subsystem with three pulleys and a cable allows the emulation of different friction scenarios during the phases of the sheet metal bending cycle. By adjusting the load, the desired friction scenario is achieved and measured directly by a force gauge attached to the pressing plate, which reflects the resistance force during the movement of the hydraulic cylinder.

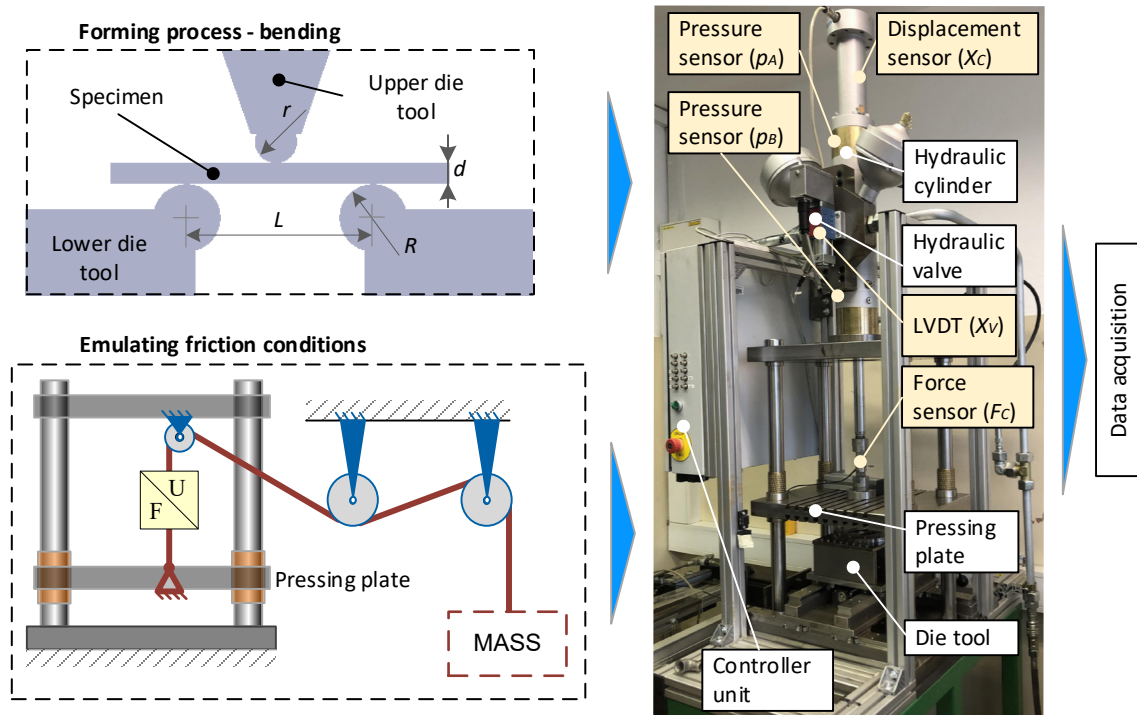


Fig. 1 Hydraulic press condition analysis

2.2 Experimental analysis

The data is visualised using variables measured in real time (sampling frequency of 100 Hz): Pressure in the upper and lower hydraulic cylinder chambers p_A , p_B , hydraulic valve opening X_V , hydraulic cylinder force F_C and hydraulic cylinder displacement X_C . The visualisation and interpretation of the raw data are the most important steps in data analytics. The extraction of relevant data intervals enables data-driven models to deliver enhanced prediction performance [26]. Furthermore, Fig. 2 shows the comparison of two main causes and their effects on the hydraulic press response, with the least intensive and the most intensive scenario shown for each cause to illustrate which measured variable offers the most information in each scenario.

The sheet metal bending cycle enhances three phases:

- Fast- forward movement (1) and fast- backward movement (4) of the hydraulic cylinder to minimise the duration of the sheet metal bending cycle and enable the production of as many parts as possible within the allotted time frame.
- In the bending phase (2), a bending force is applied to deform a sheet metal workpiece and bring it into the desired shape.
- Levelling phase (3) includes processes aimed at achieving uniformity and flatness of the sheet metal workpiece after the bending phase, with the die tool remaining stationary in its position.

The presented scenarios depict a selected sheet metal bending cycle, with the hydraulic cylinder velocity set to 50 mm/s during the movement phase (phases 1 & 4) and 15 mm/s during the bending phase (phase 2). Looking at the cause of the forming process shown in Fig. 2(a), the first scenario represents the intensity of unrestricted movement without a test specimen, while the second scenario shows the highest forming intensity, involving the placement of a specimen with a width of 40 mm is inserted into the die tool. The obvious changes occur when the die tool comes into contact with the specimen, resulting in a significant increase in the forming force F_C , which changes during the bending phase and remains semi-stationary during the levelling phase. The same tendency is observed for the variable p_A , while the variable p_B decreases as expected. During this event, the opening of the hydraulic valve X_V increases to compensate for the

hydraulic press response error ΔX_C , which persists until the end of the levelling phase, when the pressing plates release the elastic tension in the sample.

The analysis of different friction scenarios shown in Fig. 2(b) indicates non-visible changes in the hydraulic cylinder displacement X_C , which can be seen by averaging the value over the phase range. The first scenario signifies more favourable friction conditions with no restraining force, while the second scenario conditions lower friction characteristics, simulated by a pulley system emulating a restraint force of 1 kN. The most noticeable changes manifest in a visible rise in the forming force during the sheet metal bending cycle and the variable p_A , accompanied by an expected decrease in the value of the variable p_B . In addition, a slight increase in the hydraulic valve opening is observed. Clearly, the influence of the friction conditions on hydraulic press response error is less significant compared to the impact observed in the case of the forming process, especially when considering maximum force F_C required to deform the sample at 15 kN.

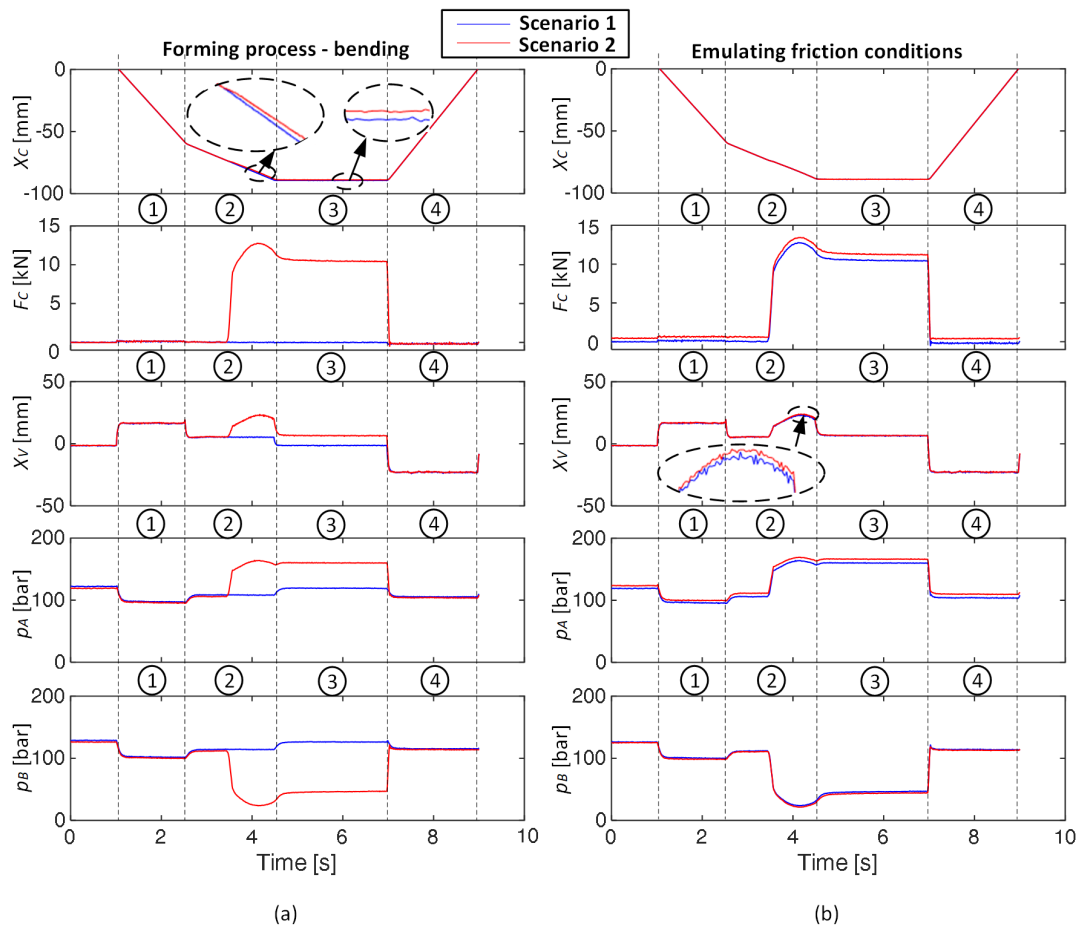


Fig. 2 Data visualization of measured variables of: (a) forming process-bending and (b) friction conditions

2.3 Data pre-processing

In regression learning, the crucial step of feature selection involves identifying and include relevant input variables, thereby enhancing the accuracy and predictive ability of the model. Moreover, regression modelling is a statistical method that reveals features that are not readily apparent in the visualisation phase, yet provides indispensable insights into data characteristics that may otherwise go unnoticed [16]. In addition, the pressure difference Δp in hydraulic cylinder chambers is assumed to be the difference between the variables p_A and p_B , providing a higher intensity of information [27]. In addition, the hydraulic cylinder velocity v_C is considered more meaningful than the measured hydraulic cylinder displacement ΔX_C , as it characterises much more dynamic conditions [28]. While the purpose of regression models is to predict the hydraulic press response error ΔX_C , the input data for regression models include the variables ΔX_C , F_C , X_V , v_C and Δp .

Industry 4.0 underlines the importance of using smart data to minimize latency and optimize the efficiency of data transmission [29]. The case study includes an investigation of 76 different conditions under which the hydraulic press was examined. However, it's crucial that the raw data for each condition includes information from different scenarios. To address this, the data must be pre-processed separately and combined into one dataset covering all scenarios so that the regression model can account for all scenarios. In addition, employing the data bundling enables the packaging of only pertinent sections into a unified dataset for each phase of the sheet metal bending cycle, as shown in Fig. 3. During the pre-processing stage, the variable sections were carefully selected to illustrate the semi-stable operating condition of the hydraulic press over a time scale. Furthermore, the extracted sections were selected considering the visualization aspects described in chapter 2.2. In addition, a feature scaling was performed to standardize the importance of the five variables and to prevent any single variable from being prioritized [30]. The scaling was performed to normalize the y-axis by setting it in the range from 0 to 1. Finally, the extracted datasets representing different scenarios were partitioned into training and test datasets to evaluate the regression models using known metrics, including Root Mean Square Error (RMSE), coefficient of determination R^2 , Mean Squared Error (MSE) and Mean Average Error (MAE). Furthermore, training datasets are used during the model training phase to determine the required training time for different regression model approaches. In addition, test datasets were used to evaluate the prediction accuracy of each model after the training phase, considering the testing time required to perform the predictions.

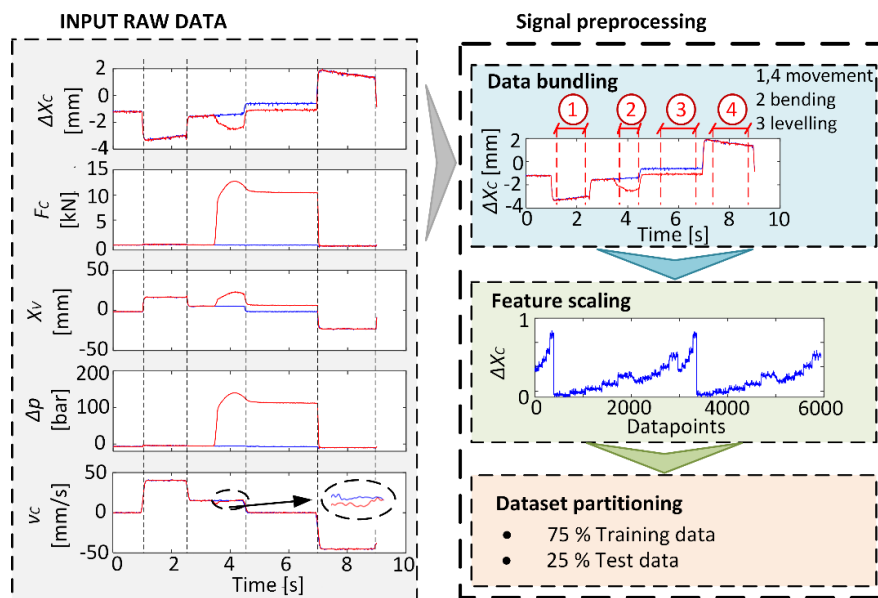


Fig. 3 Smart data forming

3. Regression models

The availability of regression modelling methods offers a wide range of possibilities to achieve semi-mirrored data-driven model for the chosen research problem. Grid search and random search are two commonly employed methods for optimizing hyperparameters. While grid search exhaustively explores all possible parameter combinations, it incurs high computational costs, especially when a large number of parameters are involved [18]. In contrast, the random search examines potential parameter combinations randomly within fixed ranges and offers a significant reduction in computational cost. In our study, both approaches are used for hyperparameter optimization. Grid search is applied to LR, GPR and SVR, while random search is applied to DT and NN depending on the number of hyperparameters involved. The MATLAB application Regression Learner was used to train and test the regression models.

Five-fold cross-validation was used to determine the hyperparameters that provide the best generalization performance for our models. In each iteration, a specific subset was selected for testing, while the remaining subsets were used for training. The investigated hyperparameters for the models are summarized for each model in its subsection.

3.1 Linear regression

Linear regression (LR) uses an assumption function described in Eq. 1 to model and capture the relationships between independent variables x , where β is the coefficient vector, n is the number of samples, and y is the target value [12]. In our approach, we include interaction terms in the linear regression model; nevertheless, we also use the classical LR approach. This decision is based on previous studies that indicate a varying degree of influence of different independent variables on the dependent variable [24]. Including interaction terms allows the model to capture and account for the joint effects of these variables, taking into account the potential dependencies and relationships identified in previous research [12].

The objective of minimizing the objective function is to reduce the sum of squared differences between the observed values and the corresponding predicted values, as described in Eq. 2 by the least squares method.

$$y(x) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_2(x_1 \cdot x_2) + \dots + \beta_nx_n \tag{1}$$

$$J(\beta) = \frac{1}{2n} \sum_{i=1}^n (y(x_i) - y_i)^2 \tag{2}$$

3.2 Decision trees

Decision trees assume nonlinear patterns and relative relationships in the data and perform feature selection for the most accurate prediction [13]. In addition, DT have simple extraction rules, are very accurate and provide good interpretability of the models. In our approach, the search for the optimal parameters for DT involves varying the leaf size (4, 12 and 36).

3.3 Support vector regression

The SVR is an extension of support vector machine developed for single-output regression [14]. In addition, the SVR determines a linear dependence between the independent variable x and the dependent variable y , as shown in Eq. 3, where w is the weight vector and b is the intercept.

$$y = wx + b \tag{3}$$

The goal of objective function in SVR is to find the weight vector w and the intercept b such that configures the least deviation in between the predicted and actual values as expressed in Eq. 4. Here, a regularization parameter C balances the trade-off between w and the slack variables ξ, ξ^* , under the conditions expressed in Eq. 5. Fig. 4 describes the linear kernel function approximation considering the data accuracy of ϵ .

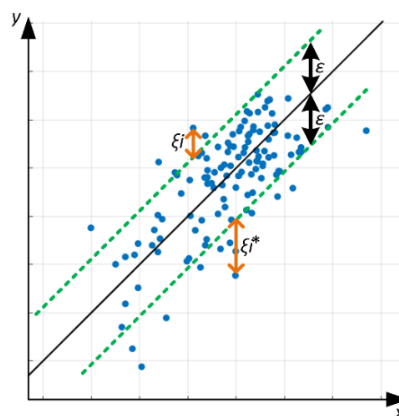


Fig. 4 Linear kernel function for SVR

$$J(w, b, \xi, \xi^*) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (4)$$

$$\begin{cases} y_i - wx_i - b \leq \varepsilon + \xi_i \\ wx_i + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* > 0 \end{cases} \quad (5)$$

The decision function in SVR with a linear kernel is determined in Eq. 6.

$$y = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x, x_i) + b \quad (6)$$

Enhancing the accuracy of nonlinear predictions, we consider a number of kernel options described in Eq. 6 through 10, including linear, quadratic K_q , cubic K_c , exponential K_e , and squared exponential K_{se} , where c is a constant term. These kernels are systematically evaluated using a random search approach to uncover the most appropriate SVR model.

$$K_q = (x + c)^2 \quad (7)$$

$$K_c = (x + c)^3 \quad (8)$$

$$K_e = \exp\left(-\frac{\|x_i - x_j\|}{l}\right) \quad (9)$$

$$K_{se} = \exp\left(-\frac{\|x_i - x_j\|^2}{2l^2}\right) \quad (10)$$

3.4 Gaussian process regression

Due to the non-parametric Bayesian approach to regression problems, the GPR offers a high degree of flexibility for the targeted prediction of complex relationships in the data [16]. It assumes a Gaussian data distribution, where the model hyperparameters, i.e. the variance of the distribution σ_f and the length parameter l , are determined by the mean function $m(x)$ and the kernel K , as shown in Eq. 11 [31]. The purpose of the mean function is to represent the expected trend of the decision function y , which is assumed to be a zero-mean function in our case [24]. In addition, the purpose of the kernel function is to assign importance rates to the training datasets and to fit the model.

$$y \sim GP(m(x), K) \quad (11)$$

To extract the most accurate GPR model, a σ_f of 0.001-0.100 was varied. In the training process, Automatic Relevance Determination (ARD) allows the model to automatically determine the relevance of each predictor by adjusting the length parameter l , considering different types of kernels as shown in Eq. 6 to 10.

3.5 Neural network

Neural networks consist of neurons organized in layers that determine the connections between neurons through a series of weights and biases [22]. The input data is processed to provide an output of each neuron, which allows flexibility in determining the number of neurons in each layer and the number of layers in the neural network to best fit the data. Neural networks are composed of an input layer, one or more hidden layers and an output layer of artificial neurons, as shown in Fig. 5. Artificial Neural Network (ANN) is a broad term that encompasses NN models of varying complexity, while Deep Neural Network (DNN) specifically refers to networks with multiple hidden layers that emphasize the depth of the architecture, which is usually associated with an increased ability to learn complicated representations from data [23]. Models built in our case with NN include five neurons in the input layer, since the available predictors are five. The output layer consists of one neuron, since only one output estimate is predicted.

In the configuration of the neural network, we varied the number of neurons in the hidden layer between 10, 25 and 100, while the number of layers ranged from one to three. ReLU was chosen as the activation function, with a maximum of 1000 iterations and a regularization strength of 0.

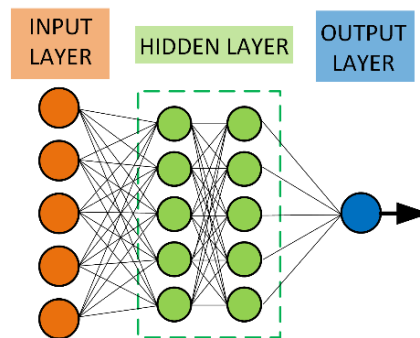


Fig. 5 Neural network structure

4. Results and discussion

4.1 Characteristics of fitted regression models

The performance of regression models of different types depends on model complexity, data quality and quantity, feature selection, assumption violations, pre-processing stage and hyperparameter selection [18]. The results shown in Table 1 represent the best-fit regression models with the most optimal selection of hyperparameters for the given case study and the range of regression model trained and test fitted in the MATLAB environment. Considering that the research problem focuses on three phases of the sheet metal bending cycle, i.e. bending, levelling movement, the complexity of the regression model differs due to the nature of the data patterns, the presence of white noise and the number of conditions summarised in each phase.

Regression modelling for LR yields superior results when interaction terms are included across all phases of the sheet metal bending cycle than the method without interaction terms. In addition, the DT type of regression modelling shows the importance of selecting hyperparameter variation. In the bending and levelling phase, the necessary depth of the DT is significantly higher than in the case of the movement phase. In addition, the medium size tree is mostly appropriate for the bending phase, while the coarse tree size is the most appropriate for the levelling and movement phases, considering the number of leaves and the size of the parents. The results in the table for the SVR regression modelling type show that the squared exponential kernel function is most effective in representing complex relationships and patterns among the input data. In addition, the hyperparameters C and ϵ reflect the trade-off between the accurate fit of the training data and the margin of tolerance, which is comparable between phases. Moreover, selected hyperparameters with low values indicate an easy fit of the model and a relatively narrow range, emphasising the minimization of errors in the predictions. In addition, the SVR regression model shows that less complexity is required in the case of the bending phase and the margin of error is even narrower than in the SVR regression models for the levelling and movement phases. Moreover, the results of the GPR regression models show that the Exponential kernel function consistently provides the most optimal fit among the different kernels available. In the levelling phase, it is evident that a higher standard deviation of noise is required as the data has a higher amount of white noise compared to the movement and bending phase. For the application of regression modelling using NN, it has been shown that a single layer is sufficient for the levelling and bending phases. However, for optimal performance in the levelling phase, it is best to use three layers with a layer size of 10.

Table 1 Best fitted regression models summary

Regression model	Hyperparameters and characteristics	Bending phase	Levelling phase	Movement phase
Linear regression (LR)	Interactions terms	ON	ON	ON
	Robust option	OFF	OFF	OFF
Decision Trees (DT)	Max. depth	410	390	162
	Min. leaf size	12	36	36
	Min. Parent size	24	72	72
	Max. Number Splits	4454	9067	8666
	Split Criterion	MSE	MSE	MSE
	Prune	ON (RMSE)	ON (RMSE)	ON (RMSE)
	Surrogate	OFF	OFF	OFF
Support Vector Regression (SVR)	Cost C	0.1848	0.2708	0.2768
	Epsilon ϵ	0.0185	0.0271	0.0277
	Kernel type	K_{SE}	K_{SE}	K_{SE}
	Kernel Scale	0.5	0.5	2
	Iteration limit	10^6	10^6	10^6
Gaussian Process Regression (GPR)	Kernel type	K_E	K_E	K_E
	Standard deviation of noise σ_n	0.0100	0.0438	0.0038
	Quasi-Newton optimization	ON	ON	ON
Neural Network (NN)	Number of layers	1	1	3
	Layer size	100	100	10, 10, 10
	Activation	ReLU	ReLU	ReLU
	Iteration Limit	1000	1000	1000
	Regularization strength	0	0	0

4.2 Regression model performance evaluation

The overall performance of the regression models was assessed using the validation and test results to confirm that the data-driven models do not overfit. The results for the bending phase are presented in Table 2 and show significant performance accuracy in the validation stage by the R^2 metric, which is at a value between 0.9945 and 0.9961. In addition, the complexity of the regression modelling affects the performance of the regression models. This is particularly the case for less complex regression modelling approaches such as LR and DT. In such examples, the values of the LR regression model for the metrics RMSE, R^2 , MSE, MAE are 0.01364, 0.9945, 0.000186, 0.009998 in the validation and 0.01368, 0.9941, 0.000187, 0.009836 in the test. As the complexity of the regression modelling increases, the performance of the regression model increases, which is evident in the regression models GPR and NN by higher evaluation metrics RMSE, R^2 , MSE, MAE. In comparison, the best fitted regression model in the case of the bending phase is the GPR model with metric values RMSE, R^2 , MSE, MAE at validation 0.01144, 0.9961, 0.000131, 0.007810 and at test 0.01150, 0.9958, 0.000132, 0.007768. Although more complex regression models, e.g., GPR and NN, perform better, less complex regression models, e.g., LR and DT, accurately predict the desired response error of the hydraulic press in different scenarios. In addition, the performances of the regression models built by different approaches provide solid results in the test stage, where the coefficient of determination is between 0.9941 and 0.9958.

Table 2 Regression model performance evaluation for bending phase

Model	Validation				Test			
	RMSE	R^2	MSE	MAE	RMSE	R^2	MSE	MAE
LR	0.01364	0.9945	0.000186	0.009998	0.01368	0.9941	0.000187	0.009836
DT	0.01357	0.9946	0.000184	0.009316	0.01265	0.9950	0.000160	0.008854
SVR	0.01252	0.9954	0.000157	0.009066	0.01238	0.9952	0.000153	0.008754
GPR	0.01144	0.9961	0.000131	0.007810	0.01150	0.9958	0.000132	0.007768
NN	0.01168	0.9960	0.000136	0.008105	0.01156	0.9958	0.000134	0.007890

The validation and test results for the levelling phase are shown in Table 3. From the metrics, the best regression model in validation is the DT model, which has the lowest RMSE, MSE and MAE at 0.04393, 0.001930 and 0.030141, but the highest R^2 at 0.9618. In addition, the other fitted regression models considering the metrics of RMSE, R^2 , MSE, MAE are in the order of GPR,

SVR, NN, LR. In addition, the test results show that the same sequence of regression models enables the best regression model performance. Moreover, the overall accuracy of the different regression models is solid, as R^2 is between 0.9517 and 0.9618 for the validation results and 0.9483 and 0.9581 for the test results. Compared to the bending phase, the performance of the regression models is lower due to the white noise contained in the experimental data.

In the movement phase, the best-fitting regression models, i.e. DT, GPR and NN, have a robust coefficient of determination R^2 of 0.9995 for both validation and test results, as shown in Table 4. In addition, DT is shown to be the best performing model with the lowest values for RMSE, MSE and MAE. The next best performing regression model for the movement phase is NN, with GPR being similarly accurate to NN in terms of the RMSE, MSE and MAE metrics. The lowest performing regression model is SVR with an R^2 value of 0.9986 in validation and 0.9987 in test. In addition, the LR model is more accurate than SVR in prediction. Overall, the prediction accuracy of the regression models is higher in the movement phase than in other phases of the sheet metal bending results, which is evident by an R^2 value of more than 0.999 in the validation and test. Overall, regression models are built accurately and are not overfitting through all phases of the sheet metal bending cycle, i.e. bending, levelling and movement, considering that the RMSE, R^2 , MSE and MAE metrics show similar values between validation and test results.

Table 3 Regression model performance evaluation for levelling phase

Model	Validation				Test			
	RMSE	R^2	MSE	MAE	RMSE	R^2	MSE	MAE
LR	0.04944	0.9517	0.002445	0.036453	0.05070	0.9483	0.002571	0.037562
DT	0.04393	0.9618	0.001930	0.030141	0.04568	0.9581	0.002086	0.031166
SVR	0.04770	0.9550	0.002275	0.034339	0.04801	0.9537	0.002305	0.034739
GPR	0.04628	0.9576	0.002142	0.033074	0.04680	0.9560	0.002190	0.033308
NN	0.04829	0.9539	0.002332	0.035171	0.04830	0.9531	0.002333	0.035264

Table 4 Regression model performance evaluation for movement phase

Model	Validation				Test			
	RMSE	R^2	MSE	MAE	RMSE	R^2	MSE	MAE
LR	0.08885	0.9991	0.007894	0.068134	0.08894	0.9991	0.007909	0.067491
DT	0.06629	0.9995	0.004395	0.048779	0.06658	0.9995	0.004433	0.048933
SVR	0.11099	0.9986	0.012318	0.089846	0.10939	0.9987	0.011966	0.089033
GPR	0.06781	0.9995	0.004598	0.050394	0.06801	0.9995	0.004626	0.050394
NN	0.06719	0.9995	0.004515	0.050851	0.06709	0.9995	0.004501	0.049864

4.3 Training and prediction rate analysis

The comparison of the training and test times in between the regression models LR, DT, SVR, GPR and NN is shown in Fig. 6. In terms of training and prediction time, the regression models LR, DT and SVR show better results for all phases of the sheet metal bending cycle, i.e. bending, levelling movement, compared to more complex regression models such as GPR and NN.

Examining the bending phase shown in Fig. 6(a), it is evident that LR, DT and SVR models exhibit short training times below 5 s and short prediction times, which are less than 8.5 ms. The best training and prediction rate were achieved by the DT model with a training time of 3.8 s and a prediction time of 3.4 ms. In contrast, the training time of the GPR and NN models is 190.6 s and 183.7 s, which corresponds to a 50 times higher rate compared to the less complex DT model. In addition, the GPR model requires 59.9 ms to predict the response error of the hydraulic press in the bending phase, compared to the NN model prediction time of 3.9 ms. Overall, the prediction rate of the NN model is comparable to other low-complexity regression models, as evidenced by the test time of LR, DT and SVR of up to 8.4 ms.

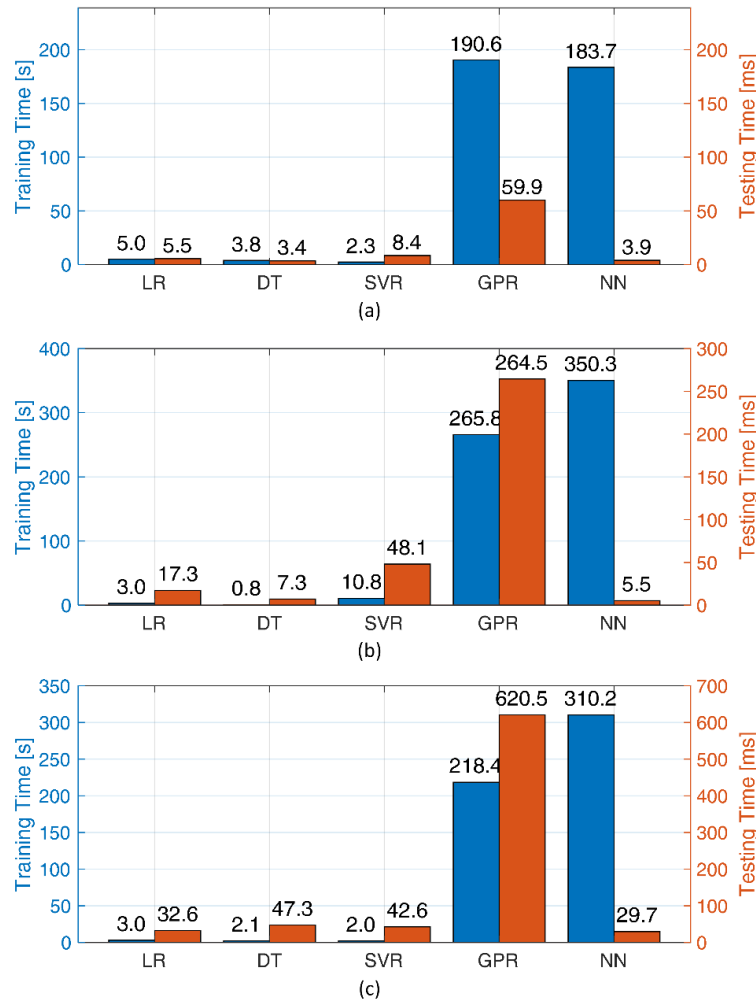


Fig. 6 Regression model training and testing time comparison for: (a) bending phase, (b) levelling phase, (c) movement phases

In the case of the levelling phase shown in Fig. 6(b), the training rate for the LR, DT and SVR models is up to 437 times better than for the GPR and NN models. In addition, the training time of the DT model is solid at 0.8 s for the entire training dataset. The training time of the NN model is the longest at 350.3 s, but the prediction time is the shortest at 5.5 ms compared to the other models. In comparison, the SVR model requires 6 times more time to predict the outcome than the DT model. The GPR model has the highest prediction rate with a prediction time of 264.5 ms. Examining the movement phase shown in Fig. 6(c), the LR, DT and SVR models require a training time of up to 3 s, which is on average better than in the bending and levelling phase. However, the LR, DT and SVR models require higher prediction times on average compared to the bending and levelling phase. In the movement phase, datasets size is higher compared to the bending and levelling phase, but the required complexity of the regression model is lower. In addition, due to the larger datasets, the number of predictions is also higher, which leads to higher prediction times. In terms of prediction time, the NN model has the lowest prediction time of 29.7 ms. In addition, the GPR model requires the longest prediction time of 620.5 ms, which is 20 times higher compared to the NN model.

4.4 A general framework for enhancing production systems using machine learning

The study also aims to provide general guidelines for transforming of conventional production systems to a smart production system using ML techniques. The main steps from 1 to 4 can be used to define the most suitable machine learning strategy and approach for particular production use case such as robotic assembly, CNC machining, etc. (Fig. 7).

Step 1: The effectiveness of these methods depends on the availability and quality of the data required for model training (Step 1, Data collection). Focusing on our use case, the results of the hydraulic press show the importance of tailored approaches to improve process control and the performance of the different data samples from production processes in terms of pattern, complexity, presence of Gaussian noise, dimensionality and transition types. In general, the most important is to design experiments and data collection approaches, focusing on selecting appropriate sensors and determining their location to extract the relevant information from the real production system and ensure the validity of the information collected. The sensors serve as the primary data source for monitoring specific characteristics and requirements relevant to the research. Furthermore, in step 4, the input variables for ML modelling are determined, which influence the predictive accuracy of the outcome variables and summarize the specific characteristics and requirements of the respective research.

Step 2: Here, the data analysis is performed by visualizing the various conditions of the production system under investigation. Learning and testing datasets should be extracted from different condition scenarios at intervals to observe the most relevant sections of the measured data for predicting outcomes of ML model.

Step 3: This step represents the signal pre-processing, which involves the condition recognition of production system. In the case of hydraulic presses condition recognition, the most important task is data pre-processing that involves bundling and scaling. In addition, when observing other production processes their characteristics are different, thereby other pre-processing methods, such as logarithmic transformation may be more appropriate to consider.

Step 4: Finally, when selecting a machine learning strategy, different ML approaches should be tested to determine the most appropriate one for the application in terms of given criteria and objective function. For hydraulic presses, the edge computing decision-making algorithm integrated into the control system is expected to be fast and accurate in order to improve the performance of the hydraulic press in real-time. The prediction rate significantly impacts the ability of the hydraulic press control system to adjust responses autonomously, increasing production reliability and reducing product uncertainty.

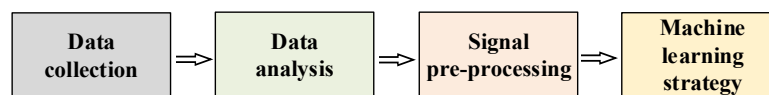


Fig. 7 Guidelines for accurate machine learning integration into production processes

5. Conclusion

This research addresses the trade-off between model complexity, prediction accuracy, and computational efficiency. In real-time applications, rapid response times are crucial for effective reactions. Therefore, the efficiency of a decision-making system that reacts to recognised conditions in the production processes depends on the prediction rate and the accuracy of the ML model. In addition, the adaptability of ML models to newly introduced conditions are determined by the learning rate. While more complex models such as GPR and NN deliver higher prediction accuracy, the associated training and prediction times need to be carefully considered, especially in different phases of the sheet metal bending cycle i.e. forming, levelling and movement. The varying performance of LR, DT, and SVR models showcases the importance of selecting models tailored to the specific requirements and constraints of the application. In conclusion, the metric R^2 yielded robust results in validation and test performance accuracy in each phase of the sheet metal bending cycle: bending 0.995, levelling 0.950, and movement 0.999. The findings suggest that more complex regression models exhibit superior performance, but they are associated with a higher training time, which is up to 437 times higher than for less complex models such as LR, DT and SVT. Notably, GPR requires even longer prediction times compared to other ML models. The difference in prediction time could be attributed to the explicit represen-

tation and computational intricacies involved in GPR during predictions, whereas NN often benefit from more implicit representations and parallelization, leading to faster predictions. Furthermore, the memory requirement with models NN is much lower compared to GPR, as GPR needs to store information for the entire training dataset, including covariance matrices. Models LR, DT, and SVR can be treated as lower-performance models, however evidently need shorter training and prediction times for the given case study. Therefore, the optimal ML models for integration into the decision-making system are either the less complex LR model, offering higher training and prediction rate, or either more intricate NN, which requires a higher training rate. However, despite higher training rates, NN showcases superior prediction accuracy and prediction rate comparable to other less complex ML models.

In summary, the design of the control system for the hydraulic press allows the consideration of predetermined operating conditions, however the occurrence of new conditions would require the optimization of machine learning hyperparameters. The methodology presented in this study provides a valuable reference point for the selection of modelling and solution methods suitable for the diagnosis of autonomous production process conditions. Given the complexity of production processes and the desired prediction results, an intelligent combination of different solution methods can effectively address these challenges. By applying multiple machine learning methods such as LR, GPR, NN and SVR, this study presents the most suitable approach that enables accurate and rapid prediction of hydraulic press response.

Acknowledgement

The work was carried out in the framework of the Slovenian Research and Innovation Agency (research program funding No. P2-0248, research project No. J2-4470 and research program for young researchers No. 53512) funded by the Republic of Slovenia – Ministry of Education, Science and Sport. The work was carried out in the framework of research project STAGE – Sustainable Transition to the Agile and Green Enterprise funded under the Grant agreement ID: 101058693.

References

- [1] Li, J., Lu, Y., He, F., Miao, L. (2023). Motion control of hydraulic actuators with nonlinear friction compensation: Applied to variable valve systems of diesel engine, *ISA Transactions*, Vol. 137, 561-573, [doi: 10.1016/j.isatra.2023.01.037](https://doi.org/10.1016/j.isatra.2023.01.037).
- [2] Stavropoulos, G., Violos, J., Tsanakas, S., Leivadreas, A. (2023). Enabling artificial intelligent virtual sensors in an IoT environment, *Sensors*, Vol. 23, No. 3, Article No. 1328, [doi: 10.3390/s23031328](https://doi.org/10.3390/s23031328).
- [3] Hariri-Ardebili, M.A., Mahdavi, G., Nuss, L.K., Lall, U. (2023). The role of artificial intelligence and digital technologies in dam engineering: Narrative review and outlook, *Engineering Applications of Artificial Intelligence*, Vol. 126, Part A, Article No. 106813, [doi: 10.1016/j.engappai.2023.106813](https://doi.org/10.1016/j.engappai.2023.106813).
- [4] Dias, I., Ruan, L., Ranaweera, C., Wong, E. (2023). From 5G to beyond: Passive optical network and multi-access edge computing integration for latency-sensitive applications, *Optical Fiber Technology*, Vol. 75, Article No. 103191, [doi: 10.1016/j.yofte.2022.103191](https://doi.org/10.1016/j.yofte.2022.103191).
- [5] Chen, X., Liu, H., Nikitas, N. (2023). Internal pump leakage detection of the hydraulic systems with highly incomplete flow data, *Advanced Engineering Informatics*, Vol. 56, Article No. 101974, [doi: 10.1016/j.aei.2023.101974](https://doi.org/10.1016/j.aei.2023.101974).
- [6] Kang, F., Wu, Y., Li, J., Li, H. (2021). Dynamic parameter inverse analysis of concrete dams based on Jaya algorithm with Gaussian processes surrogate model, *Advanced Engineering Informatics*, Vol. 49, Article No. 101348, [doi: 10.1016/j.aei.2021.101348](https://doi.org/10.1016/j.aei.2021.101348).
- [7] Pan, Q., Li, Y., Huang, M. (2018). Control-oriented friction modeling of hydraulic actuators based on hysteretic nonlinearity of lubricant film, *Mechatronics*, Vol. 53, 72-84, [doi: 10.1016/j.mechatronics.2018.05.003](https://doi.org/10.1016/j.mechatronics.2018.05.003).
- [8] Feng, H., Qiao, W., Yin, C., Yu, H., Cao, D. (2019). Identification and compensation of non-linear friction for a electro-hydraulic system, *Mechanism and Machine Theory*, Vol. 141, 1-13, [doi: 10.1016/j.mechmachtheory.2019.07.004](https://doi.org/10.1016/j.mechmachtheory.2019.07.004).
- [9] Karanović, V., Jocanović, M., Baloš, S., Knežević, D., Mačužić, I. (2019). Impact of contaminated fluid on the working performances of hydraulic directional control valve, *Journal of Mechanical Engineering*, Vol. 65, No. 3, 139-147, [doi: 10.5545/sv-jme.2018.5856](https://doi.org/10.5545/sv-jme.2018.5856).
- [10] Zhu, Y., Li, G., Tang, S., Wang, R., Su, H., Wang, C. (2022). Acoustic signal-based fault detection of hydraulic piston pump using a particle swarm optimization enhancement CNN, *Applied Acoustics*, Vol. 192, Article No. 108718, [doi: 10.1016/j.apacoust.2022.108718](https://doi.org/10.1016/j.apacoust.2022.108718).
- [11] Ji, X., Ren, Y., Tang, H., Shi, C., Xiang, J. (2020). An intelligent fault diagnosis approach based on Dempster-Shafer theory for hydraulic valves, *Measurement*, Vol. 165, Article No. 108129, [doi: 10.1016/j.measurement.2020.108129](https://doi.org/10.1016/j.measurement.2020.108129).

- [12] Ahmadzadeh, C., Menon, C. (2019). Investigation of regression methods for reduction of errors caused by bending of FSR-based pressure sensing systems used for prosthetic applications, *Sensors*, Vol. 19, No. 24, Article No. 5519, [doi: 10.3390/s19245519](https://doi.org/10.3390/s19245519).
- [13] Shi, M., Hu, W., Li, M., Zhang, J., Song, X., Sun, W. (2023). Ensemble regression based on polynomial regression-based decision tree and its application in the in-situ data of tunnel boring machine, *Mechanical Systems and Signal Processing*, Vol. 188, Article No. 110022, [doi: 10.1016/j.ymssp.2022.110022](https://doi.org/10.1016/j.ymssp.2022.110022).
- [14] Tang, X., Wu, C., Xu, X. (2022). Learning-based nonlinear model predictive controller for hydraulic cylinder control of ship steering system, *Journal of Marine Science and Engineering*, Vol. 10, No. 12, Article No. 2033, [doi: 10.3390/jmse10122033](https://doi.org/10.3390/jmse10122033).
- [15] Tang, S., Zhu, Y., Yuan, S. (2022). An adaptive deep learning model towards fault diagnosis of hydraulic piston pump using pressure signal, *Engineering Failure Analysis*, Vol. 138, Article No. 106300, [doi: 10.1016/j.engfailanal.2022.106300](https://doi.org/10.1016/j.engfailanal.2022.106300).
- [16] Kang, F., Wu, Y., Ma, J., Li, J. (2023). Structural identification of super high arch dams using Gaussian process regression with improved salp swarm algorithm, *Engineering Structures*, Vol. 286, Article No. 116150, [doi: 10.1016/j.engstruct.2023.116150](https://doi.org/10.1016/j.engstruct.2023.116150).
- [17] Sami, I., Ullah, S., Ullah, N., Ro, J.-S. (2021). Sensorless fractional order composite sliding mode control design for wind generation system, *ISA Transactions*, Vol. 111, 275-289, [doi: 10.1016/j.isatra.2020.11.001](https://doi.org/10.1016/j.isatra.2020.11.001).
- [18] Su, W., Ren, W., Sun, H., Liu, C., Lu, X., Hua, Y., Wei, H., Han, J. (2022). Data-based flow rate prediction models for independent metering hydraulic valve, *Energies*, Vol. 15, No. 20, Article No. 7699, [doi: 10.3390/en15207699](https://doi.org/10.3390/en15207699).
- [19] Lei, Y., Jiang, W., Jiang, A., Zhu, Y., Niu, H., Zhang, S. (2019). Fault diagnosis method for hydraulic directional valves integrating PCA and XGBoost, *Processes*, Vol. 7, No. 9, Article No. 589, [doi: 10.3390/pr7090589](https://doi.org/10.3390/pr7090589).
- [20] Zhong, Q., Xu, E., Shi, Y., Jia, T., Ren, Y., Yang, H., Li, Y. (2023). Fault diagnosis of the hydraulic valve using a novel semi-supervised learning method based on multi-sensor information fusion, *Mechanical Systems and Signal Processing*, Vol. 189, Article No. 110093, [doi: 10.1016/j.ymssp.2022.110093](https://doi.org/10.1016/j.ymssp.2022.110093).
- [21] Liu, Y., Li, W., Lin, S., Zhou, X., Ge, Y. (2023). Hydraulic system fault diagnosis of the chain jacks based on multi-source data fusion, *Measurement*, Vol. 217, Article No. 113116, [doi: 10.1016/j.measurement.2023.113116](https://doi.org/10.1016/j.measurement.2023.113116).
- [22] Shweta, R., Sivagnanam, S., Kumar, K.A. (2023). IoT-based deep learning neural network (DLNN) algorithm for voltage stability control and monitoring of solar power generation, *Advances in Production Engineering & Management*, Vol. 18, No. 4, 447-461, [doi: 10.14743/apem2023.4.484](https://doi.org/10.14743/apem2023.4.484).
- [23] Guo, Y., Zeng, Y., Fu, L., Chen, X. (2019). Modeling and experimental study for online measurement of hydraulic cylinder micro leakage based on convolutional neural network, *Sensors*, Vol. 19, No. 9, Article No. 2159, [doi: 10.3390/s19092159](https://doi.org/10.3390/s19092159).
- [24] Jankovič, D., Šimic, M., Herakovič, N. (2024). A data-driven simulation and Gaussian process regression model for hydraulic press condition diagnosis, *Advanced Engineering Informatics*, Vol. 59, Article No. 102276, [doi: 10.1016/j.aei.2023.102276](https://doi.org/10.1016/j.aei.2023.102276).
- [25] Ma, K., Wang, J., Gu, L. (2018). Experimental study on friction of hydraulic cylinder in different sealing systems, *MATEC Web of Conferences*, Vol. 153, Article No. 06012, [doi: 10.1051/mateconf/201815306012](https://doi.org/10.1051/mateconf/201815306012).
- [26] Xin, C., Motz, T., Fuhl, W., Hartel, A., Kasneci, E. (2023). Deep learning-based position detection for hydraulic cylinders using scattering parameters, *Expert Systems with Applications*, Vol. 232, Article No. 120892, [doi: 10.1016/j.eswa.2023.120892](https://doi.org/10.1016/j.eswa.2023.120892).
- [27] Siwulski, T. (2023). Comparative study of the influence of the system architecture on the accuracy of hydraulic cylinder working movements, *Applied Sciences*, Vol. 13, No. 3, Article No. 1594, [doi: 10.3390/app13031594](https://doi.org/10.3390/app13031594).
- [28] Zhou, R., Meng, L., Yuan, X., Qiao, Z. (2022). Research and experimental analysis of hydraulic cylinder position control mechanism based on pressure detection, *Machines*, Vol. 10, No. 1, Article No. 1, [doi: 10.3390/machines10010001](https://doi.org/10.3390/machines10010001).
- [29] Resman, M., Pipan, M., Šimic, M., Herakovič, N. (2019). A new architecture model for smart manufacturing: A performance analysis and comparison with the RAMI 4.0 reference model, *Advances in Production Engineering & Management*, Vol. 14, No. 2, 153-165, [doi: 10.14743/apem2019.2.318](https://doi.org/10.14743/apem2019.2.318).
- [30] Makansi, F., Schmitz, K. (2022). Data-driven condition monitoring of a hydraulic press using supervised learning and neural networks, *Energies*, Vol. 15, No. 17, Article No. 6217, [doi: 10.3390/en15176217](https://doi.org/10.3390/en15176217).
- [31] Dudek, A., Baranowski, J. (2022). Gaussian processes for signal processing and representation in control engineering, *Applied Sciences*, Vol. 12, No. 10, Article No. 4946, [doi: 10.3390/app12104946](https://doi.org/10.3390/app12104946).

Evolutionary game analysis of green innovation in E-commerce closed-loop supply chain WEEE recycling

Ma, R.M.^a, Yao, L.F.^b, Wang, H.^{c,*}

^aSchool of Management, Guangzhou University, Guangzhou 510006, P.R. China

^bSchool of Geography and Tourism, Guangdong University of Finance and Economics, Guangzhou 510320, P.R. China

^cSchool of Finance, Guangdong University of Finance and Economics, Guangzhou 510320, P.R. China

ABSTRACT

The accumulation of waste electrical and electronic equipment (WEEE) has become a critical global issue. E-commerce platforms offer new opportunities for WEEE recycling, making it a subject of interest for researchers. This study focuses on the E-commerce Closed-Loop Supply Chain (E-CLSC) WEEE recycling system, led by remanufacturers, and develops a dual-sided evolutionary game model with remanufacturers and platforms as participants. The model considers the influence of factors such as green innovation, service level, recycling price, and government subsidies. A profit matrix is constructed to analyze the strategic choices of remanufacturers and platforms. Then, this paper conducts a simulation using MATLAB, obtaining data based on the sales and recycling prices of smartphones. Based on evolutionary numerical analysis, the following findings were obtained: (1) Government subsidy policies are formulated based on the required investments for green innovation and service levels, which differ at each stage. (2) The decision of whether remanufacturers engage in green innovation depends largely on the extent to which the technology can reduce remanufacturing costs. They are more inclined to choose green innovation if it can significantly lower costs. (3) Consumer sensitivity to recycling prices also influences the strategic choices of remanufacturers. The more sensitive consumers are to prices, the more waste products remanufacturers can recycle, making green innovation more attractive.

ARTICLE INFO

Keywords:
Recycling;
Waste electrical and electronic equipment (WEEE);
E-commerce;
Closed-loop supply chain;
Green innovation;
Evolutionary game;
Decision-making;
MATLAB

**Corresponding author:*
wanghua1012@gdufe.edu.cn
(Wang, H.)

Article history:
Received 8 January 2024
Revised 17 January 2024
Accepted 7 February 2024



Content from this work may be used under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

With the increasing accumulation of waste electrical and electronic equipment (WEEE), the global environment faces significant challenges [1-5]. According to the Global E-Waste Monitor 2020 report [6], a staggering 53.6 million tons of e-waste was generated worldwide in 2019. China alone produced 10.1 million metric tons of WEEE, making it the largest generator of WEEE globally. The hazardous substances found in WEEE pose a significant environmental and public health threat [7-10], while the valuable resources contained within are often underutilized [11, 12]. Consequently, the collection and recycling of WEEE have become a necessity and urgency [13].

Governments worldwide have responded to WEEE with corresponding legislation [14-16]. The Waste Electrical and Electronic Equipment (WEEE) Directive of the European Union [17],

Regulations on the Administration of the Recycling and Disposal of Waste Electrical Appliances and Electronic Products, and the Law of the People's Republic of China on the Prevention and Control of Solid Waste Pollution [18] aim to reduce the environmental impact of WEEE by promoting the recovery and reuse of valuable materials from discarded products. Additionally, many socially responsible enterprises have played a critical role in promoting the recycling and reuse of WEEE. For example, Apple has launched the "Apple Reuse and Recycling Program" [19]. Alibaba has launched the "credit recovery" program that encourages consumers to recycle their used products by providing credits redeemable for new products.

Remanufacturing, an effective method of environmental protection, has gained attention from governments and enterprises worldwide. It involves the process of taking end-of-life products and restoring them to their original condition, thereby extending their lifespan and reducing the need for new products [20]. Environmental regulations and growing consumer consciousness have made remanufacturing an indispensable aspect for many manufacturers. Leading companies such as Apple, HP, Sony, Huawei, and Lenovo have started incorporating recycled components into their production processes instead of relying solely on new raw materials. This shift towards remanufacturing is driven not only by environmental concerns but also by the significant cost savings of up to 30-50 % that can be achieved [21].

However, the remanufacturing industry faces many challenges, such as the lack of a quality control system for remanufacturing, insufficient investment in technology research and development, and low rates of remanufacturing for old parts. To support green technology innovation, the U.S. government allocated \$2.4 billion to electric vehicle companies [22]. The Chinese government has issued Implementation Plans to further improve the market-oriented green technology innovation system [23]. The rise of third-party online recycling platforms provides new opportunities for the development of reverse supply chains, which allows logistics, information flow, and capital flow to be managed electronically. The development of Internet-based third-party network recycling platforms has created new opportunities for the reverse supply chain, enabling capital flow, information flow, and logistics to be managed electronically. This has led to the formation of an E-commerce closed-loop supply chain (E-CLSC) and provided a new channel for WEEE recycling [24-26].

The recycled WEEE through the E-CLSC is mainly due to product replacement and updates. However, for some consumers (such as mobile phones), these products still have residual and useful value. The platform provides recycling services and has strict recycling testing procedures. After the recycling of WEEE, there are three directions for its reuse: reselling as second-hand products, remanufacturing and dismantling for raw materials. Despite the scale effect and data advantage of platforms, little attention is paid to WEEE that require remanufacturing and disassembly and processing with lower profits. The available WEEE cannot flow smoothly to the remanufacturer for remanufacturing, resulting in instability for the entire E-CLSC dominated by the remanufacturer. Promoting technological innovation by remanufacturers and encouraging platforms to actively participate in the recycling of WEEE for remanufacturing are therefore crucial. It is worth exploring how to incentivize such innovation to ensure that these stakeholders play a more active role in WEEE recycling and remanufacturing.

The main goal of this article is to construct an evolutionary game model consisting of remanufacturers and platforms to understand in detail the influencing factors and evolution process of the E-CLSC system. Section 2 provides a comprehensive review of the relevant literature. Section 3 presents detailed descriptions of the model, including the assumptions made, and focuses on the construction of the evolutionary game model. Section 4 comprises a thorough numerical analysis. Finally, in Section 5, we conclude our findings.

2. Literature review

The literature review focuses on exploring the existing research on green innovation, government subsidies, and the E-CLSC of recycling. We highlight how our research differs from previous studies and outline the specific contributions our work brings to the field.

2.1 Green innovation and government subsidies

Innovation referred to changes in design/production/offering of products or services [27], in response to intense competition, companies worldwide have widely recognized technology innovation as an effective measure [28]. Green innovation could benefit remanufacturing because of products becoming more recyclable and easier to remanufacture [29]. In their study, Reimann *et al.* [30] adopted a game-theoretical framework to investigate how remanufacturing is interconnected with the possibility of decreasing variable remanufacturing costs through process innovation. Lee [31] examined a closed-loop supply chain (CLSC) comprising a manufacturer, a retailer, and a collector. The study highlights the necessity for all supply chain members to engage in green innovation initiatives simultaneously in order to attain a mutually beneficial outcome within the CLSC. In order to encourage enterprises to engage in green innovation, governments had adopted subsidies as a means to incentivize innovative behaviors.

In traditional recycling systems, it has been widely acknowledged that government subsidies play a significant role in boosting the participation rate and enhancing the effectiveness of recycling outcomes [32]. The research conducted by Wang *et al.* [33] indicated that the implementation of green insurance subsidies and government subsidies can serve as catalysts for driving innovation within enterprises. He *et al.* [21] examined the optimal channel structure, pricing decisions of manufacturers, and the government's optimal subsidy level across three different channel structures. In a similar vein, Liu *et al.* [34] investigated the impact of government subsidies on the profits of supply chain members. Wan and Hong [35] also showed that government subsidies promote consumption and increase recycling, benefiting the entire recycling chain. According to Aldieri *et al.* [36], the implementation of a green innovation subsidy policy can assist enterprises in adopting green technology and enhancing their employment levels. Huang *et al.* [37] examined the different modes of green credit, manufacturing subsidy, and sales subsidy in a supply chain involving a green manufacturer with limited capital. By conducting a sensitivity analysis on the green degree of recycled products, In the study conducted by Guo *et al.* [38], they provided empirical evidence that supports the adoption of a secondary subsidy strategy by the government for low-green products. This particular strategy was found to serve as an additional incentive for enterprises, effectively encouraging and promoting their endeavors in recycling and remanufacturing. According to Zhou *et al.* [39], the effectiveness of a subsidy greatly depended on its amount. Therefore, in order to ensure successful implementation of policies, the government should establish a scientific and dynamic mechanism for adjusting the subsidy amount. The research of Bai *et al.* [40] aimed to examine the influence of government subsidies on various online channel strategies and how consumer preference for high-quality services factors into this equation. However, existing literature lacks analysis of participants' decision-making processes in an E-CLSC system.

2.2 E-CLSC of recycling

The internet economy has been highly successful in both the forward sales channels and the reverse channels for WEEE [41, 42]. The traditional recycling systems face issues such as a complex chain, ambiguous procedures, and unpredictable costs, but the online recycling industry has made significant progress in addressing these concerns [43].

Various studies have been conducted on the E-CLSC of recycling, focusing on several aspects. Firstly, in terms of recycling channel selection, Feng *et al.* [44] identified four recycling modes and found that the dual-channel approach, combining online and offline channels, outperforms single-channel approaches. Li *et al.* [45] conducted a comparative analysis of online recycling, offline recycling, and mixed recycling models, highlighting the advantages for remanufacturers in online recycling but potential drawbacks for recyclers. Wang *et al.* [46] compared different

decisions in the forward and reverse channels, including wholesale vs. direct sales and entrusted recycling vs. direct recycling. Secondly, regarding sales mode selection, Jia and Li [47] proposed that e-retailers should focus on selling new products, while online platforms should prioritize remanufactured products. Zhang and Hou [48] discovered that when online retailers own their brands, the preferred sales mode of manufacturers and e-retailers tends to be opposite, even with asymmetric cost information. Lastly, studies on recycling prices have also been conducted. Matsui [49] examined the timing issue of recycling prices in both online and offline dual channels. The research indicated that offline recycling price should have an advantage over online recycling price at the time of establishment. However, these studies have often neglected the impact of the platform's recycling service level, which plays a significant role in determining the amount of WEEE that can be recycled.

Although extensive research has been conducted on the E-CLSC, there has been limited focus on the dynamic evolution of strategy selection between remanufacturers and platforms in the recycling of WEEE. As a result, we propose the assumption of bounded rationality among remanufacturers and platforms and aim to examine the significant role of government subsidies policy in guiding WEEE recycling practices. This study aims to fill the gap in the existing literature by shedding light on the dynamics of strategy selection and the impact of government interventions in the context of WEEE recycling.

3. Construction of the evolutionary game model

This study investigates a closed-loop e-commerce recycling supply chain, predominantly led by remanufacturers. The remanufacturers collaborate with e-commerce recycling platforms to retrieve discarded products. Consumers, after browsing the recycling information on the platform, provide relevant information about their discarded products. The platform evaluates the value of the discarded products and provides feedback to the consumers. Consumers have the option to either ship the products or opting for doorstep pickup services. The platform delivers the quality-checked waste products to the remanufacturers, who then engage in remanufacturing and sales activities. The remanufacturers reduce remanufacturing costs through green innovations while also minimizing environmental pollution. The E-CLSC structure in which the remanufacturer is dominate of this study is illustrated in Fig. 1.

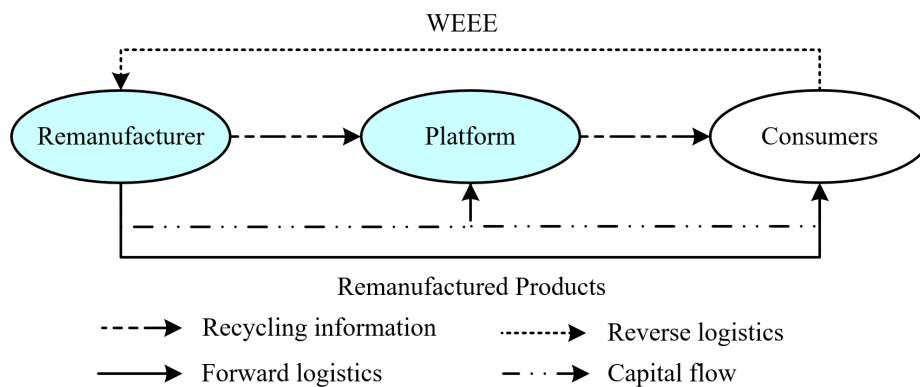


Fig. 1 Structure of E-CLSC model is led by remanufacturer

3.1 Assumptions and notations

To meet the requirements of the evolutionary game theory and our research problems, we make the following assumptions about the model we have constructed in this study. The definitions and explanations of the relevant parameters are presented in Table 1.

Assumption 1. In our model, we incorporate the notion of bounded rationality, assuming that each player has limited rationality. Due to information asymmetry, remanufacturer and platform cannot accurately determine the recycling decisions for WEEE that would maximize their indi-

vidual interests. Instead, they possess learning and imitating abilities, allowing them to adapt their strategies based on past experiences and observations.

Assumption 2. This paper constructs an evolutionary game model containing two players: remanufacturer and platform. The remanufacturer makes a strategic choice between two options: Green Innovation (GI) and No Green Innovation (NGI), with respective probabilities of x and $1 - x$. Similarly, the platform selects between two strategies: Actively Participate (AP) and Inactively Participate (IP), with probabilities of y and $1 - y$.

Assumption 3. In the case where the remanufacturer opts for the No Green Innovation (NGI) strategy, we denote the remanufacturing cost as c . However, it is worth noting that green innovation undertaken by the remanufacturer has been shown to effectively decrease the production cost of remanufactured products [29]. The remanufactured cost When the remanufacturer chooses Green Innovation (GI) strategy is $c_1 = (c - re)$, e denotes green innovation effort level, r is the marginal effect of green innovation effort on remanufactured products' production cost. When remanufacturer chooses the GI strategy, an additional cost in green innovation is required. Based on literatures [50-52], the cost can be defined as $K = 1/2me^2$. m is the cost sensitivity of green innovation effort level.

Assumption 4. The quantity of WEEE recycled is influenced by the recycling price and the recycling service level. Following the assumptions made by Giovanni and Zaccour [53], we assume that the function describing the recycling quantity q , can be expressed as $q = \beta p_c + \theta s$.

Assumption 5. The higher the level of service provided by a platform, the higher the cost involved. Assuming the platform chooses to actively participate in recycling, the additional cost required is represented by $L=1/2ns^2$ [54]. n is the cost sensitivity of service level.

Table 1 Summary of parameters

Symbol	Description
p	The selling price of the remanufactured product
p_c	The recycling price of a unit of WEEE. Decision variable
μ	The unit commission charged by platform to remanufacturer
r	The marginal effect of green innovation effort on remanufactured cost
c	Unit reproduced costs when green innovations are not implemented
e	The level of green innovation of the remanufacturer, $e > 0$, Exogenous variables
β	Sensitivity of consumer recycling price
θ	Sensitivity of consumer recycling service levels.
s	The level of recycling service
d	The inherent subsidy provided by the government to platforms on a per-unit
f	The inherent subsidy provided by the government to remanufacturer on a per-unit
m	Cost sensitivity of green innovation effort level
n	Cost sensitivity of service level
g	Government subsidies for green innovation in remanufacturers
w	Government subsidies for platforms to improve their service levels
π_{ri}	Profits of remanufacturer
π_{pi}	Profits of platform. $i = 1,2,3,4$: four strategies used by the remanufacturer and the platform

3.2 Model construction

By considering various combinations of strategies between the remanufacturer and the platform, we have determined the benefits matrix, as illustrated in Table 2.

(1) Strategy combination 1: (GI, AP)

In this scenario, remanufacturer chooses to engage in green innovation, while platform opts to actively participate in recycling. The profit functions of the remanufacturer and the platform can be formulated as follows:

$$\pi_{r1} = (p - c_1 - \mu + f - p_c)q - (1 - g)K \quad (1)$$

$$\pi_{p1} = (\mu + d)q - (1 - w)L \quad (2)$$

We thus derive:

$$\begin{cases} p_{c1} = \frac{\beta(p - c_1 - \mu + f) - \theta s}{2\beta} \\ q_1 = \frac{\beta(p - c_1 - \mu + f) + \theta s}{2} \\ \pi_{r1} = \frac{[\beta(p - c_1 - \mu + f) + \theta s]^2}{4\beta} - (1 - g)K \\ \pi_{p1} = (\mu + d) \frac{\beta(p - c_1 - \mu + f) + \theta s}{2} - (1 - w)L \end{cases} \quad (3)$$

(2) Strategy combination 2: (GI, IP)

In this scenario, remanufacturer chooses to engage in green innovation, while platform opts to inactively participate in recycling. At this point, the platform does not need to pay more for service levels, for ease of calculation, we let $s = 0, w = 0$. We thus derive:

$$\begin{cases} p_{c2} = \frac{p - c_1 - \mu + f}{2} \\ q_2 = \frac{\beta(p - c_1 - \mu + f)}{2} \\ \pi_{r2} = \frac{\beta(p - c_1 - \mu + f)^2}{4} - (1 - g)K \\ \pi_{p2} = (\mu + d) \frac{\beta(p - c_1 - \mu + f)}{2} \end{cases} \quad (4)$$

(3) Strategy combination 3: (NGI, AP)

In this scenario, remanufacturer makes no effort with regard to green innovation, while platform opts to actively participate in recycling. At this point, the remanufacturer does not need to pay more for green innovation, for ease of calculation, let $e = 0, g = 0$. We can obtain:

$$\begin{cases} p_{c3} = \frac{\beta(p - c - \mu + f) - \theta s}{2\beta} \\ q_3 = \frac{\beta(p - c - \mu + f) + \theta s}{2} \\ \pi_{r3} = \frac{[\beta(p - c - \mu + f) + \theta s]^2}{4\beta} \\ \pi_{p3} = (\mu + d) \frac{\beta(p - c - \mu + f) + \theta s}{2} - (1 - w)L \end{cases} \quad (5)$$

(4) Strategy combination 4: (NGI, IP)

In this scenario, remanufacturer makes no effort with regard to green innovation, and platform opts to inactively participate in recycling. We can get:

$$\begin{cases} p_{c4} = \frac{p - c - \mu + f}{2} \\ q_4 = \frac{\beta(p - c - \mu + f)}{2} \\ \pi_{r4} = \frac{\beta(p - c - \mu + f)^2}{4} \\ \pi_{p4} = (\mu + d) \frac{\beta(p - c - \mu + f)}{2} \end{cases} \quad (6)$$

Then, the benefit matrix can be obtained in Table 2.

Table 2 Benefit matrix of remanufacturer and platform

Remanufacturer	Platform	
	AP(y)	IP($1 - y$)
	GI(x)	(π_{r1}, π_{p1})
NGI($1 - x$)	(π_{r3}, π_{p3})	(π_{r4}, π_{p4})

Because $q_i > 0, \forall i$, we can obtain $p - c - \mu + f > 0, \pi_{r1} > \pi_{r2}, \pi_{r3} > \pi_{r4}, \pi_{p1} > \pi_{p3}, \pi_{p2} > \pi_{p4}$.

3.3 Evolutionary process

Combined with the benefit matrix in Table 2, the expected benefits for remanufacturer to adopt the GI strategy are

$$U_{11} = y\pi_{r1} + (1 - y)\pi_{r2} \quad (7)$$

The expected benefits for remanufacturer to adopt the NGI strategy are

$$U_{12} = y\pi_{r3} + (1 - y)\pi_{r4} \quad (8)$$

Average expected benefits of remanufacturer are

$$\bar{U}_1 = xU_{11} + (1 - x)U_{12} \quad (9)$$

The expected benefits for platform to adopt the AP strategy are

$$U_{21} = x\pi_{p1} + (1 - x)\pi_{p3} \quad (10)$$

The expected benefits for platform to adopt the IP strategy are

$$U_{22} = x\pi_{p2} + (1 - x)\pi_{p4} \quad (11)$$

Average expected benefits of platform are

$$\bar{U}_2 = yU_{21} + (1 - y)U_{22} \quad (12)$$

Hence, we can derive the replicated dynamic equations for the remanufacturer and the platform as follows:

$$\begin{cases} F(x) = \frac{dx}{dt} = x(U_{11} - \bar{U}_1) = x(1 - x)[y(\pi_{r1} - \pi_{r3}) + (1 - y)(\pi_{r2} - \pi_{r4})] \\ H(y) = \frac{dy}{dt} = y(U_{21} - \bar{U}_2) = y(1 - y)[x(\pi_{p1} - \pi_{p2}) + (1 - x)(\pi_{p3} - \pi_{p4})] \end{cases} \quad (13)$$

The dynamic system replication equation (13) describes the evolution of the remanufacturer and e-commerce platform selection strategy system, according to the principle of differential equation stability, when the replication dynamic equation on both sides of the game is equal to zero, the system tends to a steady state. Let $F(x) = 0, H(y) = 0$, because $\pi_{p1} - \pi_{p2} - \pi_{p3} + \pi_{p4} = 0$, we can obtain four evolutionary equilibria (EE), which are (0,0), (0,1), (1,0), (1,1).

3.4 Evolutionary equilibrium stability analysis

Based on the work of Friedman [55], we can determine the stability of equilibrium points by utilizing the Jacobian matrix of the dynamic system. The Jacobian matrix J is as follows:

$$J = \begin{bmatrix} \frac{\partial F(x)}{\partial x} & \frac{\partial F(x)}{\partial y} \\ \frac{\partial H(y)}{\partial x} & \frac{\partial H(y)}{\partial y} \end{bmatrix} = \begin{bmatrix} j_{11} & j_{12} \\ j_{21} & j_{22} \end{bmatrix}$$

where,

$$\begin{aligned}
 j_{11} &= (1 - 2x)[y(\pi_{r1} - \pi_{r3}) + (1 - y)(\pi_{r2} - \pi_{r4})] \\
 j_{12} &= x(1 - x)(\pi_{r1} - \pi_{r3} - \pi_{r2} + \pi_{r4}) \\
 j_{21} &= y(1 - y)(\pi_{p1} - \pi_{p2} - \pi_{p3} + \pi_{p4}) \\
 j_{22} &= (1 - 2y)[x(\pi_{p1} - \pi_{p2}) + (1 - x)(\pi_{p3} - \pi_{p4})]
 \end{aligned}$$

To evaluate the stability of each equilibrium point, we can calculate the determinant and trace of matrix J . The stability of an equilibrium point is determined based on the signs of its determinant and trace. In particular, if the determinant is positive and the trace is negative, the equilibrium point demonstrates local stability. This signifies that the equilibrium point represents an evolutionary stable strategy (ESS).

According to the method of local stability analysis of Jacobian matrix, stability analysis was performed on four equilibrium points. Through calculation, the $\det J$ and $tr J$ corresponding to the four replication dynamic equilibrium points can be obtained, and the local asymptotic stability analysis results are obtained according to the results of $\det J$ and $tr J$, the replication dynamic mechanism to find the system evolution stability strategy conditions, as depicted in Table 3.

Table 3 $\det J$ and $tr J$ of four equilibrium points

EE	$\det J$	$tr J$
(1,1)	$(\pi_{r1} - \pi_{r3})(\pi_{p1} - \pi_{p2})$	$-(\pi_{r1} - \pi_{r3}) - (\pi_{p1} - \pi_{p2})$
(1,0)	$-(\pi_{r2} - \pi_{r4})(\pi_{p3} - \pi_{p4})$	$-(\pi_{r2} - \pi_{r4}) + (\pi_{p3} - \pi_{p4})$
(0,1)	$-(\pi_{r1} - \pi_{r3})(\pi_{p3} - \pi_{p4})$	$(\pi_{r1} - \pi_{r3}) - (\pi_{p3} - \pi_{p4})$
(0,0)	$(\pi_{r2} - \pi_{r4})(\pi_{p3} - \pi_{p4})$	$(\pi_{r2} - \pi_{r4}) + (\pi_{p3} - \pi_{p4})$

For equilibrium points where the determinant ($\det J$) of the Jacobian matrix is greater than 0 and the trace ($tr J$) is less than 0, we classify them as partial equilibrium stability points. Due to the constraints of length, it is unnecessary to analyze all Evolutionary Stable Strategies (ESSs). In the context of this study on the manufacturer-led E-CLSC system, our main focus is on how to promote green innovation in remanufacturing and encourage e-commerce platforms to actively participate in WEEE recycling through government subsidies. In other words, our focus is on how to adjust subsidy policies in order to encourage remanufacturers and platforms to lean towards choosing the combined strategy (GI, AP). Therefore, we will focus on discussing how to make the stable points of the system tend towards (1,1).

When (1,1) is the partial equilibrium stability point, we can obtain that $(\pi_{r1} - \pi_{r3})(\pi_{p1} - \pi_{p2}) > 0$ and $-(\pi_{r1} - \pi_{r3}) - (\pi_{p1} - \pi_{p2}) < 0$. According to the two conditions, we can derive that $\pi_{r1} - \pi_{r3} > 0$ and $\pi_{p1} - \pi_{p2} > 0$.

Combined with Equations (1)-(6), it is obtained

$$\begin{cases}
 f > \frac{2me(1 - g) - 2r\theta s - \beta r^2 e - 2\beta r(p - c - \mu)}{2\beta r} \\
 d > \frac{ns(1 - w) - \theta\mu}{\theta}
 \end{cases} \tag{14}$$

4. Simulation

In this section, we utilize MATLAB R2016a software to perform a numerical study that aims to validate the evolutionary game model and investigate the modifications in the behavioral strategies of the remanufacturer and the platform through sensitivity analysis of the model parameters. The parameter values used in the study are predominantly based on the sales price of Huawei smartphones in China in 2021. It is noted that Huawei sold smartphones at an average price of 4 thousand yuan per unit during that year. The manufacturing cost accounts for approximately 50 %, which is around 2 thousand yuan/unit. We assume the commission paid to platform is 0.1 per unit, and the green innovation effort level and service effort level are both 0.1. According to the "Several Measures of Beijing Municipality on Further Improving the Market-oriented Green Technology Innovation System" issued by the Beijing Municipal Development

and Reform Commission, the proportion of government subsidies for enterprise technological innovation generally does not exceed 30 % of the total investment [56], i.e. $g = 0.3$, $w = 0.3$.

(1) The influence of green innovation on cost sensitivity

Through the above analysis, the parameters in this situation are set as follows: $p = 4$, $c = 2$, $\mu = 0.5$, $e = 0.1$, $s = 0.1$, $r = 10$, $\beta = 0.8$, $\theta = 0.8$, $f = 0.5$, $n = 8$, $g = 0.3$, $w = 0.3$, $d = 0.1$. From Fig. 2, we can see that the choice of green innovation by remanufacturers varies with changes in the investment cost of green innovation. According to the theory of industrial life cycle [57], China's WEEE recycling industry can be classified into three distinct stages: the initial stage, intermediate stage, and mature stage. Each stage represents a different phase of development and progress in the WEEE recycling industry in China [18]. Similarly, the development of green innovation for remanufacturers can also be divided into three stages. In the early stage of green innovation, technology is in its infancy and requires significant investment in laboratory construction and recruitment of personnel. The level of investment required to improve the unit's green innovation efforts is substantial, denoted as $m = 300$ million. In the intermediate stage of green innovation, where the initial infrastructure is already established ($m = 200$ million), and in the mature stage of green technological innovation, where the innovation foundation and environment are well-developed ($m = 100$ million), the required investment for green technology innovation decreases. Fig. 2 illustrates that the government's subsidy ratio for green innovation by remanufacturers remains consistent at 0.5 across these three stages. However, the choices made by remanufacturers differ, reflecting the current attitude of Chinese enterprises towards green technology innovation. It is observed that most companies tend to adopt green innovation technologies into product production when they have already gained market presence and reached a mature stage.

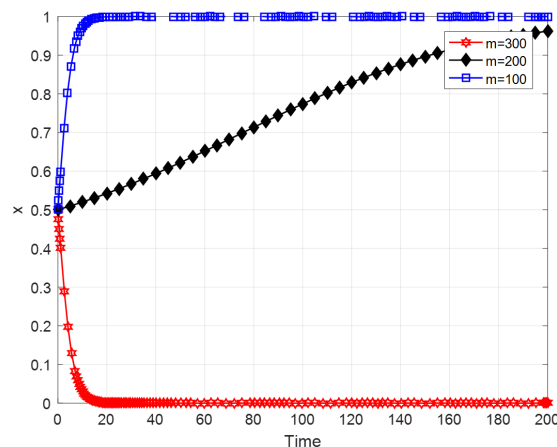


Fig. 2 The influence of green innovation on cost sensitivity

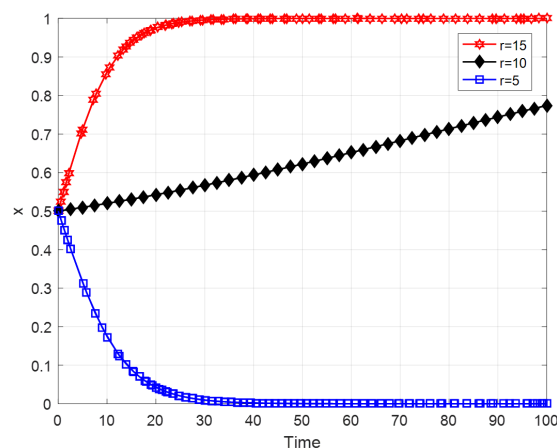


Fig. 3 Green innovation on remanufactured cost sensitivity

(2) The influence of green innovation on remanufactured cost sensitivity

Keeping other parameters unchanged, let's consider the case where $m = 200$ and examine the impact of the degree of cost reduction resulting from green innovation on the decision-making of remanufacturers. By setting r to values of 5, 10, and 15, we can observe Fig. 3. In this scenario, remanufacturers evaluate the adoption of green innovation based on its potential to reduce production costs per unit. With a fixed coefficient for the cost of investing in green innovation, they assess whether the implementation of green innovation can effectively lower their overall remanufactured costs. The higher the cost reduction achieved per unit of green innovation effort, the more likely remanufacturers are to choose green innovation.

(3) The influence of government subsidies

Keeping other parameters unchanged, let $m = 200$ and substitute the respective parameters into equation (14). We can calculate that when $f > 0.4$ and $d > 0.2$, which means the government provides a production subsidy of 0.4 units for each green innovation product, and the platform also receives a recycling subsidy of 0.2 units for each product recycled, the system's stable point will tend toward (1, 1). As shown in Fig. 4(a), during the intermediate stage of technological innovation, when f takes a value of 0.4, over time, remanufacturers will ultimately lean towards choosing NGI (Non-Green Innovation). However, when f takes a value of 0.5, although the curve is relatively flat, with the passage of time, remanufacturers will still choose GI (Green Innovation) after a longer period. Once a technology enters the mature stage ($m = 100$), where the required cost for each unit of green innovation significantly decreases, as depicted in Fig. 4(a), even if the government no longer provides production subsidies, the remanufacturers' GI selection curve will quickly converge to 1.

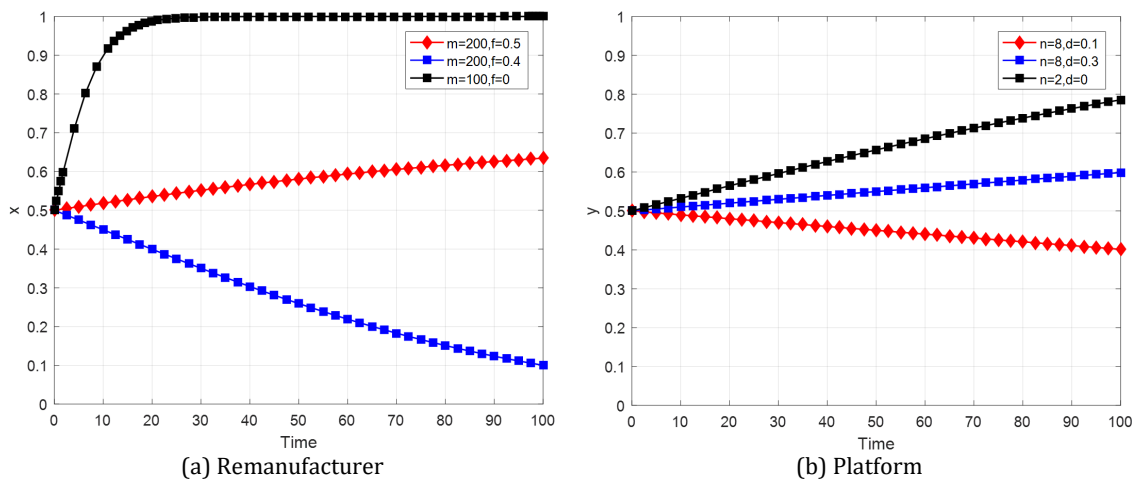


Fig. 4 The influence of government subsidies on remanufacturer and platform selection

Similarly, for the platform, as shown in Fig. 4(b), during the initial stage of product recycling, there is a need for increased investment to enhance the service level ($n = 8$). When d takes a value of 0.1, over time, the platform will ultimately lean towards choosing the IP (Inactively Participate) strategy. However, when d takes a value of 0.3, after a longer period, the platform will eventually choose the AP (Actively Participate) strategy. Once the platform's recycling approach becomes well-established among consumers, reaching the mature stage of platform recycling ($n = 2$), the platform no longer requires substantial investment to attract consumers to choose the e-commerce platform for recycling. At this stage, even if the government no longer provides recycling subsidies ($d = 0$), the platform will actively participate in the platform recycling model.

Similarly to the previous discussion, we can also discuss the government's subsidies for technological innovation and platform services. As shown in the Fig. 5, during the early to mid-term of technological investment ($m = 200$), when $g = 0.3$, remanufacturers will choose the GI strategy. However, if the government reduces subsidies for technological investment, i.e., $g = 0.2$, remanufacturers will opt for the NGI strategy. As technology continues to mature ($m =$

150), despite the decrease in technological investment subsidies ($g = 0.2$), remanufacturers will still choose the GI strategy. The same logic applies to the platform's strategy selection.

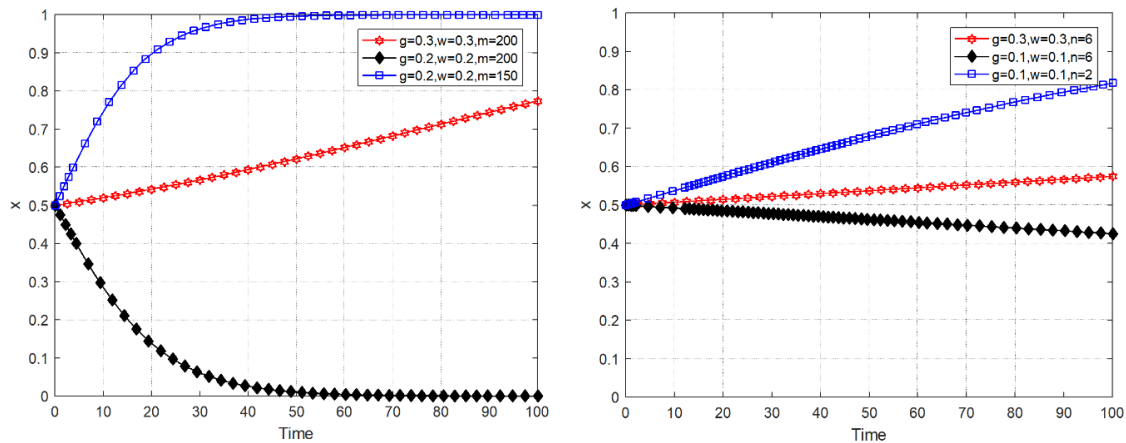


Fig. 5 The influence of government technology subsidy ratio

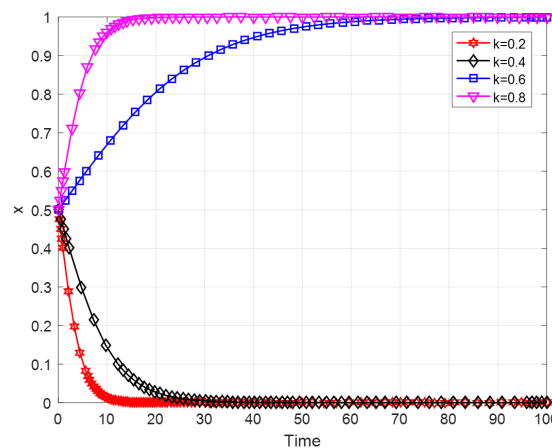


Fig. 6 The influence of recycling price coefficient

(4) The influence of recycling price coefficient

Keeping other parameters unchanged, let $m = 200$, $n = 8$, $r = 10$, $f = 0.5$. By setting different values for k , namely 0.2, 0.4, 0.6, and 0.8, we can obtain Fig. 6. From Fig. 5, it can be observed that the recycling price coefficient significantly influences the decisions of remanufacturers. As the value of k increases, the remanufacturers' choices shift from NGI to GI. Moreover, as the value of k becomes larger, the probability curve of remanufacturers choosing the GI strategy converges at a faster rate.

5. Conclusion

This study focuses on the research of the E-CLSC WEEE recycling system, with remanufacturers as the main actors. This study adopts an evolutionary game approach to analyze the decision-making processes of remanufacturers with regards to green innovation and the platform's selection of actively participating in recycling. By employing evolutionary stability strategies and stable conditions, we derive the profit matrix and evolution path for both sides. Furthermore, we provide a concise and detailed analysis of how the main parameters affect these decisions. Overall, this research sheds light on the dynamics of decision-making in the context of green innovation and active recycling within an evolutionary game framework. In the model developed in this study, we consider the impact of green innovation on remanufactured costs. This allows us to analyze how the adoption of green innovation affects the overall costs involved in the remanu-

facturing process. Additionally, the remanufacturers' fixed capital investment in green innovation is considered. Furthermore, based on the practical application of recycling on e-commerce platforms, the price of recycling and the platform's service level are factors that affect the quantity of recycling. This study introduces a functional relationship between the quantity of recycling, price, and service level. Building upon these conditions, the study further investigates how government subsidies will impact the strategic choices of both decision-making parties. Based on the findings of the research analysis, the following management suggestions are derived:

- Both the development of green innovation and platform recycling models can be categorized into three distinct stages. In the initial stage, significant capital investment is required for the development of both green innovation and platform recycling models. During this stage, the government needs to provide substantial subsidies to remanufacturers for unit product remanufacturing and green innovation investments. Only then will remanufacturers choose to adopt green innovation (GI). Similarly, in the initial stage of the platform recycling model, the platform needs to invest a considerable amount of funds in promotion, channel development, etc. Without sufficient government subsidies, the platform is often reluctant to actively participate (AP) in recycling. However, in the mature stage, green innovation has already established a certain technological foundation, and the required investments in green innovation and platform services have significantly decreased. During this stage, even without government subsidies, remanufacturers will choose green innovation (GI), and the platform will actively engage in recycling. Therefore, when formulating green innovation subsidy policies, the government needs to consider the current innovation environment, innovation foundation, and adjust subsidy policies accordingly during different stages of development.
- The degree to which green innovation technology reduces remanufacturing costs is a crucial factor in the decision-making process of remanufacturers regarding whether to adopt green innovation. If the green innovation technology offers only minimal reduction in remanufacturing costs, regardless of the amount of investment required, remanufacturers will not choose the green innovation strategy.
- The sensitivity of consumers to recycling prices has a significant impact on the strategy choices of remanufacturers in an E-CLSC system. When consumers have a low sensitivity to recycling prices, a majority of them may choose to discard or keep their used products rather than sell them to the platform. As a consequence, this could lead to a relatively low volume of waste products being recycled within the system, which in turn may limit the profitability of remanufacturers from reselling these products. In this situation, remanufacturers are unlikely to choose green innovation. However, when consumers have a higher sensitivity to recycling prices, it indicates that they are attracted by the prices offered for their used products and are willing to sell them. This results in a larger quantity of waste products being collected by remanufacturers, who can then profit from selling them after undergoing remanufacturing. In this scenario, remanufacturers are more likely to choose green innovation.
- The model proposed in this paper can be applied to various scenarios in different countries and regions. However, the levels of government subsidies, technological development, product sales prices, and recycling prices may vary among different countries and regions. Therefore, the analysis results will also vary with changes in these parameters.
- It is important to acknowledge certain limitations in our study. Firstly, we did not consider the influence and role of other stakeholders within the closed-loop supply chain on the governance mechanism. For instance, consumer environmental awareness can greatly influence the amount of WEEE recycled, thereby influencing the strategies pursued by remanufacturers. Secondly, this paper focused solely on government subsidies as a variable to examine the decision-making mechanism of remanufacturers and platforms, without considering the government as an active participant in the system. Thirdly, we did

not consider the impact of development level on the government subsidy intensity. Therefore, future research should aim to develop new approaches that encompass the remanufacturer, platform, and consumer within a comprehensive framework.

Acknowledgement

This work was supported by Humanities and social sciences fund of the Ministry of Education (Grant No.: 19YJC630119) and Guangdong Province Philosophy and Social Sciences Planning Project (Grant No.: GD22XGL25).

References

- [1] Jian, L., Zhen, W., Bao, J., Taebok, K. (2017). Coordination strategies in a three-echelon reverse supply chain for economic and social benefit, *Applied Mathematical Modelling*, Vol. 49, 599-611, doi: [10.1016/j.apm.2017.04.031](https://doi.org/10.1016/j.apm.2017.04.031).
- [2] Rahmanda, B., Njatrijani, R., Fadillah, R. (2023). Environmental policy in managing e-waste recycling: Promoting a clean environment in public policy, *International Journal of Sustainable Development and Planning*, Vol. 18, No. 1, 121-126, doi: [10.18280/ijstdp.180112](https://doi.org/10.18280/ijstdp.180112).
- [3] Rifa'i, A., Raharja, S.J., Rivani, R., Purbasari, R. (2023). One decade research in the field of business ecosystem: A bibliometric analysis, *International Journal of Sustainable Development and Planning*, Vol. 18, No. 4, 1219-1226, doi: [10.18280/ijstdp.180426](https://doi.org/10.18280/ijstdp.180426).
- [4] de Oliveira, B.O.S., de Medeiros, G.A., Paes, M.X., ManCini, S.D. (2021). Integrated municipal and solid waste management in the amazon: Addressing barriers and challenges in using the Delphi method, *International Journal of Environmental Impacts*, Vol. 4, No. 1, 49-61, doi: [10.2495/EI-V4-N1-49-61](https://doi.org/10.2495/EI-V4-N1-49-61).
- [5] Oyewale, J.A., Tartibu, L.K., Okokpujie, I.P. (2023). A review and bibliometric analysis of sorting and recycling of plastic wastes, *International Journal of Design & Nature and Ecodynamics*, Vol. 18, No. 1, 63-74, doi: [10.18280/ijidne.180107](https://doi.org/10.18280/ijidne.180107).
- [6] Forti, V., Baldé, C.P., Kuehr, R., Bel, G. (2020). The global e-waste monitor 2020: Quantities, flows and the circular economy potential, Global e-waste key statistics, International Telecommunication Union (ITU) & International Solid Waste Association (ISWA), Bonn/Geneva/Rotterdam, from https://collections.unu.edu/eserv/UNU:7737/GEM_2020_def_july1.pdf, accessed January 3, 2024.
- [7] Owusu-Sekyere, K., Batteiger, A., Afoblikame, R., Hafner, G., Kranert, M. (2022). Assessing data in the informal e-waste sector: The Agboghloshie Scrapyard, *Waste Management*, Vol. 139, 158-167, doi: [10.1016/j.wasman.2021.12.026](https://doi.org/10.1016/j.wasman.2021.12.026).
- [8] Ling, M., Xu, L., Chu, X. (2023). Heterogeneous effects of other-regarding interventions on household recycling: A field experimental study, *Journal of Environmental Management*, Vol. 329, Article No. 117102, doi: [10.1016/j.jenvman.2022.117102](https://doi.org/10.1016/j.jenvman.2022.117102).
- [9] Wang, Y.L., Yang, L., Chen, J.H., Li, P. (2023). Supply chain game analysis based on mean-variance and price risk aversion under different power structures, *Advances in Production Engineering & Management*, Vol. 18, No. 1, 104-115, doi: [10.14743/apem2023.1.460](https://doi.org/10.14743/apem2023.1.460).
- [10] Xu, G.Y., Liu, H., Duan, H.W. (2022). Cause-related marketing strategy in a supply chain: A theoretical analysis and a case study, *Advances in Production Engineering & Management*, Vol. 17, No. 4, 479-493, doi: [10.14743/apem2022.4.450](https://doi.org/10.14743/apem2022.4.450).
- [11] Pathak, P., Srivastava, R.R., Ojasvi. (2017). Assessment of legislation and practices for the sustainable management of waste electrical and electronic equipment in India, *Renewable and Sustainable Energy Reviews*, Vol. 78, 220-232, doi: [10.1016/j.rser.2017.04.062](https://doi.org/10.1016/j.rser.2017.04.062).
- [12] Fang, I.W., Lin, W.-T. (2021). A multi-objective optimal decision model for a green closed-loop supply chain under uncertainty: A real industrial case study, *Advances in Production Engineering & Management*, Vol. 16, No. 2, 161-172, doi: [10.14743/apem2021.2.391](https://doi.org/10.14743/apem2021.2.391).
- [13] Wang, Q., Li, B., Chen, B., Cheng, Y., Wang, Z. (2021). Implication of take-back and carbon emission capacity regulations on remanufacturing in a competitive market, *Journal of Cleaner Production*, Vol. 325, Article No. 129231, doi: [10.1016/j.jclepro.2021.129231](https://doi.org/10.1016/j.jclepro.2021.129231).
- [14] Dhanorkar, S., Muthulingam, S. (2020). Do e-waste laws create behavioral spillovers? Quasi-experimental evidence from California, *Production and Operations Management*, Vol. 29, No. 7, 1738-1766, doi: [10.1111/poms.13182](https://doi.org/10.1111/poms.13182).
- [15] Shittu, O.S., Williams, I.D., Shaw, P.J. (2021). Global E-waste management: Can WEEE make a difference? A review of e-waste trends, legislation, contemporary issues and future challenges, *Waste Management*, Vol. 120, 549-563, doi: [10.1016/j.wasman.2020.10.016](https://doi.org/10.1016/j.wasman.2020.10.016).
- [16] Yang, L., Guo, J., Zhang, C., Wan, P., Li, H. (2022). Optimization of location-routing for the waste household appliances recycling logistics under the uncertain condition, *Tehnički Vjesnik – Technical Gazette*, Vol. 29, No. 1, 127-138, doi: [10.17559/TV-20210623085943](https://doi.org/10.17559/TV-20210623085943).
- [17] Salhofer, S., Steuer, B., Ramusch, R., Beigl, P. (2016). WEEE management in Europe and China - A comparison, *Waste Management*, Vol. 57, 27-35, doi: [10.1016/j.wasman.2015.11.014](https://doi.org/10.1016/j.wasman.2015.11.014).
- [18] Li, B., Wang, Q., Chen, B., Sun, T., Wang, Z., Cheng, Y. (2022). Tripartite evolutionary game analysis of governance mechanism in Chinese WEEE recycling industry, *Computers & Industrial Engineering*, Vol. 167, Article No.

- 108045, doi: [10.1016/j.cie.2022.108045](https://doi.org/10.1016/j.cie.2022.108045).
- [19] Liu, J., Zhai, X., Chen, L. (2019). Optimal pricing strategy under trade-in program in the presence of strategic consumers, *Omega*, Vol. 84, 1-17, doi: [10.1016/j.omega.2018.03.005](https://doi.org/10.1016/j.omega.2018.03.005).
- [20] Yang, R., Tang, W., Zhang, J. (2021). Technology improvement strategy for green products under competition: The role of government subsidy, *European Journal of Operational Research*, Vol. 289, No. 2, 553-568, doi: [10.1016/j.ejor.2020.07.030](https://doi.org/10.1016/j.ejor.2020.07.030).
- [21] He, P., He, Y., Xu, H. (2019). Channel structure and pricing in a dual-channel closed-loop supply chain with government subsidy, *International Journal of Production Economics*, Vol. 213, 108-123, doi: [10.1016/j.ijpe.2019.03.013](https://doi.org/10.1016/j.ijpe.2019.03.013).
- [22] Gong, H., Wang, M.Q., Wang, H. (2013). New energy vehicles in China: Policies, demonstration, and progress, *Mitigation and Adaptation Strategies for Global Change*, Vol. 18, 207-228, doi: [10.1007/s11027-012-9358-6](https://doi.org/10.1007/s11027-012-9358-6).
- [23] National development and reform commission (NDRC) People's Republic of China. Implementation plan for further improving the market oriented green technology innovation system (2023-2025), from https://www.ndrc.gov.cn/xxgk/zcfb/tz/202212/t20221228_1344205.html, accessed December 11, 2022.
- [24] Dhanorkar, S., Donohue, K., Linderman, K. (2015). Repurposing materials and waste through online exchanges: Overcoming the last hurdle, *Production & Operations Management*, Vol. 24, No. 9, 1473-1493, doi: [10.1111/poms.12345](https://doi.org/10.1111/poms.12345).
- [25] Mansyur, A., Subrahmanyam, S., Ponkratov, V., Zonyfar, C., Akhmadeev, R., Manoharan, K. (2022). A new mathematical model for the integration of production and distribution in mobile centers, *Mathematical Modelling of Engineering Problems*, Vol. 9, No. 6, 1466-1470, doi: [10.18280/mmep.090604](https://doi.org/10.18280/mmep.090604).
- [26] Wang, Y., Yu, Z., Shen, L., Dong, W. (2021). Impacts of altruistic preference and reward-penalty mechanism on decisions of E-commerce closed-loop supply chain, *Journal of Cleaner Production*, Vol. 315, Article No. 128132, doi: [10.1016/j.jclepro.2021.128132](https://doi.org/10.1016/j.jclepro.2021.128132).
- [27] Sarkar, B., Ullah, M., Sarkar, M. (2022). Environmental and economic sustainability through innovative green products by remanufacturing, *Journal of Cleaner Production*, Vol. 332, Article No. 129813, doi: [10.1016/j.jclepro.2021.129813](https://doi.org/10.1016/j.jclepro.2021.129813).
- [28] Zhou, X., Liu, Z., Liu, J., Ku, Z. (2022). The choice of cooperative technology innovation strategies in a supply chain under governmental subsidy, *RAIRO - Operations Research*, Vol. 56, No. 4, 2669-2700, doi: [10.1051/ro/2022121](https://doi.org/10.1051/ro/2022121).
- [29] Zheng, X., Govindan, K., Deng, Q., Feng, L. (2019). Effects of design for the environment on firms' production and remanufacturing strategies, *International Journal of Production Economics*, Vol. 213, 217-228, doi: [10.1016/j.ijpe.2019.03.019](https://doi.org/10.1016/j.ijpe.2019.03.019).
- [30] Reimann, M., Xiong, Y., Zhou, Y. (2019). Managing a closed-loop supply chain with process innovation for remanufacturing, *European Journal of Operational Research*, Vol. 276, No. 2, 510-518, doi: [10.1016/j.ejor.2019.01.028](https://doi.org/10.1016/j.ejor.2019.01.028).
- [31] Lee, D. (2020). Who drives green innovation? A game theoretical analysis of a closed-loop supply chain under different power structures, *International Journal of Environmental Research and Public Health*, Vol. 17, No. 7, Article No. 2274, doi: [10.3390/ijerph17072274](https://doi.org/10.3390/ijerph17072274).
- [32] Ma, W.-M., Zhao, Z., Ke, H. (2013). Dual-channel closed-loop supply chain with government consumption-subsidy, *European Journal of Operational Research*, Vol. 226, No. 2, 221-227, doi: [10.1016/j.ejor.2012.10.033](https://doi.org/10.1016/j.ejor.2012.10.033).
- [33] Wang, C., Nie, P.-Y., Peng, D.-H., Li, Z.-H. (2017). Green insurance subsidy for promoting clean production innovation, *Journal of Cleaner Production*, Vol. 148, 111-117, doi: [10.1016/j.jclepro.2017.01.145](https://doi.org/10.1016/j.jclepro.2017.01.145).
- [34] Liu, Y., Quan, B.-T., Xu, Q., Forrest, J.Y.-L. (2019). Corporate social responsibility and decision analysis in a supply chain through government subsidy, *Journal of Cleaner Production*, Vol. 208, 436-447, doi: [10.1016/j.jclepro.2018.10.121](https://doi.org/10.1016/j.jclepro.2018.10.121).
- [35] Wan, N., Hong, D. (2019). The impacts of subsidy policies and transfer pricing policies on the closed-loop supply chain with dual collection channels, *Journal of Cleaner Production*, Vol. 224, 881-891, doi: [10.1016/j.jclepro.2019.03.274](https://doi.org/10.1016/j.jclepro.2019.03.274).
- [36] Aldieri, L., Carlucci, F., Cirà, A., Ioppolo, G., Vinci, C.P. (2019). Is green innovation an opportunity or a threat to employment? An empirical analysis of three main industrialized areas: The USA, Japan and Europe, *Journal of Cleaner Production*, Vol. 214, 758-766, doi: [10.1016/j.jclepro.2019.01.016](https://doi.org/10.1016/j.jclepro.2019.01.016).
- [37] Huang, S., Fan, Z.-P., Wang, N. (2020). Green subsidy modes and pricing strategy in a capital-constrained supply chain, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 136, Article No. 101885, doi: [10.1016/j.tre.2020.101885](https://doi.org/10.1016/j.tre.2020.101885).
- [38] Guo, J., Yu, H., Gen, M. (2020). Research on green closed-loop supply chain with the consideration of double subsidy in e-commerce environment, *Computers & Industrial Engineering*, Vol. 149, Article No. 106779, doi: [10.1016/j.cie.2020.106779](https://doi.org/10.1016/j.cie.2020.106779).
- [39] Zhou, Y., He, Z., Zhao, S. (2021). How do government subsidies affect the strategic choices of enterprises and individuals in agricultural waste recycling?, *Sustainable Production and Consumption*, Vol. 28, 1687-1698, doi: [10.1016/j.spc.2021.09.014](https://doi.org/10.1016/j.spc.2021.09.014).
- [40] Bai, S., Ge, L., Zhang, X. (2022). Platform or direct channel: Government-subsidized recycling strategies for WEEE, *Information Systems and e-Business Management*, Vol. 20, 347-369, doi: [10.1007/s10257-021-00517-4](https://doi.org/10.1007/s10257-021-00517-4).
- [41] Lakhout, A., Shaban, M., Alatawi, A., Abbas, S.Y.H., Asiri, E., Al Juhni, T., Elsayw, M. (2023). Machine-learning approaches in geo-environmental engineering: Exploring smart solid waste management, *Journal of Environmental Management*, Vol. 330, Article No. 117174, doi: [10.1016/j.jenvman.2022.117174](https://doi.org/10.1016/j.jenvman.2022.117174).
- [42] Zhu, M., Li, X., Zhu, L., Zhan, X., Ma, J. (2021). Dynamic evolutionary games and coordination of multiple recycling channels considering online recovery platform, *Discrete Dynamics in Nature and Society*, Vol. 2021, Article Id.

- 9976157, doi: [10.1155/2021/9976157](https://doi.org/10.1155/2021/9976157).
- [43] Lin, Y., Yu, Z., Wang, Y., Goh, M. (2023). Performance evaluation of regulatory schemes for retired electric vehicle battery recycling within dual-recycle channels, *Journal of Environmental Management*, Vol. 332, Article No. 117354, doi: [10.1016/j.jenvman.2023.117354](https://doi.org/10.1016/j.jenvman.2023.117354).
- [44] Feng, L., Govindan, K., Li, C. (2017). Strategic planning: Design and coordination for dual-recycling channel reverse supply chain considering consumer behavior, *European Journal of Operational Research*, Vol. 260, No. 2, 601-612, doi: [10.1016/j.ejor.2016.12.050](https://doi.org/10.1016/j.ejor.2016.12.050).
- [45] Li, C., Feng, L., Luo, S. (2019). Strategic introduction of an online recycling channel in the reverse supply chain with a random demand, *Journal of Cleaner Production*, Vol. 236, Article No. 117683, doi: [10.1016/j.jclepro.2019.117683](https://doi.org/10.1016/j.jclepro.2019.117683).
- [46] Wang, Y., Yu, Z., Shen, L., Jin, M. (2022). Operational modes of E-closed loop supply chain considering platforms' services, *International Journal of Production Economics*, Vol. 251, Article No. 108551, doi: [10.1016/j.ijpe.2022.108551](https://doi.org/10.1016/j.ijpe.2022.108551).
- [47] Jia, D., Li, S. (2020). Optimal decisions and distribution channel choice of closed-loop supply chain when e-retailer offers online marketplace, *Journal of Cleaner Production*, Vol. 265, Article No. 121767, doi: [10.1016/j.jclepro.2020.121767](https://doi.org/10.1016/j.jclepro.2020.121767).
- [48] Zhang, X., Hou, W. (2022). The impacts of e-tailer's private label on the sales mode selection: From the perspectives of economic and environmental sustainability, *European Journal of Operational Research*, Vol. 296, No. 2, 601-614, doi: [10.1016/j.ejor.2021.04.009](https://doi.org/10.1016/j.ejor.2021.04.009).
- [49] Matsui, K. (2022). Optimal timing of acquisition price announcement for used products in a dual-recycling channel reverse supply chain, *European Journal of Operational Research*, Vol. 300, No. 2, 615-632, doi: [10.1016/j.ejor.2021.08.010](https://doi.org/10.1016/j.ejor.2021.08.010).
- [50] Wang, Y., Yu, Z., Shen, L. (2018). Study on the decision-making and coordination of an e-commerce supply chain with manufacturer fairness concerns, *International Journal of Production Research*, Vol. 57, No. 9, 2788-2808, doi: [10.1080/00207543.2018.1500043](https://doi.org/10.1080/00207543.2018.1500043).
- [51] Basiri, Z., Heydari, J. (2017). A mathematical model for green supply chain coordination with substitutable products, *Journal of Cleaner Production*, Vol. 145, 232-249, doi: [10.1016/j.jclepro.2017.01.060](https://doi.org/10.1016/j.jclepro.2017.01.060).
- [52] Zhang, L., Wang, J., You, J. (2015). Consumer environmental awareness and channel coordination with two substitutable products, *European Journal of Operational Research*, Vol. 241, No. 1, 63-73, doi: [10.1016/j.ejor.2014.07.043](https://doi.org/10.1016/j.ejor.2014.07.043).
- [53] De Giovanni, P., Zaccour, G. (2019). Optimal quality improvements and pricing strategies with active and passive product returns, *Omega*, Vol. 88, 248-262, doi: [10.1016/j.omega.2018.09.007](https://doi.org/10.1016/j.omega.2018.09.007).
- [54] Wu, C.-H. (2012). Price and service competition between new and remanufactured products in a two-echelon supply chain, *International Journal of Production Economics*, Vol. 140, No. 1, 496-507, doi: [10.1016/j.ijpe.2012.06.034](https://doi.org/10.1016/j.ijpe.2012.06.034).
- [55] Friedman, D. (1991). Evolutionary games in economics, *Econometrica* Vol. 59, No. 3, 637-666, doi: [10.2307/2938222](https://doi.org/10.2307/2938222).
- [56] Beijing municipal development and reform commission. Measures of Beijing municipality on further improving the market-oriented green technology innovation system, from https://fgw.beijing.gov.cn/gzdt/tztg/202111/t20211104_2529410.htm, accessed January 7, 2024.
- [57] Lumpkin, G.T., Dess, G.G. (2001). Linking two dimensions of entrepreneurial orientation to firm performance: The moderating role of environment and industry life cycle, *Journal of Business Venturing*, Vol. 16, No. 5, 429-451, doi: [10.1016/S0883-9026\(00\)00048-3](https://doi.org/10.1016/S0883-9026(00)00048-3).

FDM process parameter selection by hybrid MCDM approach for flexural and compression strength maximization

Begic-Hajdarevic, D.^a, Klančnik, S.^b, Muhamedagić, K.^a, Čekić, A.^a, Cohodar Husic, M.^a, Ficko, M.^{b,*}, Gusel, L.^b

^aUniversity of Sarajevo, Faculty of Mechanical Engineering, Sarajevo, Bosnia and Herzegovina

^bUniversity of Maribor, Faculty of Mechanical Engineering, Maribor, Slovenia

ABSTRACT

Fused deposition modelling (FDM) is one of the mostly used additive technologies, due to its ability to produce complex parts with good mechanical properties. The selection of FDM process parameters is crucial to achieve good mechanical properties of the manufactured parts. Therefore, in this paper, a hybrid multi-criteria decision-making (MCDM) approach based on Preference Selection Index (PSI) and Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) is proposed for the selection of optimal process parameters in FDM printing of polylactic acid (PLA) parts. Printing temperature, layer thickness and raster angle were considered as input process parameters. In order to prove the effectiveness of the proposed hybrid PSI – TOPSIS method, the obtained results were compared with the results obtained with different MCDM methods. The obtained best option of process parameters was confirmed by other MCDM methods. The optimal combination of process parameters to achieve the maximal flexural strength, maximal flexural modulus and maximal compressive strength is selected using the hybrid PSI-TOPSIS method. The results show that the hybrid PSI-TOPSIS approach could be used for optimisation process parameters for any machining process.

ARTICLE INFO

Keywords:

Fused deposition modelling (FDM);
Multi-criteria decision-making (MCDM);
Hybrid PSI-TOPSIS method;
Process parameters;
Mechanical properties;
Optimization

*Corresponding author:

mirko.ficko@um.si
(Ficko, M.)

Article history:

Received 6 November 2023

Revised 6 March 2024

Accepted 13 March 2024



Content from this work may be used under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

Additive manufacturing (AM) represents a way of production that is based on making products by adding materials “layer by layer” This method of production was initially used only for the rapid prototyping (RP), but today additive technologies are also used for the production of highly functional products in small quantities. This principle of making parts remains the same regardless of the degree of geometric complexity of the part, which is the main advantage of this technology.

According to the physics of the process and material type a large number of different additive manufacturing processes have been developed. Some of the most commonly used procedures are: fused deposition modelling (FDM), stereolithography (SLA), Ink jet modelling, selective laser sintering (SLS), etc. The FDM process is used for prototyping and production of fully functional parts for engineering applications. Some of most commonly used material for FDM process are: Polylactic-Acid (PLA), Polyethylene-Terephthalate (PET), Acrylonitrile Butadiene-

Styrene (ABS), propylene (PP), Polyamide (PA), and Thermoplastic-Polyurethane (TPU) [1,2]. Quality and mechanical properties of FDM produced parts are key factors for their use in industrial applications. In order to achieve the appropriate quality and mechanical properties of parts made in this way, it is necessary to carefully design the FDM process in terms of the correct selection of input parameters. Due to the large number of input and output parameters, it is often necessary to solve complex optimization problems. In order to solve such problems and avoid the need to perform a large number of experiments, a systematic approach to the experiment plan and the application of various methods of multi-criteria optimization are used. Some of the most commonly used methods for experiment design, modelling and optimization of process parameters of the FDM technology are: Taguchi Method [3], Grey Relational Analysis (GRA) [4], Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) [5], Response Surface Methodology (RSM) [6], Genetic Algorithm (GA) [7], Artificial Neural Network (ANN) [6], and Adaptive Neuro Fuzzy Interface System (ANFIS) [8]. Many researchers have analysed the possibility of applying these methods to optimize the various output parameters of the FDM process.

In study [9], the authors optimized three input parameters, infill density, printing speed and printing temperature, to achieve the maximum tensile strength of samples made of PLA material. For optimisation these process parameters were used hybrid optimization techniques, genetic algorithm-artificial neural network, genetic algorithm-response surface methodology and genetic algorithm-adaptive neuro fuzzy interface system. It is shown that such hybrid models could be used for optimisation any other process parameters for any industrial application problems. Rajamani *et al.* [5] were used hybrid approach through RSM-TOPSIS method. This approach was used for improving surface quality of micro sized near-net-shaped components for end use applications using FDM additive manufacturing techniques. Production of test specimens was carried out according to the previously defined Box-Behnken experimental design. For input parameters were selected: layer thickness, part orientation, raster width and raster angle. The optimal FDM parameters for improved surface quality attributes were determined using TOPSIS method. This method proved to be a useful tool for finding optimal FDM process parameters for fabricating the components of a flapping wing micro mechanism. Also, the TOPSIS method proved to be useful in the selection of optimal process parameters in two-point incremental forming process [10], as well as for optimization of cutting parameters in turning process [11], and thanks to the proposed Fuzzy-TOPSIS approach [12], managers of manufacturing companies can access and monitor the maintenance sustainability level integrated with the industry 4.0 technologies. The RSM method is a combination of statistical and mathematical methods that is very useful for modelling and optimizing engineering scientific problems, which gives very low standard errors to experimental verification. Srinivasan *et al.* [13] used the RSM method based on central composite design to predict of tensile strength in FDM printed ABS parts. In paper [6] was successfully applied RSM and ANN to investigate the effect of the layer thickness, printing speed, raster angle and wall thickness on the tensile strength of test specimens printed with a short carbon fibre reinforced polyamide composite.

Taguchi-Grey relational analysis was used in the study [14] to optimize input parameters and improve selected output mechanical characteristics. This study is designed to capture the said gap in the literature with focus on cell geometry, nozzle diameter and strain rate by using the Taguchi design of experimentation and Grey Relational Analysis. It is shown that the GRA method significantly simplifies complex optimization problems in FDM process parameters optimization. Taguchi method is very useful in the experimental plan phase, and it can be used separately [15] or with other methods for multi-criteria optimization [14] [16]. Chohan *et al.* [16] were using Taguchi-TOPSIS based optimization of FDM process parameters for manufacturing ABS plastics parts. The results were shown that using the TOPSIS method, optimal parameters can be determined in order to improve the surface-quality of FDM parts which can be utilized for end-use products and for rapid tooling applications.

In addition to the mentioned methods, there is also the Preference Selection Index (PSI) method that can be used to solve multi-criteria optimization problems. The possibility of applying the PSI method for the selection of optimal FDM process parameters was investigated [17]. It is found that the PSI method is very simple to understand and easy to implement. The advantage

of the PSI method is that there is no need to calculate the relative weight of outputs. However, some authors [18] have observed that this method is not useful when several alternatives have criteria values that are very close to those are preferred. A hybrid TOPSIS-PSI method for selection material in marine applications was presented in study [19], the entropy method has been used to determine the weights of the selected criteria.

In this paper, a hybrid method that combines of the PSI and TOPSIS method is proposed. The proposed hybrid method considered the advantage of the PSI method that does not require the calculation of the weight factor of criteria and the advantage of the TOPSIS method that is more efficient in dealing with the criteria and the number of available alternatives. To test the proposed method, the case of selecting optimal process parameters to improve the mechanical properties of FDM printed PLA parts was considered.

2. Materials and methods

2.1 Experimental details

A 3D printer Ultimaker S5 was used to produce test samples from PLA material, that is one of the mostly used FDM material. The samples for flexural and the compressive tests were designed and tested according to ISO 178 and ISO 604 standards, respectively. The constant process parameters for printing test samples are shown in Table 1. In this paper, three input process parameters, namely printing temperature, layer thickness and raster angle were investigated in order to study their influence on the mechanical properties of the test samples using the Taguchi design of experiments. These parameters are varied at three different levels as shown in Table 2.

Table 1 The constant process parameters and their values

Parameter	Unit	Value
Nozzle diameter	mm	0.4
Infill density	%	100
Build plate temperature	°C	110
Build direction		Flat x-x direction
Printing speed	mm/s	60

Table 2 The input parameters and their levels

Parameter	Symbol	Unit	Level 1	Level 2	Level 3
Printing temperature	T	°C	180	200	220
Layer thickness	L	mm	0.1	0.2	0.3
Raster angle	A	°	0	45	90

Three samples were tested for each set of input parameters. The experimentally studied output parameters were flexural strength (FS), flexural modulus (FM) and compressive strength (CS). The previous research [2, 6-9] were focused on analysing the influence of process parameters on the tensile strength of FDM printed parts. The average value of the output parameters is reported in Table 3. Flexural and compressive tests were conducted on the 10 kN Shimadzu AGS-X universal machine.

Table 3 Experimental data

Exp. No.	Input parameters			Output parameters		
	L (mm)	T (°C)	A (°)	Flexural strength (MPa)	Flexural modulus (MPa)	Compressive strength (MPa)
1	0.1	180	0	38.55	2505.65	45.16
2	0.1	200	45	81.56	3014.92	46.78
3	0.1	220	90	92.24	2952.91	48.14
4	0.2	180	45	36.53	1794.45	39.45
5	0.2	200	90	79.77	2598.18	45.34
6	0.2	220	0	72.24	2650.26	44.37
7	0.3	180	90	29.65	2054.71	39.81
8	0.3	200	0	52.10	2221.84	41.63
9	0.3	220	45	72.38	2502.59	41.30

2.2 Hybrid PSI-TOPSIS method

Process parameters selection for any machining process is a MCDM problem that considers different competing criteria for selecting appropriate process parameters. The proposed hybrid PSI – TOPSIS method consists of the following steps.

Step 1: Determine a set of experimental trials (alternatives):

$$E = [E_1, E_2, \dots, E_m] \quad (1)$$

where m is the number of experimental trials.

Step 2: Determine a set of criteria (output parameters):

$$C = [C_1, C_2, \dots, C_n] \quad (2)$$

where n is the number of criteria.

Step 3: Creating a decision matrix:

$$D = [D_{ij} | i = 1, 2, \dots, m; j = 1, 2, \dots, n] \quad (3)$$

and D_{ij} is the value of the j -th criterion for the i -th experimental trial.

Step 4: Calculation of the normalized matrix:

a) if the larger is better (LB):

$$N_{ij} = \frac{D_{ij}}{D_j^{\max}}, \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (4)$$

b) if the smaller is better (SB):

$$N_{ij} = \frac{D_j^{\min}}{D_{ij}}, \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (5)$$

Step 5: Calculating the mean value of the normalized matrix:

$$N_j = \frac{1}{m} \sum_{i=1}^m N_{ij}, \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (6)$$

Step 6: Calculating the value of the preference variation:

$$\varphi_j = \sum_{i=1}^m [N_{ij} - N_j]^2, \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (7)$$

Step 7: Calculating the deviation in preference value:

$$\Delta_j = \left\{ 1 - \sum_{i=1}^m [N_{ij} - N_j]^2 \right\}, \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (8)$$

Step 8: Determine the overall preference value (weight factors for each criteria):

$$p_j = \frac{\Delta_j}{\sum_{j=1}^n \Delta_j}, \quad j = 1, 2, \dots, n \text{ and } \sum_{j=1}^n p_j = 1 \quad (9)$$

Step 9: Creating a weighted normalized decision matrix:

$$w_{ij} = p_j \times N_{ij} \text{ for } i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (10)$$

Step 10: Determine the positive (PS) and negative ideal solution (NS):

$$PS = [w_1^+, \dots, w_j^+, \dots, w_n^+],$$

$$\text{where } w_j^+ = \begin{cases} \max w_{ij} & \text{if } j \in LB \\ \min w_{ij} & \text{if } j \in SB \end{cases} \text{ for } j = 1, 2, \dots, n \quad (11)$$

$$NS = [w_1^-, \dots, w_j^-, \dots, w_n^-],$$

$$\text{where } w_j^- = \begin{cases} \min w_{ij} & \text{if } j \in LB \\ \max w_{ij} & \text{if } j \in SB \end{cases} \text{ for } j = 1, 2, \dots, n \quad (12)$$

Step 11: Obtain the distances of each experimental trials in relation to ideal solutions:

$$S_i^+ = \sqrt{\sum_{j=1}^n (w_{ij} - w_j^+)^2}, \quad i = 1, 2, \dots, m \quad (13)$$

$$S_i^- = \sqrt{\sum_{j=1}^n (w_{ij} - w_j^-)^2}, \quad i = 1, 2, \dots, m \quad (14)$$

Step 12: Calculate the closeness index value:

$$C_i = \frac{S_i^-}{S_i^+ + S_i^-}, \quad i = 1, 2, \dots, m \quad (15)$$

Step 13: Rank the closeness index in the descending order.

3. Results and discussion

In order to demonstrate and prove of the effectiveness of the proposed PSI-TOPIS method, practical example of the selection of FDM process parameters were presented. Also, the results obtained by the proposed hybrid method were compared with the results obtained using other MCDM methods.

The experimental data from Table 3 were normalized using Eq. 4 and the matrix was shown in Table 4. In Table 5 were presented data that were calculated using Eq. 6, 7 and 8, as well as the weight factors for each criterion (using Eq. 9). From Table 5 it can be seen that the compressive strength is most important criteria. The weighted normalized decision matrix was determined using Eq. 10 and this matrix is also shown in Table 4 due to space limitation.

Using Eqs. 11 and 12, the positive and negative ideal solution were determined. Further, the distances of each experimental trials (alternatives) in relation to positive and negative ideal solution were calculated using Eq. 13 and 14 and given in Table 6. Also, Table 6 shows the closeness index calculated using Eq. 15 and the ranking order of given alternatives.

Table 4 Matrix N_{ij} and w_{ij}

Exp. No.	Normalized matrix			Weighted normalized matrix		
	FS (MPa)	FM (MPa)	CS (MPa)	FS (MPa)	FM (MPa)	CS (MPa)
1	0.4179	0.8311	0.9381	0.0912	0.3059	0.3880
2	0.8842	1.0000	0.9717	0.1930	0.3681	0.4019
3	1.0000	0.9794	1.0000	0.2183	0.3605	0.4136
4	0.3960	0.5952	0.8195	0.0865	0.2191	0.3389
5	0.8648	0.8618	0.9418	0.1888	0.3172	0.3896
6	0.7832	0.8790	0.9217	0.1710	0.3236	0.3812
7	0.3214	0.6815	0.8270	0.0702	0.2508	0.3420
8	0.5648	0.7369	0.8648	0.1233	0.2713	0.3577
9	0.7847	0.8301	0.8579	0.1713	0.3055	0.3548

Table 5 Data determination using Eqs. 6-9

Criteria	N_j	φ_j	Δ_j	p_j
FS	0.6686	0.4898	0.5102	0.2183
FM	0.8217	0.1399	0.8601	0.3681
CS	0.9047	0.0334	0.9666	0.4136

Table 6 Closeness index and ranking

Exp. No.	S_i^+	S_i^-	C_i	Rank
1	0.1438	0.1019	0.4149	6
2	0.0278	0.2031	0.8794	2
3	0.0076	0.2180	0.9664	1
4	0.2125	0.0163	0.0712	9
5	0.0635	0.1621	0.7183	3
6	0.0726	0.1512	0.6756	4
7	0.2020	0.0319	0.1365	8
8	0.1467	0.0768	0.3436	7
9	0.0979	0.1340	0.5779	5

Results from Table 6, that were obtained using hybrid PSI-TOPSIS method, show that the alternative E_3 is the best option, while the alternative E_4 is the worst choice. Flexural and compressive stress-strain curves for the best and worst alternatives are shown in Fig. 1. The optimal combination of FDM input parameters for printing PLA parts with regard to the considered process performance are 220°C printing temperature, 0.10 mm layer thickness and 90° raster angle, as also shown in Table 7. In this table the bold value indicates level at optimal parameter settings for individual input parameters. It is clear that printing temperature has the most significant effect on the process performance, followed by layer thickness and then raster angle.

Table 7 Response table for the mean C_i .

Input parameters	Closeness index			max.-min.	Rank
	Level 1	Level 2	Level 3		
T	0.2075	0.6471	0.7400	0.5325	1
L	0.7536	0.4884	0.3527	0.4009	2
A	0.4780	0.5095	0.6071	0.1291	3

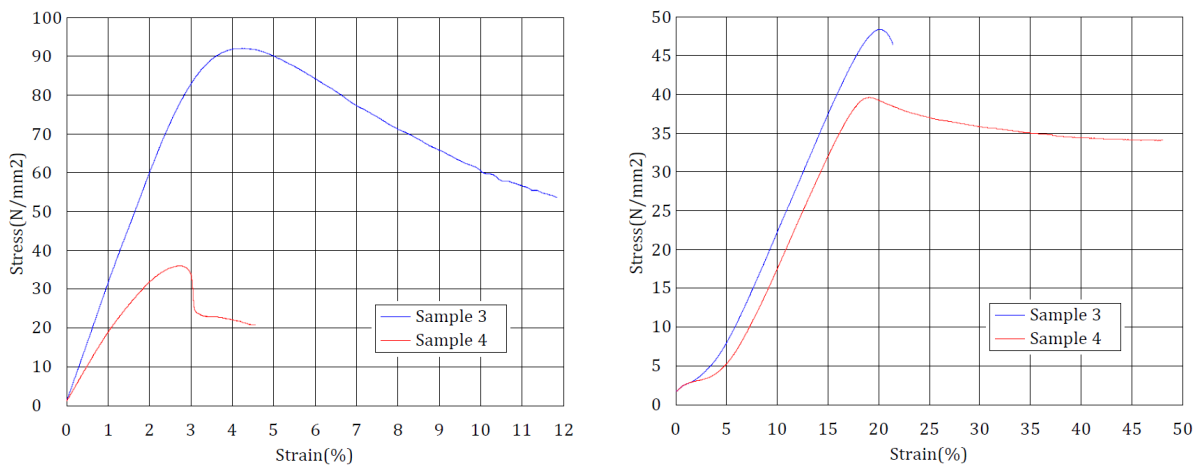


Fig. 1 The flexural (left) and compressive (right) stress-strain curves for the best (sample 3) and worst (sample 4) options

The effect of input parameters on the flexural strength and compressive strength are illustrated in Fig. 2 and Fig. 3, respectively. Results showed that flexural strength and compressive strength increased by increasing the printing temperature. This can be explained by the fact that by increasing the printing temperature, the stronger cohesive forces were realized between individual raster and layers, that resulted in higher flexural and compressive strength.

By increasing of layer thickness, flexural strength and compressive strength decrease, because the porosity between individual layers increases. Also, it can be observed that the highest values of flexural and compressive strength were achieved at a raster angle of 90°, because in this case, the direction of material deposition coincides with the direction of the load. The lowest value of the flexural strength was obtained at a raster angle of 0°, because in this case the strength of the test samples primarily depends on the cohesive force between individual raster. While the lowest value of the compressive strength was achieved at a raster angle of 45° due to the shear stresses between individual raster.

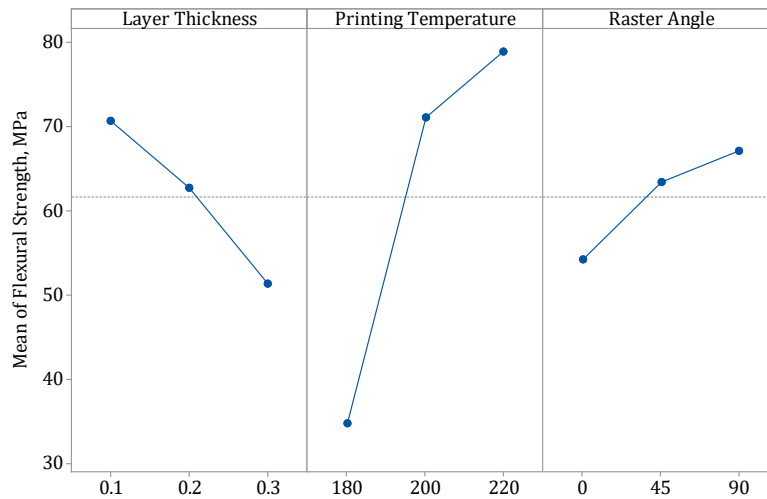


Fig. 2 Mean of the flexural strength for different levels of input parameters

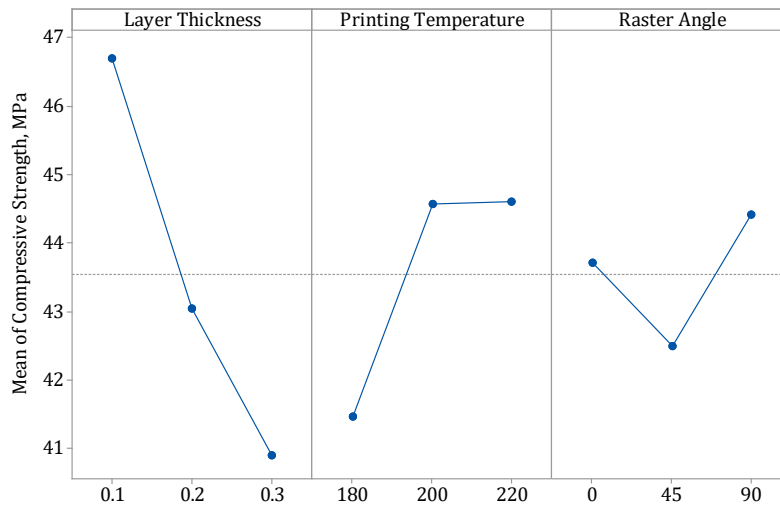


Fig. 3 Mean of the compressive strength for different levels of input parameters

Therefore, based on everything stated above, it can be concluded that the proposed hybrid PSI-TOPSIS method provides effectively very good results.

According to the proposed method, the best option (alternative no. 3) was achieved at the highest varied printing temperature, the smallest varied layer thickness and the raster angle of 90°. The best choice (alternative 3) was confirmed by all other MCDM methods (GRA, TOPSIS and TOPSIS-ENTROPY), as can be seen in Fig. 4.

The worst option, as predicted by the hybrid PSI-TOPSIS method, is alternative no. 4. This was not predicted by the other MCDM methods, as clearly shown in Fig. 4. Thus, the worst choice, as predicted by the others considered methods, is alternative no.7. The worst option, predicted by the proposed PSI-TOPSIS method, was achieved at the lowest varied printing temperature, the mean value of the layer thickness and the raster angle of 45° (it is most unfavourable angle for the compressive strength, as seen in Fig. 3). Alternative 7, as the worst choice predicted by the other methods, was also achieved at the lowest value of the printing temperature. Given that the results (as seen in Table 7) showed that the printing temperature has a most important effect on the process performance and that by decreasing the printing temperature the flexural and compressive strength decrease (as shown in Fig. 2 and Fig. 3), this proves the effectiveness of the proposed method. The effectiveness of the proposed method is also proven by the fact that the worst option (alternative 4) was achieved at the most unfavourable raster angle for the compressive strength (compressive strength is the most important criteria, as shown in Table 5). This was not predicted by the other MCDM methods. The worst option predicted by the

other MCDM methods was achieved at the raster angle of 90°. This raster angle is the most favourable angle for both considered criteria (flexural strength and compressive strength), as seen in Figs. 2 and 3.

Thus, in this paper, the determination of the best option does not depend on the MCDM methods used, it was also shown in [20]. However, the worst alternative predicted by the proposed hybrid method, unlike the other methods used, shows a good ranking order of the alternatives by the proposed method, that is an advantage proposed method in compared to the other methods used. Certainly, this advantage offered by the proposed method should be proven in other cases, that is a suggestion for future research.

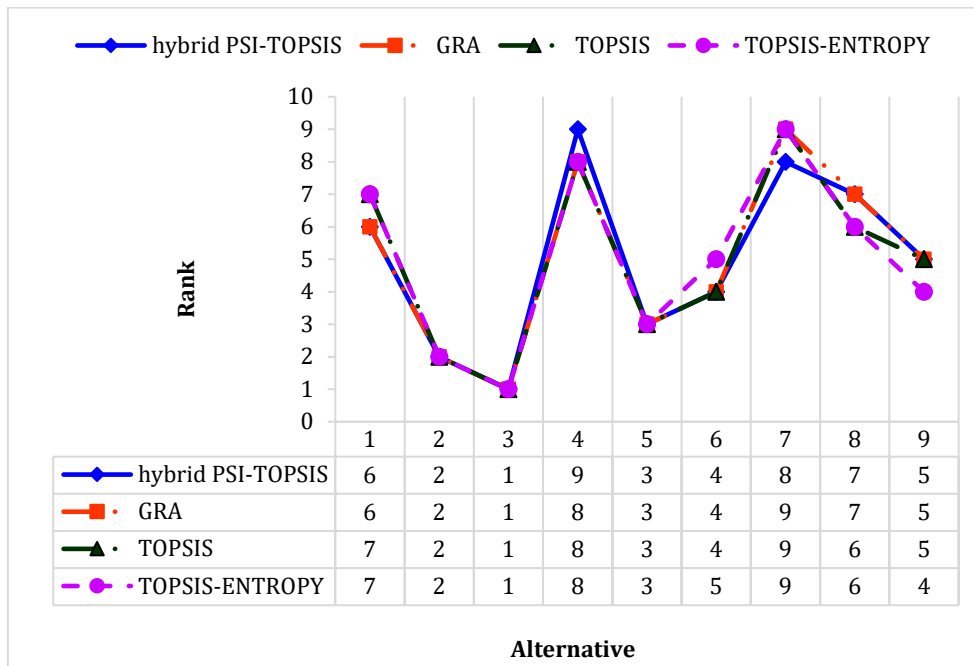


Fig. 4 Comparison of ranking with different MCDM methods

4. Conclusion

In this paper, a novel hybrid PSI-TOPSIS methodology was presented. The proposed method was tested on the example of selecting optimal process parameters during FDM printing of PLA samples. Also, the results obtained by the PSI-TOPSIS method were compared with the results that obtained by other MCDM methods. The results show that a printing temperature of 220°C, a layer thickness of 0.10 mm and a raster angle of 90° would be the best choice of process parameters according to PSI-TOPSIS analysis which has the best combination of mechanical properties of the tested samples.

In future research, a hybrid PSI-TOPSIS method will be proposed for the selection of process parameters in other non-conventional machining processes, such as laser cutting or abrasive water jet cutting.

Acknowledgement

The authors acknowledge the financial support from the Slovenian Research Agency (research core funding No. P2-0157) and the Ministry of Science, Higher Education and Youth of the Sarajevo Canton.

References

- [1] Shanmugam, V., Pavan, M.V., Babu, K., Karnan, B. (2021). Fused deposition modeling based polymeric materials and their performance: A review, *Polymer Composites*, Vol. 42, No. 11, 5656-5677, doi: [10.1002/pc.26275](https://doi.org/10.1002/pc.26275).
- [2] Tripathy, C.R., Sharma, R.K., Rattan, V.K. (2022). Effect of printing parameters on the mechanical behaviour of the thermoplastic polymer processed by FDM technique: A research review, *Advances in Production Engineering & Management*, Vol. 17, No. 3, 279-294, doi: [10.14743/apem2022.3.436](https://doi.org/10.14743/apem2022.3.436).
- [3] Alafaghani, A., Qattawi, A. (2018). Investigating the effect of fused deposition modeling processing parameters using Taguchi design of experiment method, *Journal of Manufacturing Processes*, Vol. 36, 164-174, doi: [10.1016/j.jmapro.2018.09.025](https://doi.org/10.1016/j.jmapro.2018.09.025).
- [4] Shakeri, Z., Benfriha, K., Zirak, N., Shirinbayan, M. (2022). Mechanical strength and shape accuracy optimization of polyamide FFF parts using grey relational analysis, *Scientific Reports*, Vol. 12, No. 1, Article No. 13142, doi: [10.1038/s41598-022-17302-z](https://doi.org/10.1038/s41598-022-17302-z).
- [5] Rajamani, D., Balasubramanian, E., Yang, L.-J. (2022). Enhancing the surface quality of FDM processed flapping wing micro mechanism assembly through RSM-TOPSIS hybrid approach, *Processes*, Vol. 10, No. 11, Article No. 2457, doi: [10.3390/pr10112457](https://doi.org/10.3390/pr10112457).
- [6] Muhamedagic, K., Berus, L., Potočnik, D., Cekic, A., Begic-Hajdarevic, D., Cohodar Husic, M., Ficko, M. (2022). Effect of process parameters on tensile strength of FDM printed carbon fiber reinforced polyamide parts, *Applied Sciences*, Vol. 12, No. 12, Article No. 6028, doi: [10.3390/app12126028](https://doi.org/10.3390/app12126028).
- [7] Omer, R., Mali, H.S., Singh, S.K. (2020). Tensile performance of additively manufactured short carbon fibre-PLA composites: Neural networking and GA for prediction and optimisation, *Plastics, Rubber and Composites: Macromolecular Engineering*, Vol. 49, No. 6, 271-280, doi: [10.1080/14658011.2020.1744371](https://doi.org/10.1080/14658011.2020.1744371).
- [8] Yadav, D., Chhabra, D., Gupta, R.K., Phogat, A., Ahlawat, A. (2020). Modeling and analysis of significant process parameters of FDM 3D printer using ANFIS, *Materials Today: Proceedings*, Vol. 21, Part 3, 1592-1604, doi: [10.1016/j.matpr.2019.11.227](https://doi.org/10.1016/j.matpr.2019.11.227).
- [9] Deshwal, S., Kumar, A., Chhabra, D. (2020). Exercising hybrid statistical tools GA-RSM, GA-ANN and GA-ANFIS to optimize FDM process parameters for tensile strength improvement, *CIRP Journal of Manufacturing Science and Technology*, Vol. 31, 189-199, doi: [10.1016/j.cirpj.2020.05.009](https://doi.org/10.1016/j.cirpj.2020.05.009).
- [10] Visagan, A., Ganesh, P. (2022). Parametric optimization of two point incremental forming using GRA and TOPSIS, *International Journal of Simulation Modelling*, Vol. 21, No. 4, 615-626, doi: [10.2507/IJSIMM21-4-622](https://doi.org/10.2507/IJSIMM21-4-622).
- [11] Ficko, M., Begic-Hajdarevic, D., Hadziabdic, V., Klančnik, S. (2020). Multi-response optimisation of turning process parameters with GRA and TOPSIS methods, *International Journal of Simulation Modelling*, Vol. 19, No. 4, 547-558, doi: [10.2507/IJSIMM19-4-524](https://doi.org/10.2507/IJSIMM19-4-524).
- [12] Patalas-Maliszewska, J., Losyk, H. (2022). An approach to maintenance sustainability level assessment integrated with Industry 4.0 technologies using Fuzzy-TOPSIS: A real case study, *Advances in Production Engineering & Management*, Vol. 17, No. 4, 455-468, doi: [10.14743/apem2022.4.448](https://doi.org/10.14743/apem2022.4.448).
- [13] Srinivasan, R., Pridhar, T., Ramprasath, L.S., Sree Charan, N., Ruban, W. (2020). Prediction of tensile strength in FDM printed ABS parts using response surface methodology (RSM), *Materials Today Proceedings*, Vol. 27, Part 2, 1827-1832, doi: [10.1016/j.matpr.2020.03.788](https://doi.org/10.1016/j.matpr.2020.03.788).
- [14] John, J., Devjani, D., Ali, S., Abdallah, S., Pervaiz, S. (2023). Optimization of 3D printed polylactic acid structures with different infill patterns using Taguchi-grey relational analysis, *Advanced Industrial and Engineering Polymer Research*, Vol. 6, No. 1, 62-78, doi: [10.1016/j.aiepr.2022.06.002](https://doi.org/10.1016/j.aiepr.2022.06.002).
- [15] Wankhede, V., Jagetiya, D., Joshi, A., Chaudhari, R. (2020). Experimental investigation of FDM process parameters using Taguchi analysis, *Materials Today: Proceedings*, Vol. 27, Part 3, 2117-2120, doi: [10.1016/j.matpr.2019.09.078](https://doi.org/10.1016/j.matpr.2019.09.078).
- [16] Chohan, J.S., Kumar, R., Singh, T.B., Singh, S., Sharma, S., Singh, J., Mia, M., Pimenov, D.Y., Chattopadhyaya, S., Dwivedi, S.P., Kapłonek, W. (2020). Taguchi S/N and TOPSIS based optimization of fused deposition modelling and vapor finishing process for manufacturing of ABS plastic parts, *Materials*, Vol. 13, No. 22, Article No. 5176, doi: [10.3390/ma13225176](https://doi.org/10.3390/ma13225176).
- [17] Patel, P.B., Patel, J.D., Maniya, K.D. (2018). Application of PSI methods to select FDM process parameter for polylactic acid, *Materials Today: Proceedings*, Vol. 5, No. 2, 4022-4028, doi: [10.1016/j.matpr.2017.11.662](https://doi.org/10.1016/j.matpr.2017.11.662).
- [18] Madić, M., Antucheviciene, J., Radovanović, M., Petković, D. (2017). Determination of laser cutting process conditions using the preference selection index method, *Optics & Laser Technology*, Vol. 89, 214-220, doi: [10.1016/j.optlastec.2016.10.005](https://doi.org/10.1016/j.optlastec.2016.10.005).
- [19] Yadav, S., Pathak, V.K., Gangwar, S. (2019). A novel hybrid TOPSIS-PSI approach for material selection in marine applications, *Sādhanā*, Vol. 44, Article No. 58, doi: [10.1007/s12046-018-1020-x](https://doi.org/10.1007/s12046-018-1020-x).
- [20] Trung, D.D., Thinh, H.X. (2021). A multi-criteria decision-making in turning process using the MAIRCA, EAMR, MARCOS and TOPSIS methods: A comparative study, *Advances in Production Engineering & Management*, Vol. 16, No. 4, 443-456, doi: [10.14743/apem2021.4.412](https://doi.org/10.14743/apem2021.4.412).

Integrated optimization of line planning and timetabling on high-speed railway network considering cross-line operation

Wang, R.X.^{a,b}, Nie, L.^{a,b}, Fang, W.^c, Ren, H.Q.^d, Tan, Y.Y.^{a,b,*}

^aSchool of Traffic and Transportation, Beijing Jiaotong University, Beijing, P.R. China

^bFrontiers Science Center for Smart High-speed Railway System, Beijing Jiaotong University, Beijing, P.R. China

^cChina State Railway Group Co.,Ltd., Beijing, P.R. China

^dBeijing-Shanghai High Speed Railway Co.,Ltd., Beijing, P.R. China

ABSTRACT

With the implementation of cross-line operation in high-speed railway system, accessibility of cross-line passengers on the railway network has substantially improved. However, due to limitations of capacity, it is hard to schedule a conflict-free timetable based on a train line plan with many cross-line trains. In order to generate a feasible and optimal transportation plan, a novel methodology is introduced. The approach can simultaneously optimize both the line plans of cross-line trains and train timetable, aiming at having a trade-off between operating profit and direct service. Based on an event-activity network framework, a mixed integer programming model is established. Considering service quality would decline after optimizing line plan and train schedules, the objective of the model is set to minimize deviations from ideal schedules for main-line trains while maximize direct service frequency for cross-line passengers. To solve large-scale scenarios efficiently, an enhanced heuristic genetic algorithm is developed. Two smaller-scale cases are devised to validate the efficiency of the model and approach. Finally, the model and the algorithm are applied to a real-world scenario involving the Beijing-Shanghai High-speed Railway and its connecting lines. Also, comparative experiments, including a scenario without cross-line optimization, are conducted to evaluate the advantage of the proposed approach. The result shows the approach can help to quickly find a feasible solution and have a good balance between operating profit and passenger demand.

ARTICLE INFO

Keywords:

Railway network;
Optimization;
Train line planning;
Timetable scheduling;
Cross-line operation;
Passenger demand;
Origin-destination direct service frequency;
Genetic algorithm

*Corresponding author:

yytan@bjtu.edu.cn
(Tan, Y.Y.)

Article history:

Received 21 December 2023

Revised 15 May 2024

Accepted 17 May 2024



Content from this work may be used under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

High-speed railways have garnered increasing attention because of their advantages of speed, safety, convenience, and punctuality. With the large-scale construction of railway infrastructure and the continuous growth of high-speed railway passenger flow, the number of high-speed trains operating on the railway network is increased year by year. To meet the long-distance direct demands of passengers, railway departments need to operate cross-line trains with origin-destination stations (OD) on two different railways. However, this makes railway operators in a dilemma. On the one hand, because of the long-distance causing the small adjustment time domain of departure time at origin station, when many cross-line trains are operated, the travel time domain of main-line trains will be occupied by cross-line trains, and thus it is hard to operate trains and would have an effect on travel speed of main-line trains, and decrease railway line capacity. On

the other hand, significantly reducing the number of cross-line trains makes it difficult to meet the direct travel needs of cross-line passengers, causing a large number of cross-line passengers transfer at transfer stations. Therefore, it is particularly important to optimize both train line plan and train timetable of main-line and cross-line trains comprehensively to adequately meet direct cross-line passenger demand. The problem is difficult but must be solved for well railway operation.

In this paper, we proposed a methodology to solve the problem of integration optimization train line plan planning and timetabling. Fig. 1 shows the framework of the methodology in this paper. In the model, the initial train line plan and ideal timetable provide input parameters. A timetable with ideal traveling times and ideal dwell times is termed an ideal timetable. In an ideal timetable, trains may violate track capacity constraints [1]. Because there are conflicts existing in the ideal timetable, we want to optimize train line plan for cross-line trains and the timetable to eliminate conflicts. The objectives are set to both minimize adjustment time and maximize OD direct frequency. In the methodology, we adjust origin-destination stations of some cross-line trains to change their line plan, but this will reduce some OD direct service frequency of cross-line passenger flow. Therefore, we add some short-distance trains to maintain service levels, allowing passengers to transfer between trains at stations where the line plan is changed. Based on the new line plan, we then optimize train arrival and departure times at stations considering basic train operation constraints. After several iterations, a satisfying result can be obtained.

The contribution of our paper can be summarized in three points. Firstly, we adopt a mixed integer programming model based on event-activity network to formulate train line plan and timetable collaboratively considering cross-line operation. This model will change train line plan of cross-line trains at some stations and optimize train timetable. Secondly, we develop a genetic algorithm to solve a very-large-scale problem. We can obtain an optimized train line plan and train timetable during each iteration. Thirdly, case studies based on the high-speed railway network composed of Beijing-Shanghai high-speed railway and its connecting lines are designed by using the proposed model and algorithm.

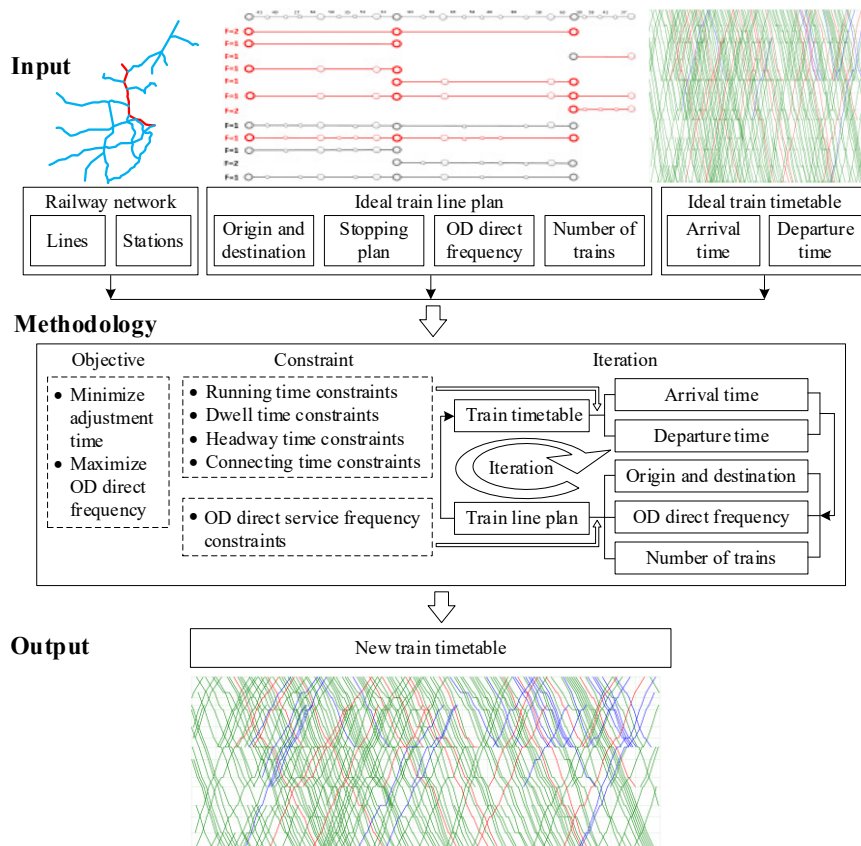


Fig. 1 Framework of integrated optimization methodology of train line plan planning and timetable scheduling

The structure of our paper is organized as follows. In the next section, we provide a summary of related work. Section 3 introduces a mixed integer programming model based on an event-activity network to describe the question. In Section 4, we develop a heuristic genetic algorithm to address the problem. Section 5 presents tests and case studies by taking Beijing-Shanghai high-speed railway and its connecting lines as an example to validate our formulas and algorithm. Finally, discussions and conclusions will be found in Section 6.

2. Literature review

2.1 Cross-line operation

With the construction of high-speed railway, the integrated operation of high-speed trains has become an inevitable trend. Many researchers studied the transportation organization modes for cross-line passengers. Wang *et al.* [2] analysed transportation modes of cross-line passenger flow, and put forward a combination mode with trains running on each line and passenger transfer. Because there were many off-line trains on Wuhan-Guangzhou high-speed corridor, which would have a great impact on capacity, Lei *et al.* [3] analysed the transportation organization mode of off-line trains on main line, and discovered that when off-line trains constitute 65 % of the traffic, capacity is most affected.

However, there were only a few researchers studying cross-line train organization problems, and most of which is metro research. Yang *et al.* [4] studied the cross-line train delays in urban rail transit, analysed cross-line train connection, and developed an optimization model aimed at curtailing delay times. Yang *et al.* [5] introduced a mixed integer non-linear programming model with the primary objective of maximizing passenger travel time while minimizing operation costs to optimize cross-line and main-line train planning, such as frequency, operation zone and stopping pattern. Still other researchers did some research on train transportation organization of main line with taking cross-line trains into consideration. Peng *et al.* [6] considered cross-line train timetable when scheduling trains, and put forward a train scheduling model aiming at minimizing train departure and arrival deviation time. Zhan *et al.* [7] focused on high-speed train stop planning when occurring a major disruption, and proposed a mixed integer programming model aimed at mitigating the frequency of train cancellations and deviations. In case study, the authors also considered the situation of cross-line trains.

2.2 Train line planning and timetable scheduling

There are two fundamental challenges in railway operation, that are train line planning operating at a strategic level, and timetable scheduling operating at a tactical level. Train line plan is the input of the train timetable, and it has a direct impact on which trains passengers choose to take. Chang *et al.* [8] devised a model with multi-objectives aimed at minimizing both the operation cost and the running time loss to solve the optimal allocation problem of high-speed train line plan. Goossens *et al.* [9] introduced a stop optimization model minimizing passenger travel time for solving line planning problems, allowing for the selection of the optimized stop combination from a predetermined stop pattern set. To meet the predicted passenger transport demand, Fu *et al.* [10] proposed a bi-level optimization model for determining line plans with the purpose of minimizing passengers' travel time at upper level while maximizing served passengers at lower level. They develop two consecutive stages heuristic algorithms corresponding to each classification. Parbo *et al.* [11] devised a bi-level programming model to present the skip-stop dilemma, and presented a heuristic approach suitable for large-scale networks. Shang *et al.* [12] introduced a skip-stop strategy, devised a space-time-state framework predicated upon multiple commodity problems, and used Lagrangian relaxation algorithm to dissect the primary problem into several subproblems.

Train timetable not only gives railway operators scheduled paths and times, but also provides different travel schemes for passengers. Therefore, a high-quality timetable can provide better service to passengers and generate more profits for the operators. Zhang *et al.* [13] introduced a minimum cycle time calculation model rooted in the PESP framework for macro-level capacity

assessment and train scheduling, Additionally, they devised an iterative approximation technique aimed at enhancing solution efficiency. Gong *et al.* [14] examined the train timetabling problem accounting for the dynamics and randomness of passenger demand, and formulated a model with the aim of minimizing the expected service cost. Zhang *et al.* [15] devised a model minimizing total train travel time to solve integrated optimization of both maintenance planning and train timetabling. Aken *et al.* [16] considered the challenge of train timetable adjustments by developing a mathematical model. They applied three network aggregation techniques and introduced flexible short-turning possibilities to solve large-scale problems. Cacchiani *et al.* [17] presented a model and designed a heuristic iterative algorithm to obtain an optimal train timetable.

Usually, these two problems are solved successively; however, in recent years, many researchers shifted their focus towards the integrated optimization of train line planning problem and timetabling problem. In order to optimize both train departure schedules and stopping patterns, Yue *et al.* [18] established a mixed integer linear model rooted in a time-space network, aiming to maximize train profits, and proposed a column generation algorithm to effectively resolve this model. Yan *et al.* [19] proposed two models including a line planning model and a timetabling model respectively, and designed a combined method with taking travel time and timetable robustness into consideration. These two models could work iteratively under the designed feedback constraints. Meng *et al.* [20] introduced an integrated team-based model to determine train stopping pattern and timetable, which was aimed at maximize the transport profit and focused on passenger responses to service intervals, stop plans, train arrival/departure times, and the infrastructure capacity. Dong *et al.* [21] established a mixed-integer nonlinear model aimed at optimizing both stop plan and timetable comprehensively. The model's objective was to minimize train running time, waiting time and delay time. To solve the model efficiently, they devised an extended adaptive large-scale neighbourhood search algorithm. To accurately compute passenger waiting times and optimize train timetable, Niu *et al.* [22] developed a timetable optimization model under predetermined skip-stop pattern and OD passenger demand. Jiang *et al.* [23] proposed a method by adjusting train dwell time and train stop plan to increase more scheduled trains on a line. To address this problem, they introduced a Lagrangian-based heuristic algorithm. Cacchiani *et al.* [24] explored the integrated issue of train timetabling and stopping pattern under uncertain demand, and established mixed integer linear models to obtain resilient solutions. Yang *et al.* [25] introduced a collaborative optimization approach for addressing timetabling and stop planning, and establish a mathematical model with multi-objectives to minimize both dwell time and delays. Burggraeve *et al.* [26] designed an iterative method to integrate timetabling and line planning from scratch with considering passenger robustness, and the goal of this method is to minimize operator cost and travel time, and from which they could generate a robust timetable.

2.3 Summary

Table 1 compares previous studies on integration optimization of timetabling and line planning with our research within some key dimensions, such as objective function, cross-line operation or not, considering OD direct service frequency or not, solution algorithm, case size. The comparison reveals that numerous scholars have delved into the amalgamation of train line planning and timetabling. However, there is none of previous researches studying on cross-line operation, and only one of them considering OD direct service frequency but not cross-line passenger flow's neither. Therefore, our research is innovative, very important and necessary.

Table 1 Overview of contemporary research on the integration of line planning and timetable scheduling

Study	Objective	CLO	ODDF	Solution algorithm	Case size
Yue <i>et al.</i> (2016)	Maximize train profit	No	Yes	Column generation heuristic algorithm	1 line and 220 trains
Yan <i>et al.</i> (2019)	Minimize travel time, empty-seat-hour and the number of lines and overtakings, maximize timetable robustness	No	No	Gurobi	1 line and 25 trains within 3 hours
Meng <i>et al.</i> (2019)	Maximize total transporting profit	No	No	Lagrangian relaxation solution	1 line and 100 trains

Table 1 (Continuation)

Dong <i>et al.</i> (2020)	Minimize passenger waiting time, delays and running time	No	No	Adaptive large neighborhood search metaheuristics	1 line and 36 trains
Niu <i>et al.</i> (2015)	Minimize passenger waiting time	No	No	GAMS	1 line and 73 trains
Jiang <i>et al.</i> (2017)	Maximize the total profit	No	No	Heuristic algorithm	1 line and 346 trains
Cacchiani <i>et al.</i> (2020)	Minimize train travel times	No	No	CPLEX	1 line and 36 trains
Yang <i>et al.</i> (2016)	Minimizing dwell time and delay time	No	No	CPLEX	line and 96 trains
Burggraeve <i>et al.</i> (2017)	Minimize operator cost and passenger travel time	No	No	Heuristic algorithm	Railway network with 7 line and 30 trains
This paper	Minimize adjustment time of main-line trains, maximize OD direct service frequency of cross-line passenger flow	Yes	Yes	Heuristic algorithm based on genetic algorithm	Railway network with 9 high-speed rail lines and 184 trains

Note: CLO = Cross-line operation; ODDF = OD direct service frequency;

3. Mathematical formulas

3.1 Problem statement

Our investigation focuses on a double-track high-speed rail network, and trains traversing in one direction entirely independent from trains operating in opposite direction. Therefore, we only consider train operation in one direction. In addition, we assume that there are adequate number of tracks in stations which can be used by trains to have stops.

Fig. 2 shows a simple high-speed railway network, which consists of 2 rail lines and 5 stations. Line 1 comprises four stations, while Line 2 consists of two stations. Station C is the connecting station of these two lines. 6 trains run on this railway network, including 2 main-line trains (T1 and T2) running from station A to station D and 4 cross-line trains (T3, T4, T5 and T6) running from A to E. The initial line plan and ideal train timetable of these simple network are as shown in Fig. 3. In ideal timetable, each train is operated at the time where the maximum benefit can be obtained. Because of this, usually there may exist many conflicts in an ideal timetable. In Fig. 3, it is obvious that there are two headway conflicts. Because of small adjustment time domain of long-distance cross-line trains and tough constraints between trains, it is hard to only adjust less arrival and departure times to make timetable feasible. Fig. 4 shows an example. In Fig. 4, all train paths are rescheduled except train T3's.

Therefore, we proposed an innovative methodology to decrease the repercussions of cross-line operation on main-line trains. Initially, our methodology, the origin or destination stations of some cross-line trains are changed. Then, some short-distance trains are added in the slot of timetable to make up the lost of train service. Finally, the new timetable is rescheduled. And Fig. 5 shows the changed train line plan and rescheduled train timetable of the small network. We change the destination station of train T3 from station E to station B and add a new train T7 to execute the lost train line plan. Now the railway network accommodates seven operational trains, including 3 main-line trains (T1, T2 and T3) and 4 cross-line trains (T4, T5, T6 and T7). In the changed train line plan, cross-line passenger flow from station A to station D can take train T3 and transfer to train T5 or train T7 to accomplish their travel. In the rescheduled train timetable, only train path of T1 is rescheduled, and train T7 is added on the slot of the timetable.

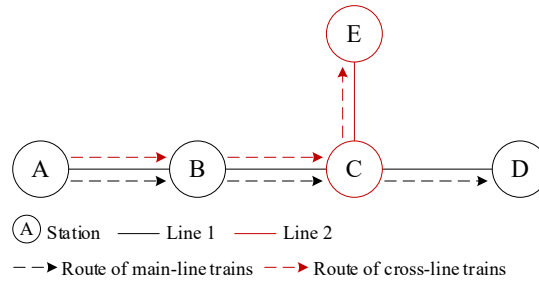


Fig. 2 A simple high-speed railway network

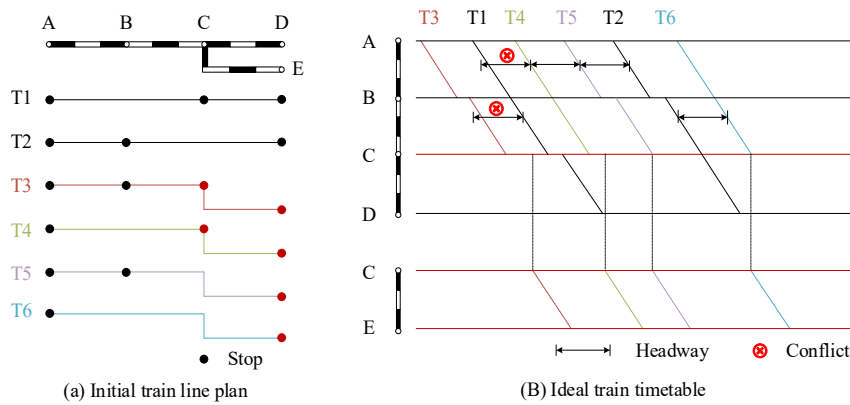


Fig. 3 Initial train line plan and ideal train timetable of the simple network

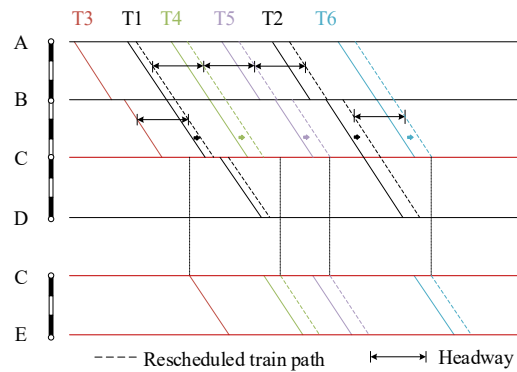


Fig. 4 An example for only adjusting arrival and departure times

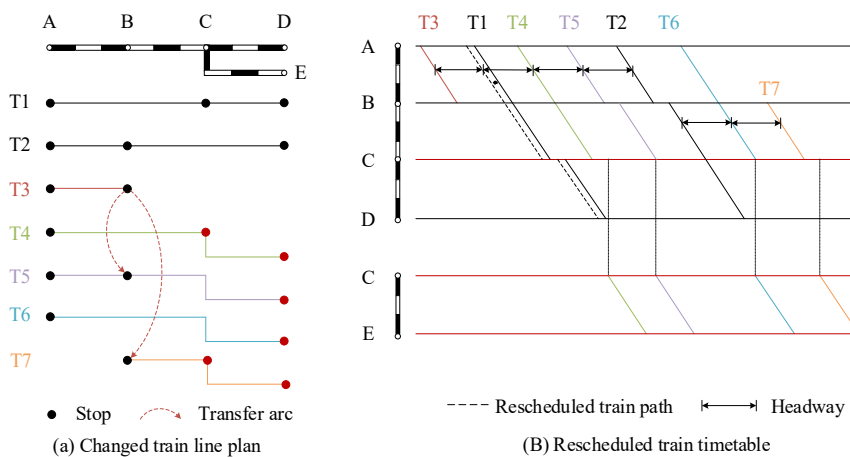


Fig. 5 Changed train line plan and rescheduled train timetable of the simple network

After the change of train line plan of cross-line trains, OD direct service frequency of cross-line passenger flow is decreased, causing passenger direct service demand cannot be well satisfied. For example, in the simple network, the initial direct service frequency from A to E and from B to E in Fig. 4 are 4 and 2, respectively. After applying the proposed methodology and changing the train line plan, we can find that in Fig. 5, these two direct service frequencies become 3 and 2, and passengers who travel from station A to station E can choose to take train T3 and transfer to train T5 or T7 at station B to accomplish their travel. Therefore, the proposed methodology should not only eliminate conflicts of timetable, but also decrease OD direct service frequency of passenger flow as less as possible.

3.2 Event-activity network

The proposed methodology is modelled on event-activity network. In the event-activity model $G = (E, A)$, events are regarded as nodes and activities are directed arcs from one event to another. In our problem of train operation organization, an event i is either a train departure event E_{dep} or a train arrival event E_{arr} , which is related to train index t_i , station index s_i , and scheduled time x_i ; while an activity connecting two events is modelled as a constraint, and there are many types of activities, including train trip activities A_{trip} , train dwell activities A_{dwell} , train headway activities A_{head} and train connection activities A_{con} , which are constructed as follows.

a. $A_{trip} = \{(i, j) | i \in E_{dep}, j \in E_{arr}, t_i = t_j, s_i = s_j - 1\}$. A train trip activity is characterized a trip of one train leaving from an upstream station to next station.

b. $A_{dwell} = \{(i, j) | i \in E_{arr}, j \in E_{dep}, t_i = t_j, s_i = s_j\}$. A train dwell activity is described a train stop at a station.

c. $A_{head} = \{(i, j) | i \in E_{dep}(E_{arr}), j \in E_{dep}(E_{arr}), t_i \neq t_j, s_i = s_j\}$. A train headway activity is described a departure or arrival interval between two different trains departing or arriving at stations.

d. $A_{con} = \{(i, j) | i \in E_{arr}, j \in E_{dep}, t_i \neq t_j, s_i = s_j\}$. A train connection activity is described a connection of two different trains connecting at the connection station where passengers can transfer from one train to the other train.

Fig. 6 illustrates the event-activity graph of some trains in Fig. 5.

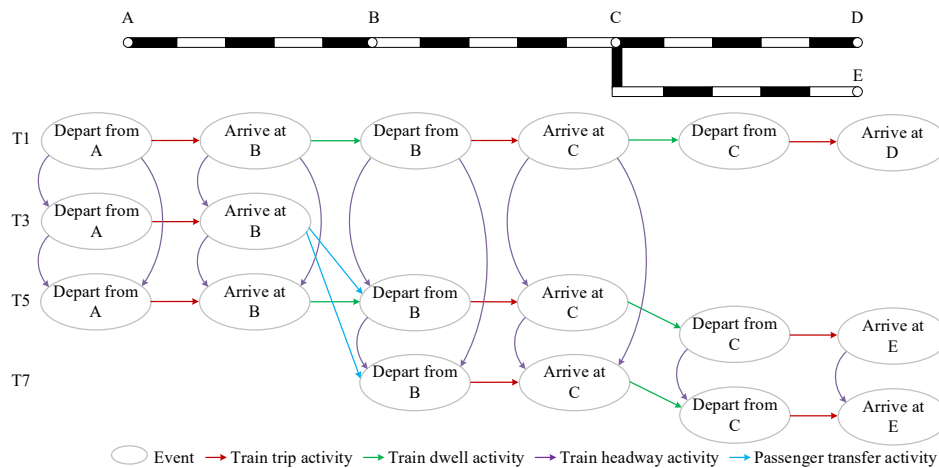


Fig. 6 Event-activity graph of some trains in Fig. 5

3.3 Notations

Table 2 presents indices, sets, parameters and variables integral to the proposed model.

Table 2 Definition of indices, sets, parameters and variables

Notation	Definition
Index	
i, j, i', j', k	Event index, $i, j, i', j', k \in E$
e	Activities index, $e = (i, j) \in A$
t_i	Train index, which denotes the train associated with event i , $t_i \in T$
s_i	Station index, which denotes the station associated with event i
Set	
E	Set of events, $E = E_{arr} \cup E_{dep}$
A	Set of activities, $A = A_{trip} \cup A_{dwell} \cup A_{head} \cup A_{con}$
T	Set of trains, $T = T_m \cup T_c$
T_m	Set of main-line trains
T_c	Set of cross-line trains
S_m	Set of stations on main line
S_i	Set of stations on cross lines where cross-line train t_i passes through
Parameter	
ori_i	Origin station on main-line of train t_i
des_i	Destination station on main-line of train t_i
ε	The maximum adjustable time of cross line trains at origin station
π_i	The scheduled departure and arrival time instant of main-line train t_i at station s_i in the ideal train timetable
c_i	The ideal time instant when cross-line train t_i departs at origin station s_i
λ_i	1 or 0, indicating whether train t_i stops or passes at station s_i
M	A sufficiently large positive integer
$trip_e^{min}$	The lower limit of running time for trip activity e , $e \in A_{trip}$
$trip_e^{max}$	The upper limit of running time for trip activity e , $e \in A_{trip}$
$dwell_e^{min}$	The lower limit of dwell time for dwell activity e , $e \in A_{dwell}$
$dwell_e^{max}$	The upper limit of dwell time for dwell activity e , $e \in A_{dwell}$
con_e^{min}	The lower limit of connecting time for connection activity e , $e \in A_{con}$
con_e^{max}	The upper limit of connecting time for connection activity e , $e \in A_{con}$
h_e	The minimum headway time of headway activity e , $e \in A_{head}$
td_i	The additional starting time for train t_i departing from station s_i
ta_i	The additional stopping time for train t_i arriving at station s_i
tws	The starting time instant of maintenance time
twe	The ending time instant of maintenance time
ξ	The discount of OD direct frequency of cross-line passenger flow
α, β	Weights of two sub-objectives
Variable	
x_i	The time instant when event i occurs in the revised train timetable, namely, when train t_i departs or arrives at station s_i in the new train timetable
ad_i	The adjustment of ideal event i , $i \in E$
θ_{ij}	1 or 0, indicating whether event i precedes or succeeds event j in the revised train timetable
ρ_i	1 or 0, indicating whether origin or destination station of cross-line train t_i was changed to station s_i

3.4 Model formulation

The mixed integer programming model has 2 part of objectives. In order to decrease the effect of cross-line operation on main-line trains, the first objective is to minimize adjustments to departure and arrival times of main-line trains, which ensures the obtained new timetable of main-line trains deviates from the ideal train timetable as little as possible, as expressed in Eq. 1. Moreover, altering the line plan of cross-line trains may lead to a reduction in the direct service frequency for cross-line passenger flow. To satisfy travel demand of cross-line passenger flow, the second objective is to maximize the OD direct service frequency after the adjustment of line plan of cross-line trains, as expressed in Eq. 2.

$$\min z_1 = \sum_{i \in E, t_i \in T_m} ad_i \quad (1)$$

$$\max z_2 = \sum_{t_i=t_j \in T_c, s_i \in S_m, s_j \in S_j} \lambda_i \cdot \lambda_j \cdot \left(1 - \sum_{t_i=t_k, s_k \in S_m} \rho_k \right) \quad (2)$$

We give different weights to these two sub-objectives and transform them into one objective, as expressed in Eq. 3.

$$\min z = \alpha \cdot z_1 - \beta \cdot z_2 \quad (3)$$

There are some constraints must be considered, which are listed in Eqs. 4-15, including train timetable scheduling constraints and train line planning constrains.

(1) Train traveling time constraints between two neighboring stations

Traveling time for a train journey between two neighboring stations is not a certain interval but a flexible time ranging from $trip_e^{min}$ to $trip_e^{max}$, which is expressed in Eq. 4.

$$\begin{aligned} trip_e^{min} + \lambda_i \cdot td_i + \lambda_j \cdot ta_j \leq x_j - x_i \leq trip_e^{max} + \lambda_i \cdot td_i + \lambda_j \cdot ta_j \\ \forall e = (i, j) \in A_{trip}, i, j \in E \end{aligned} \quad (4)$$

(2) Train dwell time constraints at stations

Trains are mandated to pause at stations as designated by the line plan, that is $\lambda_i = 1$. Besides, train dwell time is also flexible. The constraints are expressed in Eq. 5.

$$dwell_e^{min} \leq x_j - x_i \leq dwell_e^{max}, \forall e = (i, j) \in A_{dwell}, i, j \in E \quad (5)$$

(3) Train headway time constraints between two arbitrary trains

The minimum headway can not only ensure safety operation of all trains, but also save occupation time of timetable and improve railway line capacity to some degree. The constraints is expressed in Eq. 6 and Eq. 7. In addition, trains cannot be overtaken by other trains on sections, but only at station, which is expressed in Eq. 8.

$$x_j - x_i + M \cdot (1 - \theta_{ij}) \geq h_e, \forall e = (i, j) \in A_{head}, i, j \in E \quad (6)$$

$$x_i - x_j + M \cdot \theta_{ij} \geq h_e, \forall e = (i, j) \in A_{head}, i, j \in E \quad (7)$$

$$\theta_{ij} = \theta_{i'j'}, \forall (i, j), (i', j') \in A_{head}, (i, i'), (j, j') \in A_{trip} \quad (8)$$

(4) Train connecting time constraints

If origin or destination station of one cross-line train is changed, some cross-line passenger flow necessitates transferring from one train to another train. Train connecting time constraints describe passenger transfer situations, as expressed in Eq. 9. At the same time, the new origin or destination station of the cross-line train should coincide with a stop on the original train's route, and the count of times a cross-line train changing its origin or destination station should less than 1, which is expressed in Eq. 10 and Eq. 11.

$$con_e^{min} \leq x_j - x_i \leq con_e^{max}, \forall e = (i, j) \in A_{con}, i, j \in E \quad (9)$$

$$\rho_i \leq \lambda_i, \forall t_i \in T_c, s_i \in S_m \quad (10)$$

$$\sum_{s_i} \rho_i \leq 1, \forall t_i \in T_c, s_i \in S_m \quad (11)$$

(5) Adjustment constraints of ideal schedules

The Adjustment constraints record the magnitude of left or right shifts ad_i of every main-line train event i . Meanwhile, we should guarantee the time instants of cross-line train departure event i at origin station within a certain range. These constraints are expressed in Eqs. 12 and 13.

$$ad_i \geq |x_i - \pi_i|, \forall t_i \in T_m, i \in E \quad (12)$$

$$\pi_i - \varepsilon \leq x_i \leq \pi_i + \varepsilon, \forall t_i \in T_c, s_i = ori_i \quad (13)$$

(6) OD direct service frequency constraints of cross-line passenger flow

The OD direct service frequency of cross-line passenger flow would be decreased after changing train line plan of cross-line trains, therefore, to satisfy travel demand, the model should ensure the OD direct service frequency of cross-line passenger flow above a certain level, so these constraints can be expressed in Eq. 14.

$$\sum_{t_i, s_i, s_j} \lambda_i \cdot \lambda_j \cdot \left(1 - \sum_{s_k} \rho_k \right) \geq \xi \cdot \sum_{t_i, s_i, s_j} \lambda_i \cdot \lambda_j, \forall t_i = t_j = t_k \in T_c, s_i \in S_m, s_j \in S_j, s_k \in S_m \quad (14)$$

(7) Maintenance time constraints

In China, there is a fixed maintenance time on high-speed railway every night. During this time, trains are forbidden to operate on rail lines. Therefore, any event i should not appear during maintenance time, as expressed in Eq. 15.

$$twe \leq x_i \leq tws, \forall i \in E \quad (15)$$

4. Solution approach

In the proposed model, the decision variables include time instants of events, adjustment of events, sequence of events and train line plan of cross-line trains. The integration optimization of timetabling and line planning is very complex, because it is a significant challenge to acquire a feasible initial solution which could meet the proposed constraints. Hence, devising an efficient optimization algorithm holds paramount importance [27-29]. In order to solve large-scale and complex rail network problem, we introduce a heuristic genetic algorithm (GA). The GA has a flexible chromosome encoding mechanism, which has a significant advantage in solving such problems and can achieve better results within a very short time [30-33].

The algorithm proceeds through the following steps.

- Step 1: Input data. Read the information of railway network, initial line plan and ideal train timetable.
- Step 2: Encoding and population initialization. Each gene in the chromosome represents the time instant of departure event i for train t_i at origin station ori_i , and the chromosome consists of n genes, where n denotes the total number of trains including main-line trains and cross-line trains. All genes are decimal coded. The population size is very important, which is usually depends on the size of problem. After generating chromosomes, feasibility of each chromosome should be checked.
- Step 3: Fitness function and reproduction selection. The fitness function is defined as the reciprocal of objective function. Roulette wheel algorithm is employed for selecting new chromosomes. The chromosome with higher fitness is more likely to be chosen.
- Step 4: Crossover and mutation. The single-point crossover is used to swap gene pairs within two linked chromosomes with a probability. The mutation is applied to each gene with a probability, and a novel value is generated within predetermined minimum and maximum thresholds. After crossover and mutation operation, feasibility of each chromosome should be checked.
- Step 5: Termination. If the number of generations reaches at maximum number of generations, output the saved best solution.

5. Case study

In this section, firstly, We evaluate the proximity of the solution derived from the proposed heuristic methodology to the optimal solution, and verify it is a repaid and effective approach. Then, we solve a large-scale case based on actual situation and analyze the solution.

5.1 Test

We establish two small railway networks to verify the proposed model and approach. Railway network 1 contains 2 rail lines, Line 1 and Line 2, where Line 1 is a mine line and Line 2 is a cross line. Line 1 encompasses 23 stations, while Line 2 encompasses 5 stations. 24 trains run on the network, including 8 main-line while 16 cross-line. Railway network 2 contains 3 rail lines, Line 3, Line 4 and Line 5, where Line 3 is a mine railway, Line 4 and Line 5 are two cross lines. There are 23 stations on Line 3, Line 4 features 5 stations, and Line 5 incorporates 10 stations. 60 trains run on the network, including 20 main-line while 40 cross-line. These two networks are shown in Fig. 7.

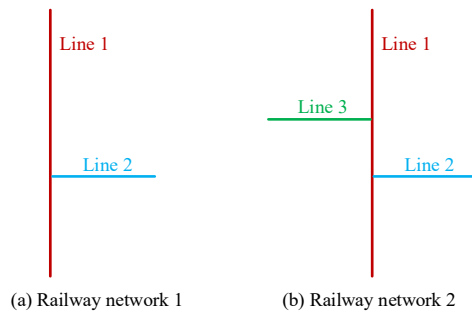


Fig. 7 Two simple railway networks

The model inputs comprise railway network, the initial train line plan and the ideal train timetable. To enhance the realism of the scenarios, we make an assumption that the new origin or destination stations of cross-line trains can be only selected in large stations. The heuristic algorithm is applied. The algorithmic parameters are configured as listed: a population size of 10, both crossover and mutation probability of 0.3, a maximum number of generations is 20. The heuristic genetic algorithm is programmed by using Python 3.7. To ascertain the robustness of the heuristic approach, we initially execute the algorithm 10 times, monitoring the fluctuation of objective values throughout each trial's iterations. Figure 8 illustrates the optimization processes across all 10 trials for both scenarios. All of them can converge at final solutions within 20 generations. We can find that 7 of the 10 trials finally converge at the solution 115 in Fig. 8 (a), and 4 of the 10 trials finally converge at the solution 691 in Fig. 8 (b). Both deviations of final solutions in these two figures are within 6 %. We choose solutions with the lowest objective values as the final results of these two cases.

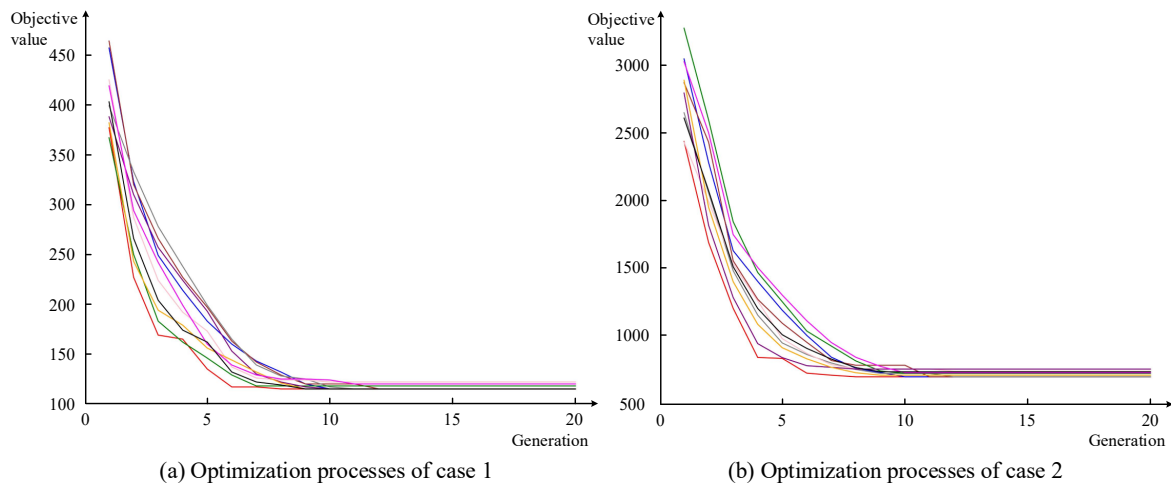


Fig. 8 Optimization processes for all 10 trials of two small cases

Alternatively, we address the mixed integer programming model by optimization software CPLEX 12.10. The Python and CPLEX programs are conducted on a computing system operating on the 64-bit Windows 10 Pro platform, equipped with an Intel(R) Core(TM) i3-10100 CPU @ 3.60 GHz and 16.0 GB RAM.

We compare the outcomes derived from distinct computational methodologies, as shown in Table 3. It is obvious that Scenario 2 and 4 are able to obtain solutions closely to Scenario 1 and 3 with their gaps lower than 1 %. Compared with Scenario 1, there are the same adjustment time of main-line train and the number of changed train line plan of cross-line trains in Scenario 2, but they have difference in OD direct service frequency of cross-line passenger flow. Compared with Scenario 3, there are the same OD direct service frequency for cross-line passengers and alterations in line plan of cross-line trains in Scenario 4, but they have different adjustment time. In general, the application of the suggested heuristic method yields satisfactory outcomes.

However, when the case size (number of rail lines, stations and trains) is very small, there is no advantage for GA approach in calculation time. As the case size escalates, the amount of decision variables and constraints grows multiply, therefore, it takes more and more calculation time by using CPLEX solver, and GA approach has significant advantages when solving large-scale problems. Because of this, we find that calculation time in Scenario 2 is higher than that in Scenario 1, while calculation time in Scenario 4 is lower than that in Scenario 3 on the contrary.

Therefore, we can believe that the proposed heuristic genetic methodology is effective and rapid compared with CPELX solver.

Table 3 Comparison of solutions from CPLEX and GA

Scenario	1	2	3	4
	Case 1 with 24 trains CPLEX	GA	Case 2 with 60 trains CPLEX	GA
Objective	114	115	685	691
Adjustment time	126	126	716	722
OD direct service frequency (initial/new)	14/12	14/11	38/31	38/31
Number of changed train line plan	4	4	9	9
Calculation time	77 s	474 s	3102 s	2176 s
Gap	-	0.8 %	-	0.9 %

5.2 Real-world case

In this section, we deploy the proposed model and solution approach to the Beijing-Shanghai High-speed Railway and its HSR lines across China. Beijing-Shanghai High-Speed Railway (HSR) connects Beijing and Shanghai, connecting 23 stations. Up to now, there are 8 connecting lines of it, which are Tianjin-Shenyang HSR, Shijiazhuang-Jinan HSR, Jinan-Qingdao HSR, Xuzhou-Lanzhou HSR, Hefei-Bengbu HSR, Nanjing-Chongqing HSR, Nanjing-Hangzhou HSR, and Shanghai-Kunming HSR, as shown in Fig. 9.

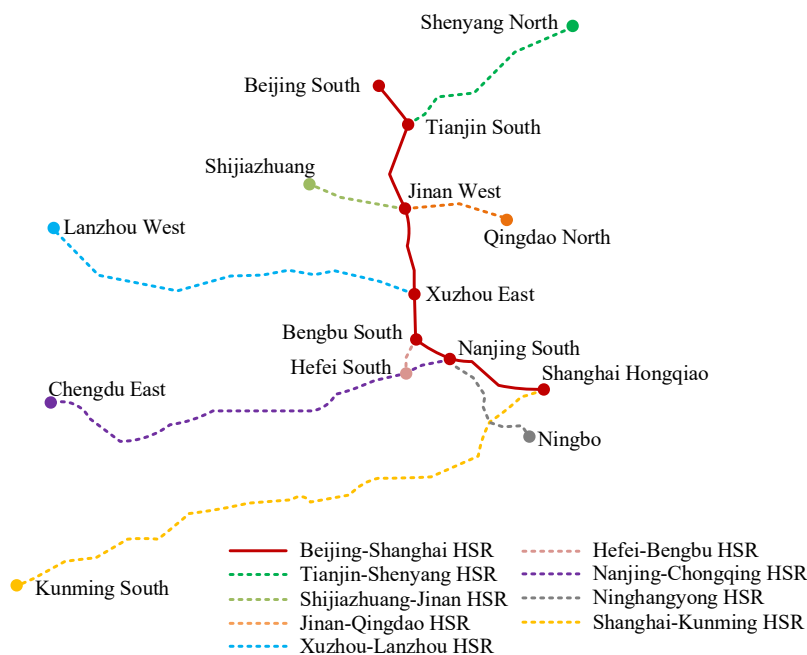


Fig. 9 Beijing-Shanghai High-speed Railway and its connecting lines

Table 4 lists some basic information of lines in this railway network, such as number of stations on each line, line length, and number of operated trains (only trains running in down direction are considered) on each line. It should be noted that trains which we count operated on each connecting line are only the cross-line trains of Beijing-Shanghai High-speed Railway.

Table 4 Basic information of lines in the railway network

Name	Number of stations	Length(km)	Number of operated trains
Beijing-Shanghai HSR	23	1318	Total 184 trains, including 61 main-line trains and 123 cross-line trains
Tianjin-Shenyang HSR	17	671	18
Shijiazhuang-Jinan HSR	11	298	1
Jinan-Qingdao HSR	10	230	26
Xuzhou-Lanzhou HSR	29	1434	24
Hefei-Bengbu HSR	5	132	13
Nanjing-Chongqing HSR	31	1630	13
Nanjing-Hangzhou HSR	11	256	21
Shanghai-Kunming HSR	52	2252	7

We apply the model and approach on the framework of railway network. Also, in order to make cases closer to the actual situation, we make an assumption that the new origin or destination stations of cross-line trains can be only selected in large stations, such as Beijing South, Tianjin South, Jinan West, Xuzhou East, Bengbu South, Nanjing South, and Shanghai Hongqiao. Fig. 10 illustrates 10 iterations of optimization process, all of them can converge at the final solutions within 20 generations, and the objective values vary from 2480 to 2601 within 5 % deviation. We can obtain the final solutions in acceptable calculation time. Additionally, we endeavour to address the problem using CPLEX. However, the model generates an immense number of variables and constraints, causing the solver to grapple with the complexity of such a colossal problem. Consequently, when the RAM reaches full capacity, the software halts without providing a solution.

We designate the solution yielding the minimum objective value as the ultimate outcome for the real-world case. The result shows that the total adjustment time is 2561 minutes, the number of changed train line plan of cross-line trains is 11, and the OD direct service frequency of cross-line passenger flow is 108 in the new train line plan while 117 in the initial train line plan. We analyse the distribution of the adjustments of main-line trains timetables in this case, as depicted in Fig. 11. The visualization demonstrates that 63.4 % of rescheduled timetable's arrival and departure times are equal to those in the ideal timetable, and only 0.4 % of them are adjusted more than 10 minutes. Besides, the line plans of cross-line trains after optimization are listed in Fig. 12.

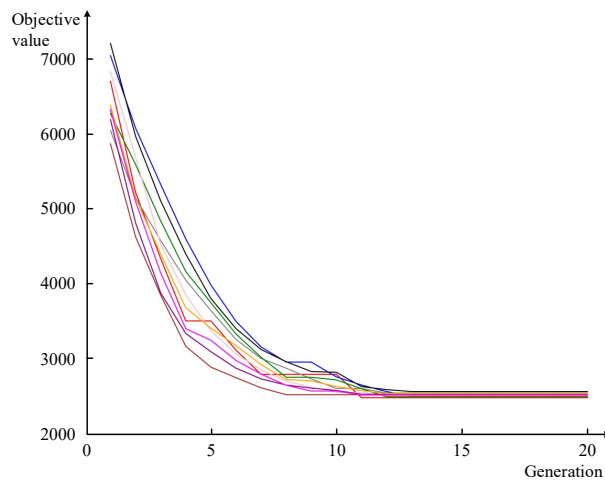


Fig. 10 Optimization process for all 10 trials of the real-world case

We also try to solve the real-world case when the line plans of cross-line trains are not modified, unfortunately, it is greatly difficult to acquire a feasible solution because the conflicts between each train paths cannot be completely resolved.

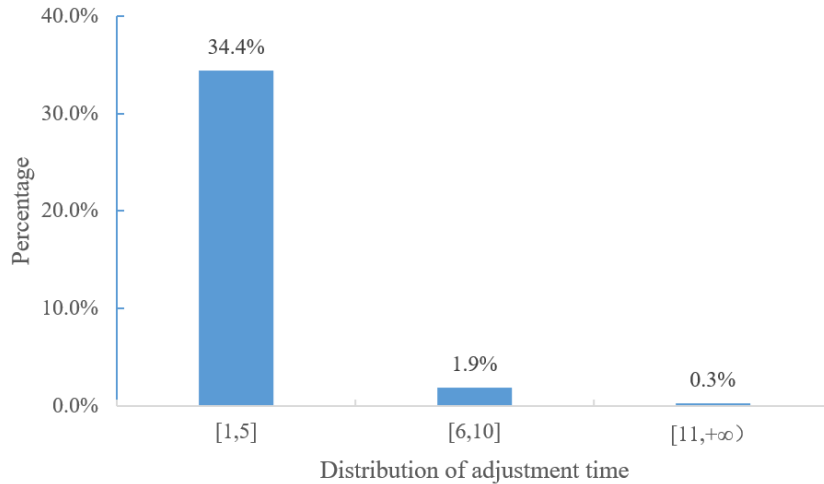


Fig. 11 Distribution of adjustment of main-line trains' departure and arrival times at stations

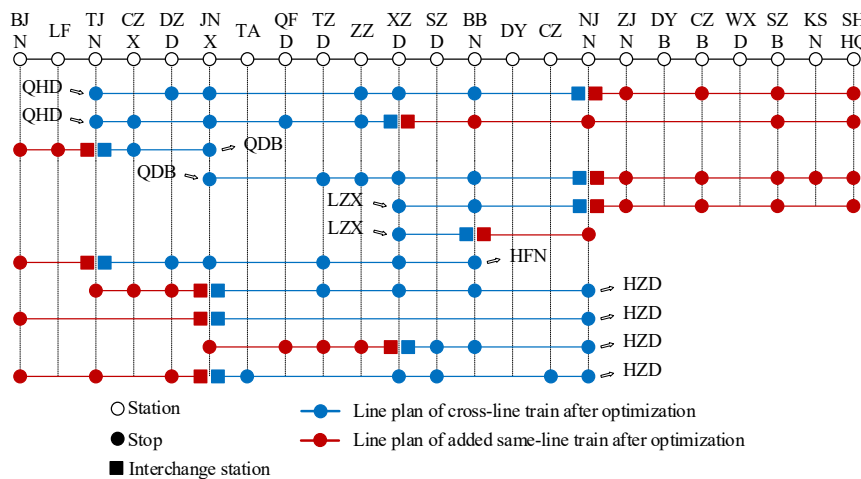


Fig. 12 Line plans of cross-line trains after optimization

6. Conclusion

This paper aims to address the comprehensive optimization challenge of train line planning and timetabling when taking cross-line operation into account. We develop an innovative method to change train line plans of cross-line trains, aiming to decrease the influence of cross-line operation on main-line trains. Leveraging an event-activity network framework, we formulate a sophisticated mixed integer programming model. The model tries to obtain a new timetable with less adjustment time of ideal main-line train timetable. At meantime, to satisfy travel demand of cross-line passengers, the model also takes into account the maximum OD direct service frequency for cross-line passengers. The constraints are formulated based on actual railway operation requirements and passenger demand. An improved heuristic algorithm is proposed to overcome difficulties in calculation efficiency when solving large-scale problems. The approach is based on the genetic algorithm, and firstly is tested on 2 small railway networks with 24 trains running on 2 lines and 60 trains running on 3 lines, respectively. The solutions are compared with those obtained by using CPLEX solver. The findings illustrate that the devised heuristic methodology effectively attains satisfactory solutions which are close to optimal solutions with much less calculation times when solving large-scale problems. Finally, the model and the methodology are deployed in an

actual network with 184 trains operating across 9 high-speed rail corridors, and we obtain a satisfactory solution. In the future, we will do more researches on cross-line operations, especially in considering passenger demand, a regional railway network and rolling stock scheduling.

Acknowledgement

This work was supported by the Fundamental Research Fund for the National Natural Science Foundation of China under Grant No.52002017; the Science and Technology Research and Development Plan Funds for China National Railway Group Co., Ltd. under Grant No.K2023X030, N2023X042; the Fundamental Research Funds for the Central Universities (Science and technology leading talent team project) under Grant No.2022JBQY005; the 111 Project under Grant No.B18004.

References

- [1] Caprara, A., Fischetti, M., Toth, P. (2002). Modeling and solving the train timetabling problem, *Operations Research*, Vol. 50, No. 5, 851-861, doi: [10.1287/opre.50.5.851.362](https://doi.org/10.1287/opre.50.5.851.362).
- [2] Wang, H., Cui, Y., Li, B. (2009). Related issues on train program of passenger special line, In: *Proceedings of the Second International Conference on Transportation Engineering 2009*, Chengdu, China, 410-414, doi: [10.1061/41039\(345\)68](https://doi.org/10.1061/41039(345)68).
- [3] Lei, Z., Yan, H., Jiang, Y. (2019). The transportation organization mode of off-line train in Wuhan-Guangzhou high-speed railway, In: *Proceedings of the Sixth International Conference on Transportation Engineering*, Chengdu, China, 907-914, doi: [10.1061/9780784482742.104](https://doi.org/10.1061/9780784482742.104).
- [4] Yang, X., Hu, H., Yang, S., Wang, W., Shi, Z., Yu, H., Huang, Y., (2020). Train operation adjustment method of cross-line train in urban rail transit based on coyote optimization algorithm, In: *Proceedings of 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, Rhodes, Greece, 1-6. doi: [10.1109/ITSC45102.2020.9294701](https://doi.org/10.1109/ITSC45102.2020.9294701).
- [5] Yang, A., Wang, B., Huang, J., Li, C. (2020). Service replanning in urban rail transit networks: Cross-line express trains for reducing the number of passenger transfers and travel time, *Transportation Research Part C: Emerging Technologies*, Vol. 115, Article No. 102629, doi: [10.1016/j.trc.2020.102629](https://doi.org/10.1016/j.trc.2020.102629).
- [6] Peng, Q., Zhao, J., Xiao, G. (2011). Train scheduling problem in railway passenger dedicated lines, In: *Proceedings of the Third International Conference on Transportation Engineering*, Chengdu, China, 1180-1185, doi: [10.1061/41184\(419\)195](https://doi.org/10.1061/41184(419)195).
- [7] Zhan, S., Wong, S.C., Peng, Q., Lo, S.M. (2021). Train stop deployment planning in the case of complete blockage: An integer linear programming model, *Journal of Transportation Safety & Security*, Vol. 13, No. 10, 1066-1092, doi: [10.1080/19439962.2019.1645773](https://doi.org/10.1080/19439962.2019.1645773).
- [8] Chang, Y.-H., Yeh, C.-H., Shen, C.-C. (2000). A multiobjective model for passenger train services planning: Application to Taiwan's high-speed rail line, *Transportation Research Part B: Methodological*, Vol. 34, No. 2, 91-106, doi: [10.1016/S0191-2615\(99\)00013-2](https://doi.org/10.1016/S0191-2615(99)00013-2).
- [9] Goossens, J.-W., van Hoesel, S., Kroon, L. (2006). On solving multi-type railway line planning problems, *European Journal of Operational Research*, Vol. 168, No. 2, 403-424, doi: [10.1016/j.ejor.2004.04.036](https://doi.org/10.1016/j.ejor.2004.04.036).
- [10] Fu, H., Nie, L., Meng, L., Sperry, B.R., He, Z. (2015). A hierarchical line planning approach for a large-scale high speed rail network: The China case, *Transportation Research Part A: Policy and Practice*, Vol. 75, 61-83, doi: [10.1016/j.tra.2015.03.013](https://doi.org/10.1016/j.tra.2015.03.013).
- [11] Parbo, J., Nielsen, O.A., Prato, C.G. (2018). Reducing passengers' travel time by optimising stopping patterns in a large-scale network: A case-study in the Copenhagen Region, *Transportation Research Part A: Policy and Practice*, Vol. 113, 197-212, doi: [10.1016/j.tra.2018.04.012](https://doi.org/10.1016/j.tra.2018.04.012).
- [12] Shang, P., Li, R., Liu, Z., Yang, L., Wang, Y. (2018). Equity-oriented skip-stopping schedule optimization in an over-saturated urban rail transit network, *Transportation Research Part C: Emerging Technologies*, Vol. 89, 321-343, doi: [10.1016/j.trc.2018.02.016](https://doi.org/10.1016/j.trc.2018.02.016).
- [13] Zhang, X., Nie, L. (2016). Integrating capacity analysis with high-speed railway timetabling: A minimum cycle time calculation model with flexible overtaking constraints and intelligent enumeration, *Transportation Research Part C: Emerging Technologies*, Vol. 68, 509-531, doi: [10.1016/j.trc.2016.05.005](https://doi.org/10.1016/j.trc.2016.05.005).
- [14] Gong, C., Shi, J., Wang, Y., Zhou, H., Yang, L., Chen, D., Pan, H. (2021). Train timetabling with dynamic and random passenger demand: A stochastic optimization method, *Transportation Research Part C: Emerging Technologies*, Vol. 123, Article No. 102963, doi: [10.1016/j.trc.2021.102963](https://doi.org/10.1016/j.trc.2021.102963).
- [15] Zhang, C., Gao, Y., Yang, L., Kumar, U., Gao, Z. (2019). Integrated optimization of train scheduling and maintenance planning on high-speed railway corridors, *Omega*, Vol. 87, 86-104, doi: [10.1016/j.omega.2018.08.005](https://doi.org/10.1016/j.omega.2018.08.005).
- [16] Van Aken, S., Bešinović, N., Goverde, R.M.P. (2017). Solving large-scale train timetable adjustment problems under infrastructure maintenance possessions, *Journal of Rail Transport Planning & Management*, Vol. 7, No. 3, 141-156, doi: [10.1016/j.jrtpm.2017.06.003](https://doi.org/10.1016/j.jrtpm.2017.06.003).
- [17] Cacchiani, V., Furini, F., Kidd, M.P. (2016). Approaches to a real-world train timetabling problem in a railway node, *Omega*, Vol. 58, 97-110, doi: [10.1016/j.omega.2015.04.006](https://doi.org/10.1016/j.omega.2015.04.006).

- [18] Yue, Y., Wang, S., Zhou, L., Tong, L., Saat, M.R. (2016). Optimizing train stopping patterns and schedules for high-speed passenger rail corridors, *Transportation Research Part C: Emerging Technologies*, Vol. 63, 126-146, doi: [10.1016/j.trc.2015.12.007](https://doi.org/10.1016/j.trc.2015.12.007).
- [19] Yan, F., Goverde, R.M.P. (2019). Combined line planning and train timetabling for strongly heterogeneous railway lines with direct connections, *Transportation Research Part B: Methodological*, Vol. 127, 20-46, doi: [10.1016/j.trb.2019.06.010](https://doi.org/10.1016/j.trb.2019.06.010).
- [20] Meng, L., Zhou, X. (2019). An integrated train service plan optimization model with variable demand: A team-based scheduling approach with dual cost information in a layered network, *Transportation Research Part B: Methodological*, Vol. 125, 1-28, doi: [10.1016/j.trb.2019.02.017](https://doi.org/10.1016/j.trb.2019.02.017).
- [21] Dong, X., Li, D., Yin, Y., Ding, S., Cao, Z. (2020). Integrated optimization of train stop planning and timetabling for commuter railways with an extended adaptive large neighborhood search metaheuristic approach, *Transportation Research Part C: Emerging Technologies*, Vol. 117, Article No. 102681, doi: [10.1016/j.trc.2020.102681](https://doi.org/10.1016/j.trc.2020.102681).
- [22] Niu, H., Zhou, X., Gao, R. (2015). Train scheduling for minimizing passenger waiting time with time-dependent demand and skip-stop patterns: Nonlinear integer programming models with linear constraints, *Transportation Research Part B: Methodological*, Vol. 76, 117-135, doi: [10.1016/j.trb.2015.03.004](https://doi.org/10.1016/j.trb.2015.03.004).
- [23] Jiang, F., Cacchiani, V., Toth, P. (2017). Train timetabling by skip-stop planning in highly congested lines, *Transportation Research Part B: Methodological*, Vol. 104, 149-174, doi: [10.1016/j.trb.2017.06.018](https://doi.org/10.1016/j.trb.2017.06.018).
- [24] Cacchiani, V., Qi, J., Yang, L. (2020). Robust optimization models for integrated train stop planning and timetabling with passenger demand uncertainty, *Transportation Research Part B: Methodological*, Vol. 136, 1-29, doi: [10.1016/j.trb.2020.03.009](https://doi.org/10.1016/j.trb.2020.03.009).
- [25] Yang, L., Qi, J., Li, S., Gao, Y. (2016). Collaborative optimization for train scheduling and train stop planning on high-speed railways, *Omega*, Vol. 64, 57-76, doi: [10.1016/j.omega.2015.11.003](https://doi.org/10.1016/j.omega.2015.11.003).
- [26] Burggraeve, S., Bull, S.H., Vansteenwegen, P., Lusby, R.M. (2017). Integrating robust timetabling in line plan optimization for railway systems, *Transportation Research Part C: Emerging Technologies*, Vol. 77, 134-160, doi: [10.1016/j.trc.2017.01.015](https://doi.org/10.1016/j.trc.2017.01.015).
- [27] Rashid, A.M., Midi, H. (2023). Improved nu-support vector regression algorithm based on principal component analysis, *Economic Computation and Economic Cybernetics Studies and Research*, Vol. 57, No. 2, 41-56, doi: [10.24818/18423264/57.2.23.03](https://doi.org/10.24818/18423264/57.2.23.03).
- [28] Shamami, N., Mehdizadeh, E., Yazdani, M., Etebari, F. (2023). A hybrid ba-vns algorithm for solving the weapon target assignment considering mobility of resources, *Economic Computation and Economic Cybernetics Studies and Research*, Vol. 57, No. 3, 59-76, doi: [10.24818/18423264/57.3.23.04](https://doi.org/10.24818/18423264/57.3.23.04).
- [29] Du, H., Zhang, Y., Zhang, L., Chen, Y. (2023). Selective ensemble learning algorithm for imbalanced dataset, *Computer Science and Information Systems*, Vol. 20, No. 2, 831-856, doi: [10.2298/CSIS220817023D](https://doi.org/10.2298/CSIS220817023D).
- [30] Xu, W., Sun, H.Y., Awaga, A.L., Yan, Y., Cui, Y.J. (2022). Optimization approaches for solving production scheduling problem: A brief overview and a case study for hybrid flow shop using genetic algorithms, *Advances in Production Engineering & Management*, Vol. 17, No. 1, 45-56, doi: [10.14743/apem2022.1.420](https://doi.org/10.14743/apem2022.1.420).
- [31] Amjad, M.K., Butt, S.I., Anjum, N., Chaudhry, I.A., Faping, Z., Khan, M. (2020). A layered genetic algorithm with iterative diversification for optimization of flexible job shop scheduling problems, *Advances in Production Engineering & Management*, Vol. 15, No. 4, 377-389, doi: [10.14743/apem2020.4.372](https://doi.org/10.14743/apem2020.4.372).
- [32] Shi, D.L., Zhang, B.B., Li, Y. (2020). A multi-objective flexible job-shop scheduling model based on fuzzy theory and immune genetic algorithm, *International Journal of Simulation Modelling*, Vol. 19, No. 1, 123-133, doi: [10.2507/IJSIMM19-1-CO1](https://doi.org/10.2507/IJSIMM19-1-CO1).
- [33] 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](https://doi.org/10.2507/IJSIMM17(4)CO17).

Experimental and numerical investigation of the deep drawing process using a tractrix die – An industrial case study focused on stress and temperature analysis

Mandic, V.^{a,*}, Milosavljevic, Dj.^b, Jurkovic, Z.^c, Adamovic, D.^a

^aFaculty of Engineering of the University of Kragujevac, Kragujevac, Serbia

^bSloboda a.d., Cacak, Serbia

^cFaculty of Engineering of the University of Rijeka, Rijeka, Croatia

ABSTRACT

The deep drawing process of thick sheet metal for vessel production is carried out by applying a tractrix die with the absence of a blank holder, which has economic benefits for industrial production. The main aim of the paper is the development of a reliable numerical thermo-mechanical model of a silicon brass vessel manufactured by a deep drawing process in a tractrix die and a subsequent ironing process, which includes the previous assembly of the dies with reinforcing rings that creates the required prestresses. The testing of the mechanical properties of silicon brass CuZn24Si was carried out by a standard uniaxial tensile test, thus a flow curve was determined to describe the material behaviour. The initial temperatures of the environment, blank and tools were measured with an infrared thermal imager. A comprehensive finite element stress analysis of the deformable tools was carried out for the assembly phase of the dies, and for workpiece and tools in the deep drawing and ironing processes. The comparison of measured and numerically estimated temperatures had a good agreement, so the developed numerical model was confirmed and validated. This research study demonstrates how different process parameters can be investigated through a reliable and precise numerical model with complementary experimental research for the optimization of industrial technology.

ARTICLE INFO

Keywords:

Deep drawing;
Tractrix die;
Reinforcing rings;
Finite element method (FEM);
Experimental-numerical approach;
Numerical modelling;
Simulation;
Infrared imaging technology;
Simufact.forming

*Corresponding author:

mandic@kg.ac.rs
(Mandic, V.)

Article history:

Received 28 December 2023
Revised 29 February 2024
Accepted 1 March 2024



Content from this work may be used under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

Deep drawing is one of the sheet metal forming processes that is widely used in the automotive, aerospace, and packaging industries for the production of parts from various materials. It most often represents the first sheet metal forming operation where a 3D shape is obtained from a relatively flat sheet metal. It is often combined with an ironing operation (DDI – Deep Drawing and Ironing) when, in order to obtain special vessels and liners from thick sheets, it is necessary to maintain an identical thickness at the bottom while significantly reducing the thickness of the vessel walls [1]. In addition to numerous advantages compared to other technologies, it must be noted that the use of this technology is economically feasible only in the case of large-scale and mass production, as it requires the production of special tools as well as the use of machines with

large nominal forces. One of the ways to successfully perform the first deep drawing operation on thick sheets with less forming loads and radial stresses is to use a tool with a tractrix die and without a blank holder [2-3].

Thick sheet forming processes are considered bulk forming rather than sheet metal forming due to large deformations, temperature increase and the appearance of significant radial stresses, where it is necessary to protect the die with reinforcing rings and apply cooling agents. Understanding the influence of numerous process parameters on the stress-deformation-temperature state in the tool and the workpiece during the process is very important for the prevention of defects (wrinkling, folding etc.), excessive radial stresses and obtaining quality workpiece and finished parts [4-6].

There are numerous studies that apply an integrated experimental-numerical approach to the analysis of metal forming processing [7-14] where reliable numerical models have been developed and verified by industrial case studies. The basic condition for a reliable model is the similarity between the real and the model process, which can only be achieved through experimental laboratory tests and measurements in the industrial process. When the conditions of similarity are proven, then a lot of information and results can be obtained from such a validated numerical model of the process about the coupled influence of the process parameters on the output performance, even outside the scope of industrial application, thus the physics of the process itself is explored more deeply.

The purpose of the paper is to develop a reliable numerical thermo-mechanical finite element (FE) model for the analysis of the deep drawing process using a tractrix die, combined with the subsequent ironing process, to obtain a silicon bronze vessel under controlled temperature conditions, in order to achieve the target structure with a significant reduction of the wall thickness. The overall contribution of this study is reflected in the proposed and demonstrated integrated experimental-numerical approach for achieving the conditions of similarity of real and model processes through the experimental identification of process input parameters and model validation. The uniaxial tensile test was used to determine the mechanical properties of silicon brass and the flow curve in mathematical form, to describe the behaviour of the material during plastic forming in FE numerical simulation. To achieve the required contact friction conditions the blank was subjected to mechanical and chemical treatments described in the paper. Using infrared imaging technology, the temperatures of blanks, tools and the environment were measured before and after the industrial process.

2. Deep drawing process using a tractrix die

The profile of the deep drawing die significantly influences the sheet metal flow, the degree of deformation, the change in forming load depending on the puller stroke and, finally, on the quality of the workpiece and the process itself. Three types of the die profiles are used depending on the type of operation, with a transition radius, with a cone and with a specially designed tractrix curve.

When it comes to thick sheets, the die with the tractrix profile gives the best results, especially for consecutive deep drawing and ironing operations in the same tool, such as the case study in this paper. In such a die, the bending arm of the sheet is constant throughout the drawing process. The normal force always acts on the edge surface of the billet creates the largest bending arm. Considering that the arm h moves along the tractrix curve tangentially, the solution boils down to finding the equation of the tractrix curve with constant tangent length. Fig. 1 shows the profile of the die with the tractrix curve in the Cartesian coordinate system, and the Eq. 1 defines the tractrix curve with a constant tangent length [15]:

$$x = h \cdot \ln\left(\frac{h + \sqrt{h^2 - y^2}}{y}\right) - \sqrt{h^2 - y^2} \quad (1)$$

The weakness of the tractrix profile is that it theoretically has an infinite height ($\lim_{y \rightarrow 0} x \rightarrow \infty$), so it is necessary to adjust the initial value h_0 , so that the deep drawing does not start at the very end of the tractrix curve but moves closer to the x axis. It is recommended that h_0 be slightly larger than the radius r_m for the drawing die with a radius. For the recommended value of the height of the tractrix curve in the die profile, it is necessary to choose the appropriate value of y_0 .

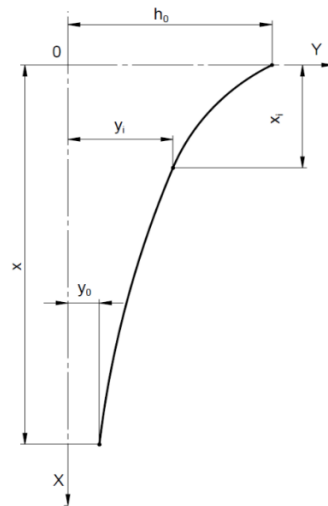


Fig. 1 Die profile with a tractrix curve

Taking into account the billet dimensions used in the industrial case study for the purposes of this paper $\phi 227 \times 15$ mm, and the outer diameter after deep drawing $\phi 163.5$ mm, the value $h_0 = 30$ mm was chosen, so that Eq. 1 has the form:

$$x = 30 \cdot \ln\left(\frac{30 + \sqrt{30^2 - y^2}}{y}\right) - \sqrt{30^2 - y^2} \tag{2}$$

Pairs of values (x, y) for the tractrix profile of the die are shown in Table 1, and the technical drawing of the die is shown in Fig. 2.

Table 1 The coordinates of the points belonging to the tractrix profile

Point	1	2	3	4	5	6	7	8	9	10	11
Y (mm)	30	28	25	22	19	16	13	10	8	6	4
X (mm)	0	0.488	2.064	4.438	7.643	11.82	17.27	24.53	30.92	39.299	51.284

3. Methods and materials

3.1 Integrative experimental-numerical approach

For the analysis of the deep drawing process, an integrative approach that includes experimental and numerical methods was applied, as a proven way to prepare high-quality input data and boundary conditions for the application of the finite element method and its verification, schematically shown in Fig. 3. In this approach, the process must be observed as a system of interconnected relevant parameters that include the following subsystems:

- *Billet*: size and shape, material, chemical composition and microstructure, flow properties under processing conditions (flow stress as a function of deformation, deformation rate and temperature), thermal and physical properties;
- *Tool*: geometry, surface condition, material and hardness, temperature, stiffness and accuracy;
- *Interface conditions*: finalization of surfaces, lubrication, friction and heat transfer;
- *Deformation zone*: mechanics of plastic deformation, material flow, strains, stresses, strain rates and temperatures;
- *Production equipment*: machine type, velocity, maximum load / available energy, stiffness and precision.

Numerical process modelling, supported by experimental data and tests, is used to predict material flow, determine distributions of strains, stresses, temperatures, tool stresses, potential sources of defects and brittle fracture, properties and microstructure of workpiece, as well as assessment of elastic springback and residual stresses.

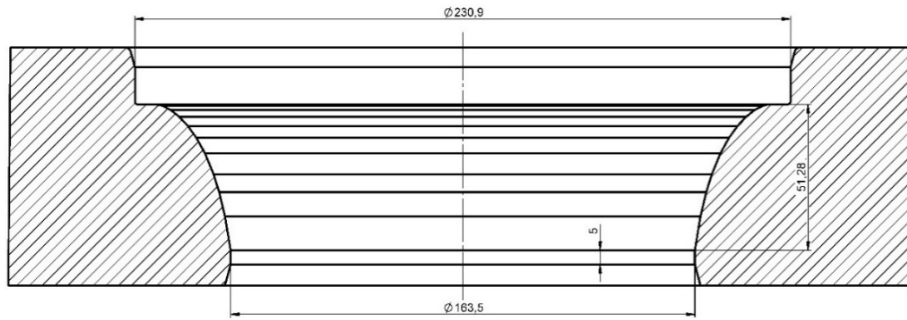


Fig. 2 Technical drawing of the die with a tractrix profile

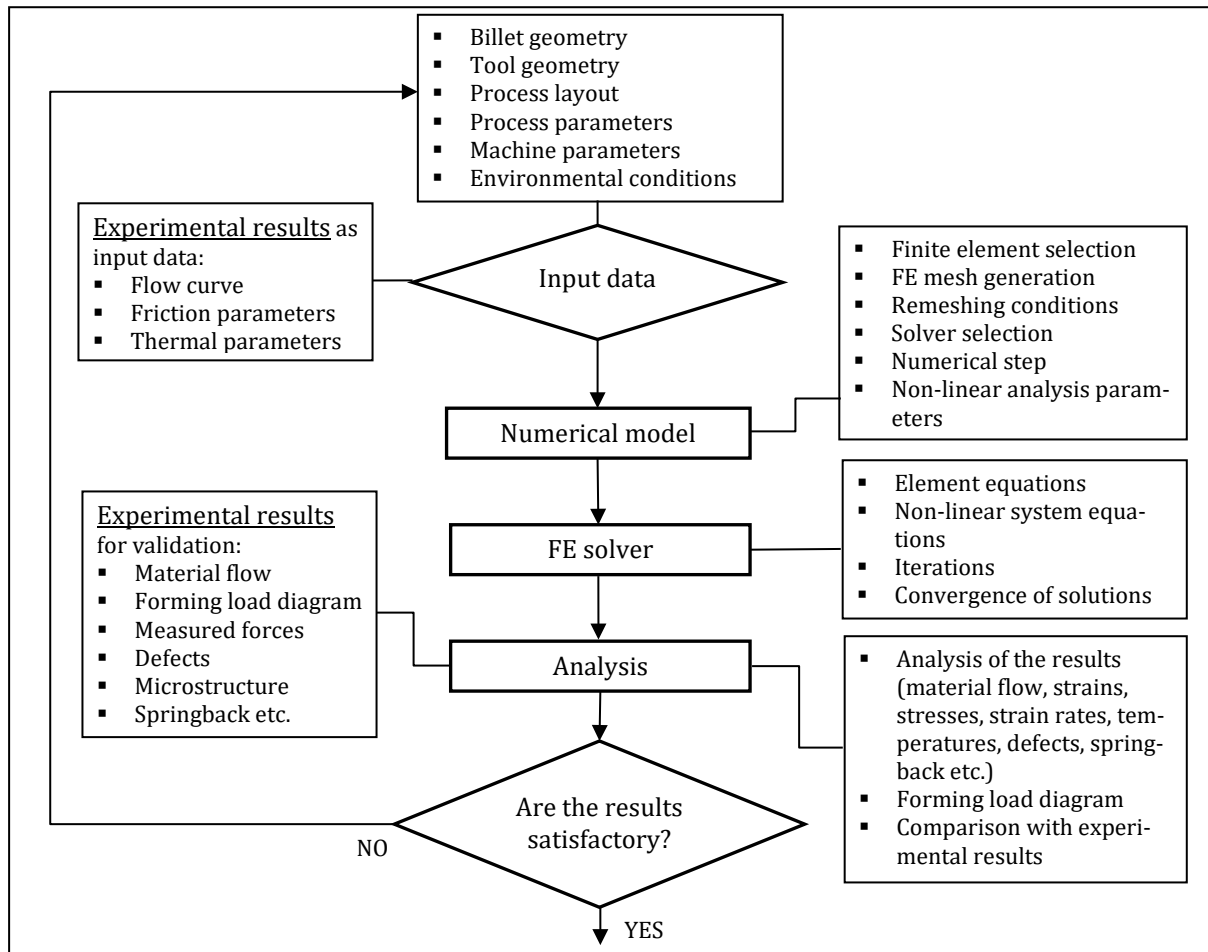


Fig. 3 Integrative experimental-numerical approach in process modelling

For the results of numerical modelling to be reliable, it is necessary to provide high-quality input data that faithfully describe the real industrial process and refer to the geometry of applied tools and billet, and process and environmental parameters. In addition, it is necessary to perform experimental research to determine the true stress – true strain relationship (flow curve) for the materials of the billet and the tool, if the tool stress analysis is taken into account. The conditions of contact friction, heat transfer and radiation are described by appropriate laws for which it is necessary to experimentally determine the friction coefficient/factor, heat transfer coefficient and emissivity coefficient.

In the pre-processing phase a numerical model of the process is developed, using all input data and through the selection of the appropriate finite element, defining the conditions for generating the initial FE mesh, as well as the subsequent remeshing, the selection of the appropriate solver, the number of numerical steps and other parameters for non-linear FE analyses. If all previous

steps are well defined, the FE calculation itself is realized automatically by means of the solver by iteratively solving the nonlinear equations of the system with the convergence of the solution.

The selection of numerical results for the analysis depends on the goals of the process simulation, but those that can be compared with the experimental ones should be chosen, in order to verify the numerical model of the process. For this purpose, laboratory or industrial experiments can be carried out. When the comparative analysis of the results shows an unsatisfactory match or the obtained results are not as expected by the set goals, it is necessary to repeat the simulation with changed input parameters and settings for FE analysis.

3.2 Material properties

One part of the experimental research was related to the characterization of the workpiece material and the determination of flow curves using the uniaxial tension test. The tested material was silicon brass (Plate EN 1652 – CuZn24Si – R500) containing 23.13 % Zn, whose complete chemical composition is given in Table 2.

Table 2 Chemical composition of silicon brass CuZn24Si

Element	Cu	Pb	Fe	Sn	Al	Ni	P	Si
%	75.9	0.004	0.002	0.007	0.003	0.003	0.006	0.940

This chemical composition of brass is suitable for deep drawing of thick sheets (in this case 15 mm) considering the higher percentage of Zn gives better material ductility. The presence of silicon results in an increase in resistance to general and stress corrosion, and, in addition, it has a very favourable effect on the mechanical properties of brass, both at room and elevated temperatures. Previous research on this silicon brass have shown that the best technological properties are obtained if silicon is added to brass with a stoichiometric composition of Cu₃Zn, i.e. to brass with 75 % of Cu, as well as with the addition of silicon over 0.5 % [16]. However, higher percentages of silicon can negatively affect the formability of the material.

Uniaxial tensile tests were performed on a Zwick/Roell Z100 universal material testing machine, which enables automatic data acquisition at a speed of 500Hz. A ceramic extensometer in the range of 11-50 mm was used to measure elongation. The strain, strain at break, yield strength and tensile strength values were calculated using testXpert software and the collected pairs of load and elongation values. A total of 9 specimens were prepared, three for each of the rolling directions (0°, 45° and 90°), due to the reproducibility of the obtained results, and according to the standard ISO 6892-1:2019, with original diameter of 8 mm and original gauge length of 100 mm. Specimens after one series of tests are shown in Fig. 4. Holloman’s hardening law given by Eq. 3 was used for the mathematical formulation of the material's behaviour during deformation:

$$k_f = C \varphi^n, \text{MPa} \tag{3}$$

where k_f is flow stress, C is strength coefficient, φ is true strain, and n is hardening exponent. The obtained mechanical properties of the material are shown in Table 3, as average values for three repeated tests.



Fig. 4 Specimens after uniaxial tension test

Table 3 Mechanical properties of CuZn24Si brass

Rolling direction	Stress (MPa)		Strain		Coefficients in Eq. 3	
	Yield stress	Ultimate stress	Strain (ultimate)	Strain at break	C	n
0°	94.63	558.04	0.5288	0.5647	779.69	0.52886
45°	94.10	552.96	0.5300	0.5683	768.63	0.52591
90°	92.11	554.01	0.5259	0.5540	768.29	0.53004

3.3 Numerical modelling

The application of numerical process modeling is a well-proven and extremely useful tool for predicting problems in industrial production and reducing time and cost in new product development. For metal forming modeling based on the finite element method, it is necessary to apply a nonlinear coupled thermo-mechanical approach, since due to plastic deformation and the influence of contact friction, there is an increase in temperature in the workpiece. In order to realistically describe industrial process, modelling must include, in addition to heat generation in the workpiece, the heat transfer to the tool and its radiation to the environment. Although the considered deep-drawing process takes place at room temperature, due to the bulk forming of the thick sheet, this process can be considered a bulk forming process, in which the temperature in the workpiece can rise by several hundred degrees. Due to all of the above, only a coupled finite element approach can provide reliable numerical modelling results.

In this study, the *Simufact.forming* software was applied as a virtual manufacturing tool, which applies the enhanced version of the *MARC* finite element solver based on the displacement method and the coupled thermal-mechanical approach [17]. The constitutive relations for this approach are defined by Eqs. 4 and 5:

$$K(T) \cdot u = f \quad (4)$$

$$C(T) \cdot \dot{T} + k(T) \cdot T = Q + Q^1 \quad (5)$$

where K is the system stiffness matrix, u is the nodal displacement, f is the force vector, C is the heat capacity matrix, T is the nodal temperature vector, \dot{T} is the time derivative of the temperature, k is thermal-conductivity matrix, Q is the thermal load vector (flux) and Q^1 is the internal heat generated due to plastic deformation. K , k and C are all temperature dependant.

The nodal force vector f in Eq. 4 includes the sum of different types of loads:

$$f = f_{point} + f_{surface} + f_{body} + f^* \quad (6)$$

where f_{point} is the point load vector, $f_{surface}$ is the surface load vector, f_{body} is the volume load vector, and f^* represents all other types of load vectors (for example thermal strains).

In metal forming processes, these equations are non-linear due to the non-linearity of the material and geometry during deformation, as well as due to the constant changes in boundary conditions (contact friction and thermodynamic processes). To solve them, a Newton-Raphson incremental method is used, within the iterative solver. A direct solver can also be used, but it requires more processing time, but offering more accurate solutions.

To establish a connection between the real 3D stress state existing in the workpiece during deformation and the flow curves obtained by the uniaxial tensile test, the Von Mises flow function, usually for plastic materials, is used. According to this criterion, plastic flow of the material occurs when the calculated effective stress reaches the yield strength determined by the tensile test. When it is necessary to include material anisotropy into the modelling process, which is especially significant in thin sheet metals, this software uses the Hill anisotropic yield function.

In a coupled thermo-mechanical numerical approach, the critical input parameters for numerical modelling are mechanical and heat properties of workpiece and tools, material model that includes an elastic-plastic behaviour with strain hardening, initial temperatures, heat transfer and emissivity parameters as well as friction properties for the interface.

No less important is the choice of the appropriate finite element that enables obtaining convergent solutions in the calculation, then automatic generation of the initial FE mesh in the workpiece and tool, and automatic remeshing according to trigger criteria and conditions for mapping parameter values from the old to the new FE mesh. Shell elements are usually used for modelling the sheet forming process, but in the case of thick sheet metal, they do not give satisfactory results. Therefore, in this study, 3D solid isoparametric eight-node hexahedral elements were applied, with automatic detection of sheet thickness and determination of the optimal number of elements through the thickness and the best in-plane element size.

When modelling multistage processes, the virtual manufacturing software should “remember” the entire history of deformation through which the workpiece is transformed, whether the first

stage is related to heating or not, followed by the transfer of workpiece to the tool, in which the workpiece is formed through a series of successive different forming operations, and sometimes interoperational thermal treatments. At the end of forming, the workpiece removal from the tool, the release of elastic strains (springback) should be modelled. During the numerical setup of the process, it is necessary to precisely define the complex kinematics of metal forming machines, because the choice and change of the machine significantly affects the material flow. The applied software for numerical modelling in this study enables all the mentioned requirements.

3.4 Infrared thermal imaging

Infrared thermography is one of the non-contact methods of temperature measurement that is increasingly used for scientific and industrial measurements of temperature fields, especially when contact methods are limited by accessibility or the range of high temperatures in technological processes. This is of particular importance for metal forming processes, because by measuring temperatures in the visible part of the workpiece and the tool itself, places with critical temperature stresses can be identified. When measurements are used to set boundary conditions in numerical modelling or to validate numerical process models, then such an experimental-numerical approach opens up possibilities for numerical estimations of temperature fields outside the experimental range, for example in hot forging processes [18], extrusion at high velocities and industrial ironing processes.

Non-contact methods of temperature measurement are based on the principle of measuring the energy of electromagnetic radiation emitted by the measured object. Infrared radiation is located in the electromagnetic spectrum between visible light and microwave radiation in the range of wavelengths from 0.75 μm to 1 mm. The main advantage of infrared thermography, compared to other methods of temperature measurement, is that a visual image of the temperature field is obtained, where points with maximum or minimum temperature values, as well as their arrangement and mutual dependence, can easily be seen.

Within the thermal interaction of the measured object with the environment, the total energy emitted from the environment is divided into absorbed, reflected and transmitted energy. For measuring the temperature of solid bodies that are impervious to thermal radiation, only the energy emitted by the body is relevant, and it is equal to the absorbed energy. The characteristic by which the radiation of any body is compared with the radiation of a black body is called emissivity expressed through the emission coefficient ε . All bodies have an emissivity greater than 0 (ideal mirror) and less than 1 (black body). The emission coefficient of an object depends on the type of material and the state of the object's surface. For precise measurement of objects made of materials with low emissivity, clean and polished metals with shiny surfaces, emissivity correction is necessary, precisely through the emission coefficient. Before measuring, it is necessary to adjust the emissivity settings in the imager in order to accurately calculate the temperature of the measured object [19].

To calculate the object's temperature, the Stephen–Boltzmann law is applied, which exhibits the relationship between the radiation energy and the temperature by the Eq. 7 [20]:

$$E = \varepsilon \cdot \sigma \cdot T^4 \quad (7)$$

where E is the radiant power of the object; ε is emissivity of material; σ is the Stephen–Boltzmann constant; T is the absolute temperature of the object.

Fluke Ti400 Infrared Thermal Imager and accompanying *SmartView* software were used for experimental measurement of temperature fields in the workpiece and deep drawing tools in this study [21]. The change in emissivity for different materials and conditions of the surface of the object can be changed directly by entering the desired value or using coefficients from the software database. The imager uses manual focusing mode or automatic focusing through the *LaserSharp* system that measures the distance of the object with a laser and the system adjusts the focus for thermal imaging processing. *SmartView* enables the analysis and processing of obtained thermal images for comparative analysis with numerically estimated temperature fields.

4. Industrial case study

4.1 Deep drawing industrial setup

The tool for industrial experiments of the deep drawing process was manufactured by the Sloboda company and mounted on a hydraulic press with a selected deformation speed of 50 mm/s. Fig. 5 shows the technical drawing of the tool assembly. The tool consists of the following components: 1) puller, 2) upper die, 3) intermediate plate, 4) lower die, 5) remover and 6) base plate.

The puller (1), made of alloyed cold-work tool steels 1.2379/X153CrMoV12 (EN ISO 4957:2000) and heat-treated to a hardness of 60HRC, participates in forming the inner shape of the drawn vessel. It has a good ratio of high outer surface hardness and high core durability. The surface of the working part of the puller is polished to the quality of surface treatment N4. In the upper die (2), the workpiece is drawn from the billet into vessel, with a smaller reduction in wall thickness. The material for the upper die is the same as for the puller and lower die. In the lower die (4), the ironing process is carried out to obtain the final shape of the vessel, the thickness and the height of the walls.

Both dies were reinforced with reinforcing rings of 1.7225/42CrMo4 tempering steel (DIN EN 10088) with a 1 mm overlap to achieve prestressing of the dies. Assembly of the rings on the dies was realised by pressing on the press. The dies strengthened in this way are finally mounted in the lower tool holder, as shown in Fig. 6.

4.2 Blank preparation

The circular blank (ϕ 227 mm) was prepared by cutting with a high-pressure water jet from a brass plate of 15 mm thickness (Fig. 7). After cutting, the blank was subjected to strict dimensional and visual control, as well as metallographic examination of the micro and macrostructure. A chemical treatment was applied by immersing the blank in a 10 % solution of sulfuric acid (H_2SO_4) at a temperature of 60-70 °C for 5 minutes. Then the blank was immersed in a 5 % solution of soap emulsion at a temperature of 60 °C. With these procedures, a better plastic flow of the material, a reduction in the impact of contact friction and less heating of the tool were achieved.

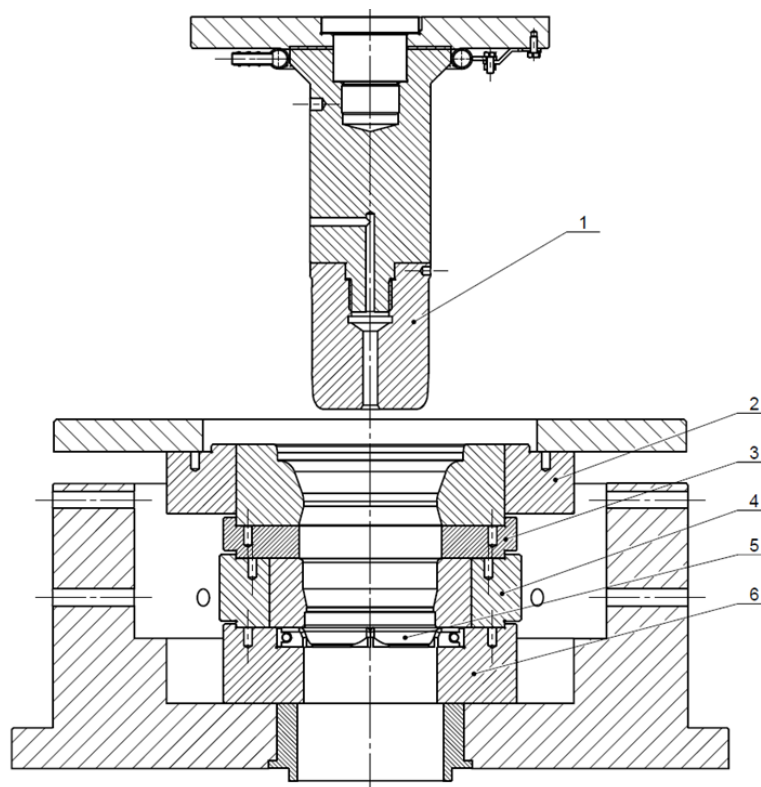


Fig. 5 Technical drawing of the tool assembly

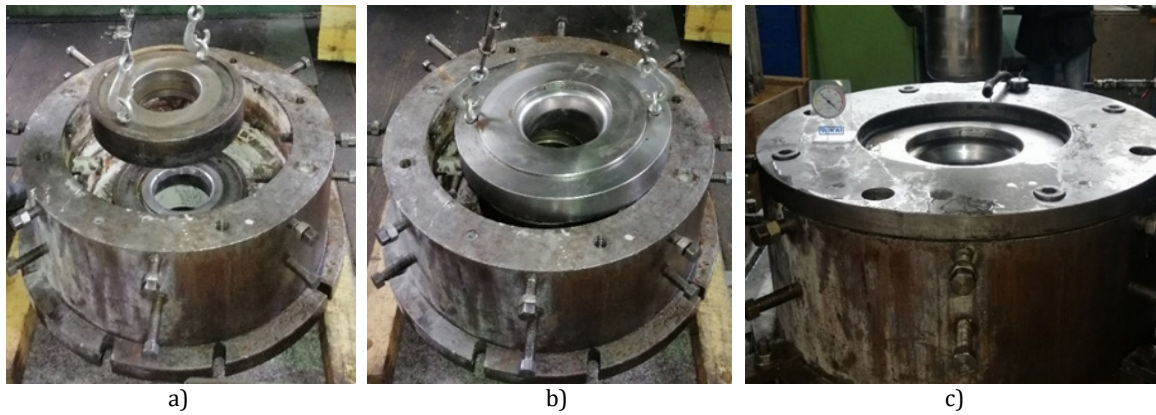


Fig. 6 Assembly of the lower deep drawing tool with reinforcing rings: a) lower die b) upper die c) assembly



Fig. 7 Cutting the blank by the water-jet process (left) and the blank after chemical treatments (right)

4.3 Measuring the initial temperatures in industrial experiments

In thermo-mechanical FE analysis using the finite element method, it is crucial to provide high-quality and reliable input data, not only for plastic flow of materials, but also for thermodynamic processes. *Fluke Ti400 Infrared Thermal Imager* was used to measure the temperatures of the blank, the environment and the tool in the industrial process, with prior adjustment of the emissivity coefficient. The initial temperature of the blank was 20.7°C, the tool 18.2°C and the environment 18.8°C. However, as the blank before deep drawing was immersed in a heated soap emulsion, in order to reduce contact friction during the process, the measured temperature of the blank placed in the tool was 30.7°C. Fig. 8 shows the registered temperatures.

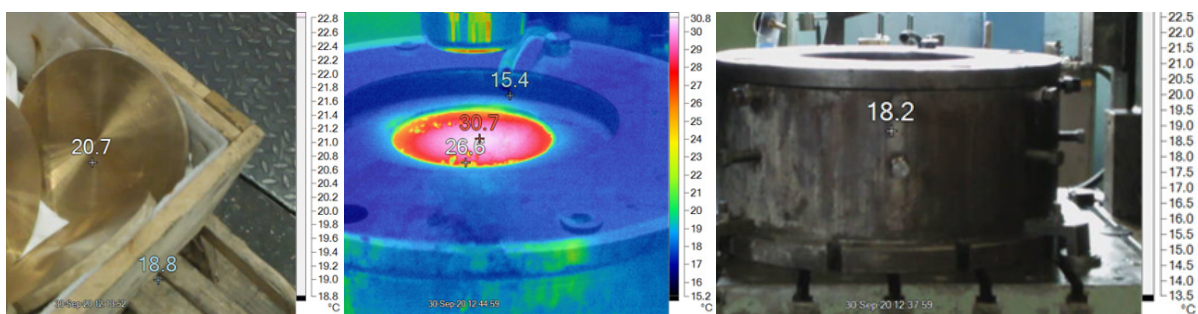


Fig. 8 Initial temperatures in industrial process (blank before and after soap emulsion, environment and tool)

4.4 Numerical model setup

The numerical model of the deep drawing process was set according to identical conditions in the industrial experiment using *Simufact.forming* software and its solver based on the finite element method. 3D CAD models of tool components in *IGES* format were applied, the assembly of which is shown in Fig. 5. The geometry of the reinforcing rings is simplified, and instead of the intermediate plate, the boundary conditions for fixing the nodes along the horizontal plane in the direction of the *y* axis were defined. Given that it is necessary to achieve absolute similarity between the

industrial and numerical process model, firstly related to the assembly of the upper and lower dies and their reinforcement with rings, the numerical model first included simulations of reinforcing and prestressing of the upper and lower dies. Then, the prestressed tools with deformation history were used for a coupled simulation of the deep drawing process, where the stress, strain and temperature fields in the workpiece and in the tools were monitored simultaneously.

The simulation of the prestressing of the dies by the reinforcing rings was carried out with a virtual overlap of 1 mm as one of the boundary conditions, while other contact conditions between the tool components and blank were defined using the *Contact table*. Instead of the properties and flow curves of the tool steel from which the die and ring were made, materials from the *Simufact material database* used for heat-treated tool steels were applied, for the die material HM560 and for the ring material ASP. The dies and rings were considered deformable while the base plate was considered rigid. The contact friction conditions between the dies and the rings were defined as dry with a friction coefficient of 0.4. Since the process is axisymmetric, a 2D numerical analysis was applied. A *Quadtree* mesh of two-dimensional finite elements with an element size of 1 mm was created on deformable tools.

For the numerical simulation of the deep drawing process, a cylindrical blank with identical dimensions as in the industrial process was modelled, on which a *Quadtree* mesh of two-dimensional finite elements with an element size of 1.3 mm was prepared. Due to significant deformations of the material in the areas of the calibrating bands in both dies, and thus greater distortion of the FE mesh, two refinement boxes were defined in which the number of FE elements was doubled during the simulation and the mesh was regenerated automatically (Fig. 9a).

The material behaviour was described by the flow curve according to Eq. 3 with coefficients' values $C=779.69$ and $n=0.52886$, for rolling direction 0° . The workpiece does not represent a finished part, so it was not necessary to take into consideration the anisotropy of the material when analysing the stresses and temperatures in the workpiece and the tool in the axisymmetric process. For the forming process, the contact friction conditions were described by Coulomb's law, where the given value of the friction coefficient was 0.12. The initial temperatures of the tool components, the blank and the environment were as registered and described in the previous section. The working stroke of the press was set at 300 mm. The numerical model of the deep drawing process is shown in Fig. 9b.

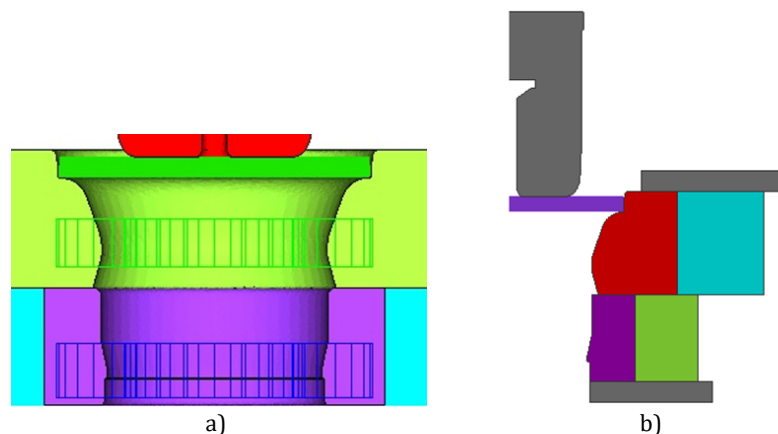


Fig. 9 Position of the refinement boxes (a) and 2D numerical model of the deep drawing process (b)

5. Results and discussion

Within this section, the results of all three numerical simulations are presented, both the assembly and prestressing of the upper and lower dies, as well as the coupled simulation of the deep drawing process with deformable tools. The analysis and discussion of the results will focus on stress fields in the workpiece and tools during the process and after workpiece release. Finally, a comparative analysis of measured temperatures in the industrial process and numerically estimated temperature fields in the workpiece and the tools will be presented and discussed.

The radial stress distributions obtained in the simulations of mounting the reinforcing rings on the upper and lower dies, with a virtual overlap of 1 mm, are shown in Fig. 10. It is noticeable that the maximum compressive radial stress in the upper die is -385 MPa and in the lower die -639 MPa. The fields of maximum stresses are located in the overlapping zones of the dies and the rings and mainly in the central part of the dies. The obtained prestress values in the upper die are lower than recommended (600-700 MPa), which is caused by the small value of the total overlap of 1 mm in relation to the outer diameter of the matrix of ϕ 330 mm.

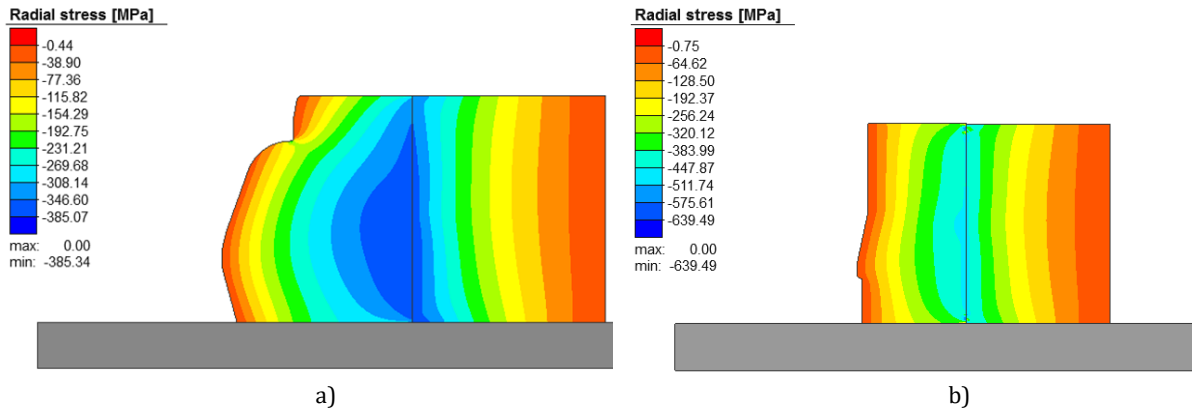


Fig. 10 Radial stress distribution of the reinforced upper die (a), lower die (b) in the assembly phase (1 mm overlap)

The recommendation for the overlap value, when installing reinforcing ring, is 0.1 mm for every 25 mm of a diameter. For the lower die, the overlap is adequate considering that the outer diameter of is ϕ 240 mm, so the obtained prestress is approximately the same as the recommended one. For the upper die, it is necessary to implement the overlap of 1.32 mm. Therefore, another simulation of the reinforcing upper die with a new value of the virtual overlap was performed. Fig. 11 shows the distribution of radial stresses in the upper die and ring with a larger overlap, where higher maximum values of radial stresses were obtained in the amount of 587 MPa. In this way, the optimal values of die reinforcement and prestressing can be determined by tool assembly finite element numerical simulation. Arbitrarily chosen values of the overlap between the ring and the die and thus the prestresses can have negative consequences on the tool life, as well as on the dimensional accuracy of the workpiece.

Simufact.forming software allows the transfer of deformation history from operation to operation. Thus, the strain and stress fields in the deformable tools (upper and lower dies and both rings) from the previous simulations of the reinforcing dies were imported into the next simulation of the deep drawing process. In this numerical simulation, the approach of simultaneous complete FE analysis of workpiece forming and elastic deformation in deformable tools was applied. Only in this way reliable results comparable to the industrial process can be obtained.

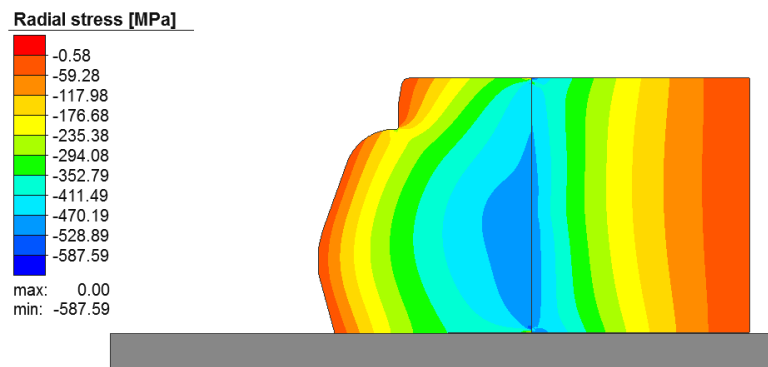


Fig. 11 Radial stress distribution of the reinforced upper die (1.32 mm overlap)

Fig. 12 shows the radial stress distributions in the workpiece and tools in the selected stages of the process. The first stage shown at 19.5 % of the puller stroke was chosen because there is a maximum contact zone between the workpiece and the tractrix profile of the upper die. It is evident that the compressive radial stresses in the upper die have decreased to the value of -400 MPa and the appearance of smaller tensile radial stresses up to 100 MPa (Fig. 12a). The state of the radial stresses in the lower die is almost unchanged. In the workpiece, maximum tensile radial stress was at the bottom of the vessel, the values of which range from 200-280 MPa.

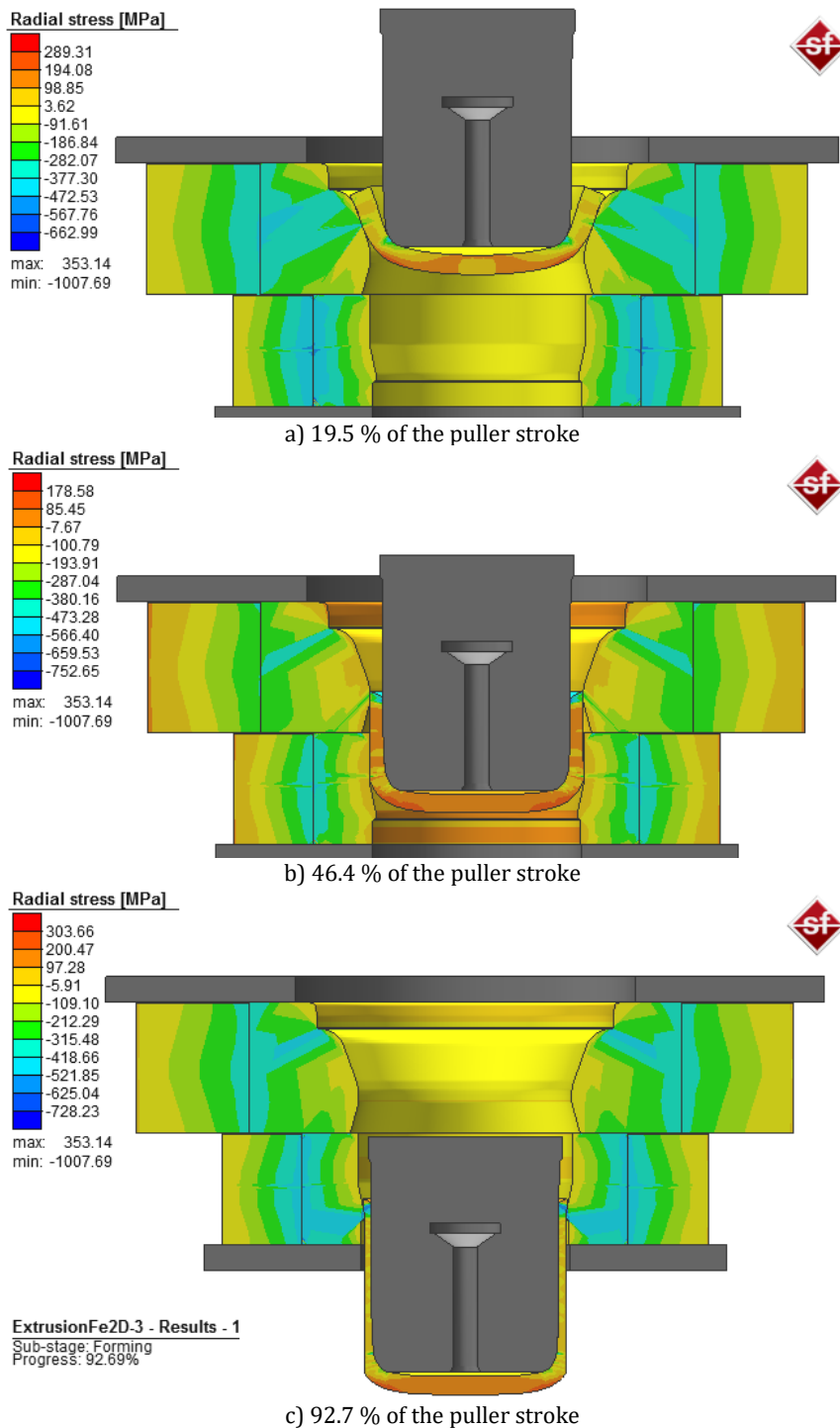


Fig. 12 Radial stress distributions in the workpiece and tools during coupled simulation of the deep drawing process

The second characteristic phase of the process (46.4 % of the stroke, Fig. 12b) includes the moment immediately before the exit of the workpiece from the calibrating zone of the upper die. It is noted that the maximum compressive radial stresses of -752 MPa are in the deformation zone of the workpiece at the very top, and that the upper part of the matrix along the tractrix curve is almost entirely relaxed. On the other hand, there is a decrease in tensile radial stresses at the bottom of the piece up to 178 MPa. In addition to the change in the shape of the workpiece, there is a small decrease in the thickness of the vessel wall.

In the last third selected phase of the process (92.7 % of the stroke), when the top of the workpiece is in contact with the calibrating zone of the lower die, maximum radial compressive stresses at the top of the vessel wall (-728 MPa) occur again. In contrast to passing material through the upper die, when the deep drawing process takes place, ironing is performed here, so the deformation of the workpiece in wall thickness is significant. This is transferred to the calibrating part of the die, but there is no large drop in compressive radial stresses (they amount to -520 MPa) because the ironing forces act differently on the die in the radial direction. There are no critical tensile stresses in the lower die and ring. However, the tensile radial stresses at the bottom of the vessel increased up to 300 MPa.

In order to more deeply analyse the thermo-mechanical processes in the numerical model of the deep drawing process, and compare them with industrial temperature measurements using infrared thermography, apart from the determination of the stresses in the workpiece and the tools, it is necessary to analyse the numerically estimated temperature fields.

Fig. 13a shows the temperature fields in the upper die obtained by the infrared thermal imager immediately after the process. For comparison, Fig. 13b shows the distribution of temperature fields obtained by FE numerical simulation. The contact friction between the tool and the workpiece, as well as the plastic deformation of the workpiece affected the temperature increase in the workpiece and the transfer of part of the temperature to the tools. The development of elevated temperatures in the upper die occurs in the output path of the tractrix curve in the range from 30°C to 77.1°C, with the maximum temperature being registered in the die calibration zone. The numerically estimated temperature field has a good coincidence as the maximum temperature of 77.9°C occurs on the same part of the die surface, and the range of the temperature field starts from 30°C. The advantage of the numerical simulation is that the temperature fields are calculated for the entire volume of the die, so the temperature fields can be analysed in any section and at any position, as opposed to physical measurements.

The measured temperatures in the lower die immediately after the process are shown in Fig. 14a, and the comparative distribution of the numerically estimated temperature fields is given in Fig. 14b. By analysing the temperatures on the lower die, it is noticed that there was a development of slightly higher temperatures in the range from 30°C to 90°C and a wider temperature field along the depth of the die, in the calibration zone where contact friction has the greatest influence. The reason for this is a longer contact with the workpiece, since the entire length of the vessel wall was in contact with this die during the ironing process. This was also confirmed by numerical simulation with the same area of elevated temperatures, with the fact that the maximum estimated temperature is slightly higher and amounts to 91.9°C.

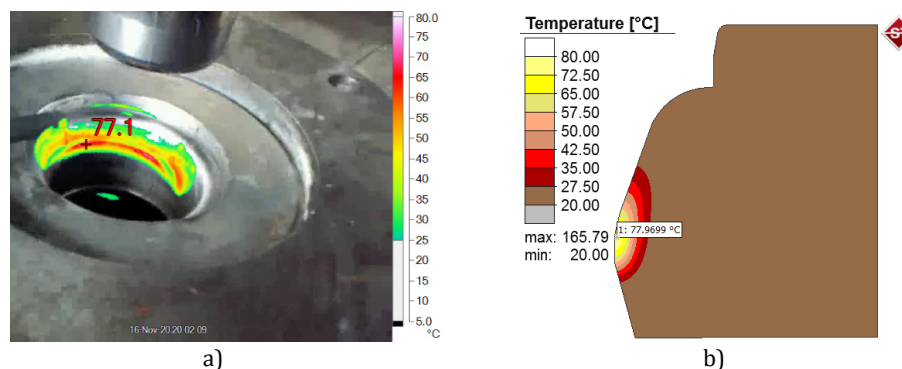


Fig. 13 Comparison of the measured temperatures in the upper die (a) and numerically estimated (b)

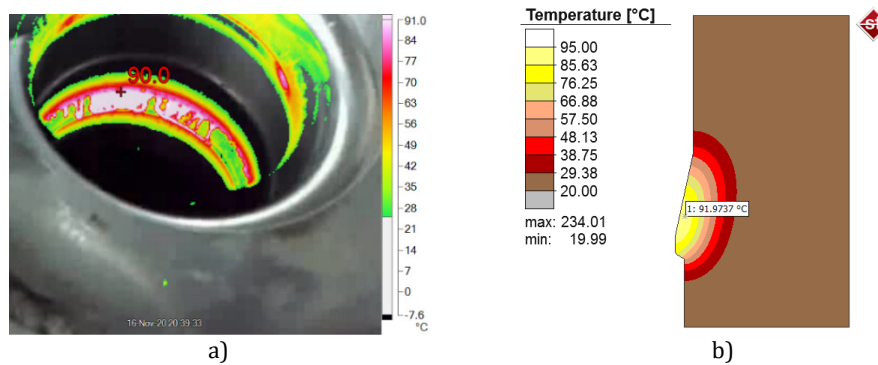


Fig. 14 Comparison of the measured temperatures in the lower die (a) and numerically estimated (b)

Based on the analysed temperature fields, it can be concluded that the tool temperature values are relatively small. However, it must be emphasized that the simulation includes the deep drawing process of one workpiece, which means that it is expected that, during the longer exploitation of the tool, it will heat up more. Therefore, it is necessary to apply a cooling agent, in order to prevent overheating of the tool, which can lead to a change in the characteristics of the material, the appearance of temperature cracks and a short tool life.

Immediately after removing the workpiece from the tool in the industrial process, the temperatures were measured in the same way as for the tool. Arbitrary points on the thermal image were selected, as shown in Fig. 15a, in order to compare the values with the numerical temperature fields in the workpiece, shown in Fig. 15b. The legend of the industrial temperature measurements shows temperatures up to a maximum of 240°C, while the legend of the numerical model shows a value up to 233°C. In similar referent points, there is a satisfactory match between the measured and estimated temperatures in the workpiece. In the zone near the top of the vessel, the measured temperature was 182.1°C, and the numerically estimated temperature was 183.2°C. The lowest temperature values were registered and evaluated at the bottom of the vessel and range from 20°C to 45°C, because there was the least plastic deformation and there was no contact with the tool and thus no influence of contact friction.

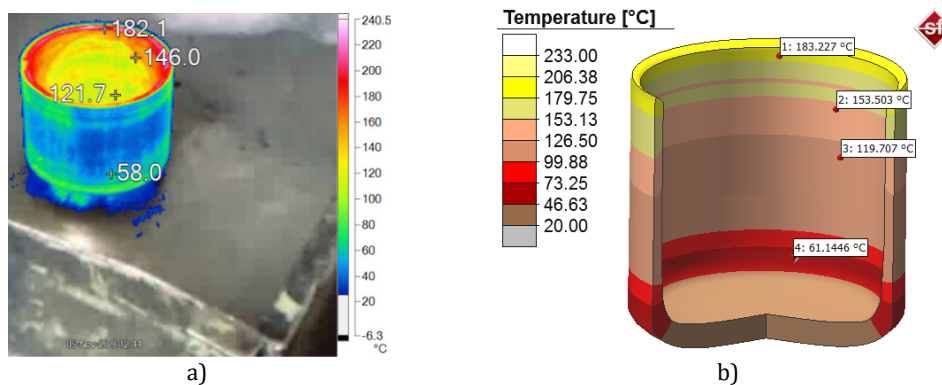


Fig. 15 Comparison of the measured temperatures in the workpiece (a) and numerically estimated (b)

6. Conclusion

Deep drawing of thick sheets using tractrix die is a widely used technology in large-scale industrial production for the initial operations of obtaining special vessels, which require a significant difference in the thickness of the bottom and vessel walls, as well as special mechanical properties of the material in the characteristic sections. Reliable technology cannot be developed only through industrial trials, so the application of an experimental-numerical approach is more than necessary.

The aim of this study was to develop, through this approach, a reliable numerical thermo-mechanical model of the process for obtaining a silicon brass vessel in successive deep drawing and

ironing operations. It is known that the forming of thick sheets, as bulk deformation, is accompanied by an increase in temperature in the workpiece and transfer to the tools, so a comparative analysis of measured and numerically estimated temperatures in the workpiece and tools was used to validate the numerical model based on the finite element method.

The reliability of the numerical model is conditioned by high-quality input data for a set of FE simulations, which refer to models of material behaviour, contact friction and thermodynamic processes. The selection of adequate finite elements for the workpiece and deformable tools with constitutive relations that correspond to the physics of the process, then the parameters of the FE analysis and the conditions of automatic mesh generation and remeshing are crucial for precise and convergent numerical solutions. The paper provides a detailed description of the above-mentioned requirements, among other things, through the results of experimental testing of silicon brass material as well as temperature measurements before and after the deep drawing industrial process using an infrared thermal imager.

The numerical model of the process included the assembly phase of the upper and lower dies with reinforcing rings with an overlap in order to prestress and protect the dies. Adequate selection of overlaps and realised prestressing of the die is of essential importance for the accuracy of the parts, the reduction of the total tensile radial stresses, the dynamic cyclic loading of the die and its lifetime. The present study showed that identical overlaps cannot be applied in the upper and lower dies, so through repeated simulation with a changed value of the overlap of 1.3 mm, optimal prestresses were obtained.

Stress and temperature fields were analysed in the combined FE analysis of deep drawing and ironing processes with deformable prestressed tools. Numerical simulations, which are mostly performed with rigid tools, with or without the heat transfer option, do not provide a realistic numerical model of the industrial process. Therefore, the study established an expensive and time-consuming coupled FE numerical thermo-mechanical model through simultaneous plastic deformation of the workpiece, elastic deformation of the tools, heat generation due to plastic deformation and contact friction, heat loss due to transfer to tools and emissions to the environment. A comparative analysis of the temperature fields obtained by industrial measurements and numerical simulations showed a good agreement, thus validating the numerical model.

The establishment of a reliable and precise numerical model is the basis for investigating the influence of process parameters on other output indicators (strains, strain rates, forming load) and for making the desired modifications to the optimization of industrial technology. Temperature fields under other processing conditions and their influence on microstructure changes can be investigated with appropriate input data describing phase transformations in the material, which is the subject of future research using an integrated experimental-numerical approach.

Acknowledgement

The part of this research is supported by the Ministry of Science, Technological Development and Innovation, Republic of Serbia, Grant TR34002 and Grant TR32036, as well as Erasmus Grant 2020-1-ES01-KA203-081978

References

- [1] Park, G., Park, R., Kwak, H., Kim, C. (2021). Design of a combined redrawing-ironing process to manufacture a CNG pressure vessel liner, *Applied Sciences*, Vol. 11, No. 18, Article No. 8295, doi: [10.3390/app11188295](https://doi.org/10.3390/app11188295).
- [2] Pernis, R., Barényi, I., Kasala, J., Ličková, M. (2015). Evaluation of limiting drawing ration (LDR) in deep drawing process, *Acta Metallurgica Slovaca*, Vol. 21, No. 4, 258-268, doi: [10.12776/ams.v21i4.642](https://doi.org/10.12776/ams.v21i4.642).
- [3] Kesharwani, R.K., Basak, S., Panda, S.K., Pal, S.K. (2017). Improvement in limiting drawing ratio of aluminum tailored friction stir welded blanks using modified conical tractrix die, *Journal of Manufacturing Processes*, Vol. 28, Part 1, 137-155, doi: [10.1016/j.jmapro.2017.06.002](https://doi.org/10.1016/j.jmapro.2017.06.002).
- [4] Loganathan, C., Narayanasamy, R. (2005). Effect of mechanical properties on the wrinkling behaviour of three different commercially pure aluminium grades when drawn through conical and tractrix dies, *Materials Science and Engineering: A*, Vol. 406, No. 1-2, 229-253, doi: [10.1016/j.msea.2005.06.037](https://doi.org/10.1016/j.msea.2005.06.037).
- [5] Park, G.-Y., Kwak, H.-S., Jang, H.-S., Kim, C. (2022). Deep drawing process using a tractrix die for manufacturing liners for a CNG high-pressure vessel (Type II), *Chinese Journal of Mechanical Engineering*, Vol. 35, No. 15, 1-12, doi: [10.1186/s10033-022-00681-9](https://doi.org/10.1186/s10033-022-00681-9).

- [6] Reddy, A.C.S., Rajesham, S., Reddy, P.R. (2015). Experimental and simulation study on the warm deep drawing of AZ31 alloy, *Advances in Production Engineering & Management*, Vol. 10, No. 3, 153-161, doi: [10.14743/apem2015.3.199](https://doi.org/10.14743/apem2015.3.199).
- [7] Mandic, V. (2012). *Physical and numerical modelling of metal forming processes (in Serbian)*, Faculty of Mechanical Engineering, Kragujevac, Serbia.
- [8] Hrnjica, B., Behrem, Š. (2022). A new multi-objective optimization approach for process parameters optimization during numerical simulation of quenching steel parts, *Advances in Production Engineering & Management*, Vol. 17, No. 1, 16-32, doi: [10.14743/apem2022.1.418](https://doi.org/10.14743/apem2022.1.418).
- [9] Mandić, V., Čosić, P. (2011). Integrated product and process development in collaborative virtual engineering environment, *Tehnički Vjesnik – Technical Gazette*, Vol. 18, No. 3, 369-378.
- [10] Gusel, L., Boskovic, V., Domitner, J., Ficko, M., Brezocnik, M. (2018). Genetic programming method for modelling of cup height in deep drawing process, *Advances in Production Engineering & Management*, Vol. 13, No. 3, 358-365, doi: [10.14743/apem2018.3.296](https://doi.org/10.14743/apem2018.3.296).
- [11] Adamovic, D., Mandic, V., Jurkovic, Z., Grizelj, B., Stefanovic, M., Marinkovic, T., Aleksandrovic, S. (2010). An experimental modelling and numerical FE analysis of steel-strip ironing process, *Tehnički Vjesnik – Technical Gazette*, Vol. 17, No. 4, 435-444.
- [12] Elplacy, F., Samuel, M. Mostafa, R. (2022), Modelling and simulation of hot direct extrusion process for optimal product characteristics: Single and multi-response optimization approach, *Advances in Production Engineering & Management*, Vol. 17, No. 1, 33-44, doi: [10.14743/apem2022.1.419](https://doi.org/10.14743/apem2022.1.419).
- [13] Tomáš, M., Evin, E., Kepič, J., Hudák, J. (2019). Physical modelling and numerical simulation of the deep drawing process of a box-shaped product focused on material limits determination, *Metals*, Vol. 9, No. 10, 1058-1074, doi: [10.3390/met9101058](https://doi.org/10.3390/met9101058).
- [14] Volk, M., Nardin, B., Dolsak, B. (2014). Determining the optimal area-dependent blank holder forces in deep drawing using the response surface method, *Advances in Production Engineering & Management*, Vol. 9, No. 2, 71-82, doi: [10.14743/apem2014.2.177](https://doi.org/10.14743/apem2014.2.177).
- [15] Musafija, B. (1979). *Metal forming by plastic deformation*, Svetlost, Sarajevo, Bosnia and Herzegovina.
- [16] Romhanji, E., Milenković, V., Drobnjak, D. (1992). The grain size and alloying influence on the strain hardening of polycrystalline α -brasses, *International Journal of Materials Research*, Vol. 83, No. 2, 110-114, doi: [10.1515/ijmr-1992-830208](https://doi.org/10.1515/ijmr-1992-830208).
- [17] Hexagon. HxGN virtual manufacturing, from <https://www.simufact.com/simufactforming-forming-simulation.html>, accessed November 25, 2023.
- [18] Popović, M., Mandić, V., Delić, M., Pavićević, V. (2021). Experimental-numerical analysis of hot forging process with monitoring of heat effects, In: Karabegović, I. (ed.), *New technologies, development and application IV, NT 2021, Lecture notes in networks and systems*, Springer, Cham, Switzerland, 341-349, doi: [10.1007/978-3-030-75275-0_38](https://doi.org/10.1007/978-3-030-75275-0_38).
- [19] Hou, F., Zhang, Y., Zhou, Y., Zhang, M., Lv, V., Wu, J. (2022). Review on infrared imaging technology, *Sustainability*, Vol. 14, No. 18, Article No. 11161, doi: [10.3390/su141811161](https://doi.org/10.3390/su141811161).
- [20] Usamentiaga, R., Venegas, P., Guerediaga, J., Vega, L., Molleda, J., Bulnes, F. (2014). Infrared thermography for temperature measurement and non-destructive testing, *Sensors*, Vol. 14, No. 7, 12305-12348, doi: [10.3390/s140712305](https://doi.org/10.3390/s140712305).
- [21] Fluke. Ti200, Ti300, Ti400, Thermal imagers, Users manual, from <https://www.fluke-direct.com/pdfs/cache/www.fluke-direct.com/ti200-60hz/manual/ti200-60hz-manual.pdf>, accessed December 15, 2023.

Calendar of events

- International Conference on Digital Manufacturing Systems ICDMS, January 7-8, 2024, Tokyo, Japan.
- International Conference on Manufacturing and Intelligent Machining ICMIM, January 14-15, 2024, Zurich, Switzerland.
- 18th International Conference on Advanced Manufacturing Engineering and Technologies, January 15-16, 2024, Montevideo, Uruguay.
- International Conference on Cyber Manufacturing Systems ICCMS, January 18-19, 2024, Sydney, Australia.
- International Conference on Manufacturing and Industrial Technologies ICMIT, January 25-27, 2024, Budapest, Hungary.
- International Conference on Mechatronics and Manufacturing ICMM, February 2-4, 2024, Penang, Malaysia.
- International Conference on Manufacturing and Optimization ICMO, February 25-26, 2024, Buenos Aires, Argentina.
- International Conference on Mechanical and Intelligent Manufacturing Technologies ICMIMT, March 7-9, 2024, Cape Town, South Africa.
- 2024 Annual Modeling and Simulation Conference (ANNSIM 2024), May 20-23, 2024, Washington D.C., USA.
- North American manufacturing research conference (NAMRC) 52, June 17-21, 2024, Knoxville, TN, USA.
- 18th International Conference on Industrial and Manufacturing Systems Engineering, August 9-10, 2024, Lagos, Nigeria.

This page intentionally left blank.

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. The content from published paper in the *APEM journal* may be used under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0). For most up-to-date information 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

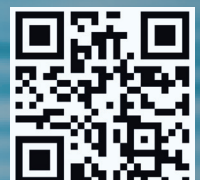
Chair of Production Engineering (CPE)
University of Maribor
APEM homepage: apem-journal.org

Volume 19 | Number 1 | March 2024 | pp 1-152

Contents

Scope and topics	4
An improved multi-objective firefly algorithm for integrated scheduling approach in manufacturing and assembly considering time-sharing step tariff Xu, E.B.; Zou, F.F.; Shan, P.P.; Wang, Z.Y.; Shi, B.X.	5
Research on recovery strategies of supply chain network under disruption propagation using memetic algorithm Li, Z.Y.; Zhao, P.X.; Wang, C.L.; Mi, Y.Z.	21
Human-robot collaboration assembly line balancing considering cross-station tasks and the carbon emissions Li, Y.C.; Wang, X.	31
Predicting the deep drawing process of TRIP steel grades using multilayer perceptron artificial neural networks Sevšek, L.; Vilkovský, S.; Majerniková, J.; Pepelnjak, T.	46
The investigation of production variance in a module-based assembly system: A Markovian Arrival Process approach Zhong, C.W.; Zhang, H.; Zhang, Z.W.; Wu, Z.Q.; Lu, M.L.	65
A comparative study of machine learning regression models for production systems condition monitoring Jankovič, D.; Šimic, M.; Herakovič, N.	78
Evolutionary game analysis of green innovation in E-commerce closed-loop supply chain WEEE recycling Ma, R.M.; Yao, L.F.; Wang, H.	93
FDM process parameter selection by hybrid MCDM approach for flexural and compression strength maximization Begic-Hajdarevic, D.; Klančnik, S.; Muhamedagić, K.; Čekić, A.; Cohodar Husic, M.; Ficko, M.; Gusel, L.	108
Integrated optimization of line planning and timetabling on high-speed railway network considering cross-line operation Wang, R.X.; Nie, L.; Fang, W.; Ren, H.Q.; Tan, Y.Y.	117
Experimental and numerical investigation of the deep drawing process using a tractrix die – An industrial case study focused on stress and temperature analysis Mandic, V.; Milosavljevic, Dj.; Jurkovic, Z.; Adamovic, D.	133
Calendar of events	149
Notes for contributors	151

Published by CPE, University of Maribor



apem-journal.org